

Algorithmic Polynomial Freiman-Ruzsa Theorems

Srinivasan Arunachalam* Davi Castro-Silva† Arkopal Dutt* Tom Gur†

September 2, 2025

Abstract

We prove algorithmic versions of the polynomial Freiman-Ruzsa theorem of Gowers, Green, Manners, and Tao (Annals of Mathematics, 2025) in additive combinatorics. In particular, we give classical and quantum polynomial-time algorithms that, for $A \subseteq \mathbb{F}_2^n$ with doubling constant K , learn an explicit description of a subspace $V \subseteq \mathbb{F}_2^n$ of size $|V| \leq |A|$ such that A can be covered by K^C translates of V , for a universal constant $C > 1$.

1 Introduction

The Freiman-Ruzsa theorem [Fre87, Ruz99] is a cornerstone of additive combinatorics with diverse applications to theoretical computer science (cf. [Lov15]). Loosely speaking, the theorem shows that sets exhibiting approximate combinatorial subgroup behaviour must be algebraically structured. To make this precise, recall that a set A has doubling constant K if $|A + A| \leq K|A|$, where $A + A = \{a + a' ; a, a' \in A\}$. Note that A has doubling constant 1 if and only if it is a subgroup or a coset of a subgroup, and in turn, the doubling constant of A can be thought of as a combinatorial measure of the approximate subgroup behaviour of sets. In this paper, we focus on subsets of \mathbb{F}_2^n . In this setting, the Freiman-Ruzsa theorem states that sets $A \subseteq \mathbb{F}_2^n$ with $|A + A| \leq K|A|$ is covered by $\exp(K)$ translates of a subspace $V \subset \mathbb{F}_2^n$ of size $|V| \leq |A|$.

Marton conjectured that the aforementioned dependency in K can be improved to a polynomial, in what became widely known as the Polynomial Freiman-Ruzsa (PFR) conjecture. Over a decade later, Sanders proved a quasipolynomial Bogolyubov-Ruzsa theorem, which implies a version of the Freiman-Ruzsa theorem with quasipolynomial dependency on the doubling constant K . In a recent breakthrough, the PFR conjecture was proved by Gowers, Green, Manners, and Tao.

Theorem 1.1 (Combinatorial PFR theorem [GGMT25]). *There exists a polynomial $P_0 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following holds. For any $n \geq 1$, if $A \subseteq \mathbb{F}_2^n$ satisfies $|A + A| \leq K|A|$, then A is covered by at most $P_0(K)$ translates of a subspace $V \subset \mathbb{F}_2^n$ of size $|V| \leq |A|$.*

A key reason for the importance of the PFR theorem in additive combinatorics is that it provides means to transition from a combinatorial notion of approximate subgroup structure, captured by constant doubling, to an algebraic notion, captured by a bounded subspace-cover, at only a polynomial cost.

*IBM Quantum.

†University of Cambridge.

DCS is supported by ESPRC Robust and Reliable Quantum Computing Grant.

TG is supported by ERC Starting Grant 101163189 and UKRI Future Leaders Fellowship MR/X023583/1.

1.1 Algorithmic PFR

The PFR theorem (and the closely-related quasi-polynomial Bogolyubov-Ruzsa theorem) also provide powerful tools that found diverse applications to theoretical computer science, including linearity testing of maps $f: \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ [Sam07], constructions of two-source extractors from affine extractors [ZBS11], communication complexity lower bounds [BSLRZ14], super-polynomial lower bounds on locally decodable codes [BDL13], constructions of non-malleable codes [ADL14], higher-order Goldreich-Levin theorems [TW14, KLT23, BC25], sparsification algorithms for 1-in-3-SAT [BNOŽ25], quantum proofs for classical theorems [DdW09], quantum and classical worst-case to average-case reductions [AGGS22, AGG+24], tolerant quantum testing of stabilizer states [AD25b], quantum learning of structured stabilizer decompositions [AD25a], and beyond.

However, when considering applications of the PFR theorem and similar tools to theoretical computer science as above, it is often necessary or desirable to have an efficient algorithmic statement, where an explicit description of the subspace can be learned efficiently, as opposed to an existential combinatorial statement. Indeed, the naive algorithm that extracts the subspace runs in time $O(2^n)$. Fortunately, for the quasi-polynomial Bogolyubov-Ruzsa theorem, Ben-Sasson, Ron-Zewi, Tulsiani, and Wolf [BSRZTW14] showed an algorithmic version, which extracts a subspace in time $O(n^3 \log n)$.

The above motivates a natural question that arose after the resolution of the PFR conjecture [GGMT25]. Namely, now that we know that a subset A of constant doubling can be covered by at most a polynomial number of translates of a subspace $H \subset \mathbb{F}_2^n$ of size $|H| \leq |A|$, *can we learn the subspace H efficiently?* We refer to this as the “algorithmic PFR question.” Our main contribution answers this question by proving an algorithmic version of the polynomial Freiman-Ruzsa theorem, where the covering subspace can be learned explicitly in $\text{poly}(n)$ -time. In the following, a query to a set $A \subseteq \mathbb{F}_2^n$ is an evaluation of the characteristic function $\mathbf{1}_A(x)$ for a chosen $x \in \mathbb{F}_2^n$, and a random sample from A is a uniformly chosen element $a \in A$.

Theorem 1.2. (*Algorithmic PFR*) *There exists a polynomial $P_1: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following holds. Let $A \subseteq \mathbb{F}_2^n$ satisfy $|A + A| \leq K|A|$ for a doubling constant $K \geq 1$. There is an $\tilde{O}(n^4)$ -time randomized algorithm that uses $O(\log|A|)$ random samples and $\tilde{O}(\log^2|A|)$ queries to A which, with probability at least $2/3$, returns a subspace $V \leq \mathbb{F}_2^n$ of size at most $|A|$ such that A can be covered by $P_1(K)$ translates of V .*

We remark that the probability of success $2/3$ is arbitrarily chosen and can be amplified via standard error-reduction techniques. The algorithm returns the subspace V specified by an explicit basis. As typically viewed in additive combinatorics, the doubling constant K is a constant independent of n (as constant doubling implies structure), and in turn our asymptotic notation suppresses factors of K .

Quantum algorithms. En route to obtaining our algorithmic PFR theorem, we obtain quantum algorithms with query complexity (which we then dequantize; see the techniques section) and time complexity¹ that is a factor- n lesser than the classical algorithms. Namely, we show the following theorem.

Theorem 1.3. (*Quantum Algorithmic PFR*) *There exists a polynomial $P_1: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following holds. Let $A \subseteq \mathbb{F}_2^n$ satisfy $|A + A| \leq K|A|$ for a doubling constant $K \geq 1$. There is*

¹In the context of quantum algorithms, by time we mean the total number of single and two-qubit quantum gates used in the quantum algorithm.

an $O(n^3)$ -time quantum algorithm that uses $O(\log|A|)$ random samples and quantum queries to A which, with probability at least $2/3$, returns a subspace $V \leq \mathbb{F}_2^n$ of size at most $|A|$ such that A can be covered by $P_1(K)$ translates of V .

Optimality. We complement our algorithms by also showing that both of our classical and quantum algorithms are asymptotically optimal in terms of the *query complexity* dependence in n , up to a logarithmic factor. In particular, we show that $\Omega(n^2)$ queries to the set A are necessary for classical algorithms in order to output the subspace V . Similarly, we show that $\Omega(n)$ quantum queries to the set A are necessary. Our lower bounds are proven via a simple information-theoretic argument. Finally, we note that random samples are necessary to hit A in the case it is sparse, and information-theoretically, it is necessary to obtain $\Omega(\log|A|)$ samples from A in order to hit at least a basis for A . See Section 5 for the lower bounds.

1.2 Homomorphism testing and structure-vs-randomness decomposition

A key reason for the power and centrality of the PFR theorem is its applications to deriving strong structural theorems regarding homomorphism testing and structure-vs-randomness decomposition. However, while the PFR theorem is known to be equivalent to the aforementioned structural theorems [Gre04a, Gre05], the equivalences are not trivially algorithmic. Nonetheless, we provide versions of these theorems that admit efficient algorithms.

The first theorem is concerned with local-to-global phenomena. Namely, it shows that if a map $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ satisfies a local affine-linear constraint with a significant probability, then it must be globally close, in fractional distance, to an affine-linear map, which can be efficiently learned. This statement is useful because it connects the PFR theorem to property testing and coding theory.

Theorem 1.4 (Homomorphism testing). *There exists a polynomial $P_2 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following holds. Let $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ satisfy*

$$\Pr_{x_1+x_2=x_3+x_4} [f(x_1) + f(x_2) = f(x_3) + f(x_4)] \geq 1/K.$$

Then, there is an affine-linear function $g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ such that $f(x) = g(x)$ for at least $2^m/P_2(K)$ values of $x \in \mathbb{F}_2^m$. Furthermore, there is an $\tilde{O}((m+n)^3)$ -time randomized algorithm that, with probability at least $2/3$, learns a concise representation of g .

The second theorem also exhibits local-to-global structure, this time by showing that maps that are locally an approximate homomorphisms admit a structured decomposition, which can be efficiently learned.

Theorem 1.5 (Structured approximate homomorphism). *There exists a polynomial $P_3 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the following holds. Let $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ satisfy*

$$|\{f(x) + f(y) - f(x+y) : x, y \in \mathbb{F}_2^m\}| \leq K.$$

Then f may be written as $g + h$, where $g : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ is linear and $|\text{Im}(h)| \leq P_3(K)$. Furthermore, there is an $\tilde{O}((m+n)^3)$ -time randomized algorithm that, with probability at least $2/3$, learns a concise representation of g .

1.3 Technical overview

A natural approach for proving an algorithmic PFR theorem is to try to algorithmize each step in the proof of the PFR conjecture given in [GGMT25], which would in principle provide a result

similar to Theorem 1.2. Unfortunately, the aforementioned proof heavily relies on entropic methods that are non-algorithmic by nature and it is unclear whether such machinery can be transformed into efficient algorithms. To overcome this barrier, we take a detour through quantum algorithms.

Stabilizer learning and the Gowers U^3 -norm of quantum states. In a recent breakthrough, Chen, Gong, Ye, and Zhang provided an efficient quantum procedure to learn the closest stabilizer state to a given quantum state [CGYZ25]. The connection between this task and additive combinatorics was first noted by Arunachalam and Dutt [AD25b], who defined the Gowers U^3 -norm of an arbitrary n -qubit quantum state $|\psi\rangle = \sum_{x \in \mathbb{F}_2^n} f(x) |x\rangle$ (where $(f(x))_x$ is an ℓ_2 -unit vector) as

$$\|\psi\|_{U^3} = 2^{n/2} \left[\mathbb{E}_{x, h_1, h_2, h_3 \in \mathbb{F}_2^n} \prod_{\omega \in \mathbb{F}_2^3} C^{|\omega|} f(x + \omega \cdot h) \right]^{1/2^3}, \quad (1)$$

where $C^{|\omega|} f = \bar{f}$ if $\omega := \sum_{j \in [3]} \omega_j$ is odd and f otherwise. This is proportional to the Gowers U^3 -norm of the function f encoded in the amplitudes of the state $|\psi\rangle$, but normalized so that $\|\psi\|_{U^3} \leq \|f\|_{\ell_2} = 1$.

Arunachalam and Dutt showed that stabilizer states are the extremizers of the Gowers U^3 -norm over quantum states, and that a quantum state has non-negligible U^3 -norm if and only if it correlates with a stabilizer state. Moreover, using the PFR theorem (Theorem 1.1), they obtained a *polynomial Gowers inverse theorem* for quantum states: the U^3 -norm of a quantum state and its maximal correlation with a stabilizer state are polynomially related (see also [BvDH25, MT25]). As such, our high-level strategy is to use the stabilizer learning protocol in [CGYZ25] to obtain a quantum algorithmic version of the polynomial Gowers inverse theorem, which is known to also be equivalent to the PFR theorem due to work of Green and Tao [GT10] and of Lovett [Lov12]. We then arrive at an algorithmic result, albeit quantum, of a statement that is combinatorially equivalent to PFR.

However, as discussed in the previous section, it is non-trivial to make such combinatorial equivalences algorithmic. In this paper, we algorithmize a proof of equivalence between these two results (inspired by the proofs of Green-Tao and Lovett), thus allowing us to employ the stabilizer learning algorithm [CGYZ25] to obtain efficient *quantum* algorithms for the PFR theorem and the structural theorems stated in Section 1.2.

Classical algorithms via dequantization. After obtaining the quantum algorithms above, the last ingredient needed is a method to dequantize these algorithm so as to obtain efficient *classical* algorithms for PFR, as stated in Theorems 1.2, 1.4 and 1.5. Towards this end, we use in our arguments the machinery developed by Briët and Castro-Silva [BC25] to replace the quantum learning algorithm in [CGYZ25] by a classical algorithm that emulates it. This allows us to obtain analogous algorithmic results in the classical setting, at the expense of quadratically worse query complexity than in the quantum setting. This quantum-to-classical blow-up is inherent, and indeed, we prove it is necessary and essentially optimal (see Theorem 5.1).

Proof outline. With the strategy above in mind, we proceed to give a high-level outline of the proofs of Theorems 1.2 and 1.3, which build on the combinatorial arguments of Green and Tao [GT10] and Lovett [Lov12]. Our other algorithmic results follow via similar methods from these two theorems and the combinatorial arguments of Green and Ruzsa [Gre05].

Suppose we have sample and query access to a set $A \subseteq \mathbb{F}_2^n$ such that $|A + A| \leq K|A|$ for a doubling constant $K \geq 1$. We shall first need to localize A inside the space \mathbb{F}_2^n , which in applications can be much larger than A itself (indeed, this is the reason why sample access to A is necessary). We do this by first sampling $O(\log|A|)$ uniformly random elements from A and taking their linear span, which we denote by $U \leq \mathbb{F}_2^n$. While U might not contain all of the original set, using the fact that A has bounded doubling we can show that their intersection $A' := A \cap U$ will likely comprise at least half of the points in A (for a careful choice of parameters). We prove this in Lemma 3.1, which allows us to shift attention from A , which can be arbitrarily sparse inside \mathbb{F}_2^n , to the localization A' , which occupies a positive fraction (at least a 2^{-2K} -fraction) of the vector space U , by the Freiman-Ruzsa theorem. Note that A' will also have small doubling constant:

$$|A' + A'| \leq |A + A| \leq K|A| \leq 2K|A'|.$$

We next wish to obtain a “dense model” of the localized set A' ; that is, a set $S \subseteq \mathbb{F}_2^m$ that is “additively equivalent” to A' , as captured by the notion of Freiman isomorphisms, but which has density at least $1/K^C$ (for a universal constant $C > 1$) inside its ambient space \mathbb{F}_2^m . We show how to do this in Lemma 3.2, which states that with high probability a uniformly random linear map $\pi : U \rightarrow \mathbb{F}_2^m$, for $m = \log|4A'| + 10$, will be a Freiman isomorphism (i.e., isomorphism of additive quadruples) from A' to $S := \pi(A')$. Informally, this means they have the same additive structure, and hence such a dense model can be efficiently obtained by sampling.

Equipped with the localized dense model, we proceed to learn the covering subspace. Denote by $f : S \rightarrow A'$ the inverse of π when restricted to S . By the definition of Freiman isomorphisms, we have that

$$\forall a, b, c, d \in S : a + b = c + d \implies f(a) + f(b) = f(c) + f(d).$$

From this approximate linearity condition of f on S , we can show that the function

$$g(x, y) = \mathbf{1}_S(x)(-1)^{f(x) \cdot y}$$

is approximately quadratic.

In particular, following the approach of Green [Gre05] and Green-Tao [GT10], we show a slight strengthening of the homomorphism testing formulation of the PFR theorem (see Lemma 3.3), which we then proceed to algorithmize (see Lemma 3.4), relying on tools such as the quadratic Goldreich-Levin theorem (see Theorem 2.12). In more detail, we first prove that there exists a quadratic function $q : \mathbb{F}_2^{m+n} \rightarrow \mathbb{F}_2$ such that

$$\left| \mathbb{E}_{x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n} \mathbf{1}_S(x)(-1)^{f(x) \cdot y} (-1)^{q(x, y)} \right| \geq \frac{1}{P(K)}, \quad (2)$$

for some polynomial $P : \mathbb{R}_+ \rightarrow \mathbb{R}_+$.

Crucially, we can efficiently *learn* such a high-correlation quadratic function q . This is done relying on the stabilizer learning algorithm of Chen et al [CGYZ25] in the quantum setting, or its dequantization by Briët and Castro-Silva [BC25] in the classical setting. In order to use those theorems, however, we need to be able to efficiently query the function $g(x, y)$. This requires making queries to the set $S = \pi(A')$, and inverting the linear map π restricted to A' . We show how this can be done using a $O(n^3)$ time pre-processing step, and an extra cost of $O(n^2)$ time for each query to g . The total time and query complexities of our algorithms follow from this step.

From Eq. (2) and algebraic manipulations, we conclude that the homogeneous bilinear form

$$B(x, y) = q(x, y) - q(x, 0) - q(0, y) + q(0, 0)$$

correlates well with $g(x, y)$. Since we obtained an explicit description of q , we can compute a matrix $M \in \mathbb{F}_2^{n \times m}$ such that $B(x, y) = y^T Mx$. By a simple Fourier analytic argument, we can then conclude there is some $v \in \mathbb{F}_2^n$ such that

$$f(x) = Mx + v \quad \text{for at least } 2^m/P'(K) \text{ values } x \in S,$$

where $P': \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is another polynomial we obtain in our proof. This implies that the subspace

$$V = \{Mx : x \in \mathbb{F}_2^m\} \leq \mathbb{F}_2^n$$

satisfies

$$|A' \cap (v + V)| = |\text{Im}(f) \cap (v + V)| \geq 2^m/P'(K) \geq |A|/P'(K).$$

By an application of Ruzsa's covering lemma, we conclude that $P'(K)$ translates of V can cover A . By our choice of m and the fact that A' has bounded doubling, we deduce that V is covered by $2^m/|A| \leq 2^{12}K^4$ translates of any of its subspaces having size $|A|$, which concludes the high-level overview of the proof.

Future directions. As discussed above, a crucial component of the proof of our *classical* algorithmic PFR theorem relies on dequantizing a *quantum* algorithm for learning the closest stabilizer state to a given arbitrary quantum state. Finding further connections between quantum algorithms and additive combinatorics is an interesting direction for future research. One concrete question that is left open from our work is to improve the dependence on the doubling constant K . As is standard in additive combinatorics, K is assumed to be a constant, but for asymptotically growing K it is an interesting open problem whether there exists an algorithm with query and time complexities that scale polynomially in K .

2 Preliminaries

In this section, we state results that we will use in our algorithmic PFR theorems. We begin with standard definition in additive combinatorics.

2.1 Additive combinatorics

For a set $S \subseteq \mathbb{F}_2^n$ and $k \geq 1$, define the k -fold sumset as $kS = \{\sum_{i \in T} a_i : a_i \in S\}_{|T|=k}$. In particular, $S + S = \{a_1 + a_2 : a_1, a_2 \in S\}$. We define the *doubling constant* of S as the smallest integer K such that $|2S| \leq K|S|$. We denote the linear span of S by $\text{span}(S)$. An additive quadruple in a set S is $(x_1, x_2, x_3, x_4) \in S^4$ such that $x_1 + x_2 = x_3 + x_4$.

Definition 2.1. *The additive energy of a set S is the number of additive quadruples in S :*

$$E(S) := |\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4\}|.$$

Note that $E(S) \leq |S|^3$.

Definition 2.2 (Freiman homomorphism). *For a set $S \subseteq \mathbb{F}_2^n$, a function $\phi : S \rightarrow \mathbb{F}_2^m$ is a Freiman homomorphism if, for every additive quadruple $x_1, x_2, x_3, x_4 \in S$ such that $x_1 + x_2 = x_3 + x_4$, we have that $\phi(x_1) + \phi(x_2) = \phi(x_3) + \phi(x_4)$.*

Definition 2.3 (Freiman isomorphism). A Freiman isomorphism is a bijective Freiman homomorphism ϕ such that its inverse is also a Freiman homomorphism; this is equivalent to requiring that

$$\forall a, b, c, d \in S : a + b = c + d \iff \phi(a) + \phi(b) = \phi(c) + \phi(d).$$

We shall need the following well-known theorems and lemmas in additive combinatorics.

Theorem 2.4 (Freiman-Ruzsa theorem [Ruz99, EZ12]). Let $A \subseteq \mathbb{F}_2^n$. If $|A + A| \leq K \cdot |A|$, then $|\text{span}(A)| \leq 2^{2K}/(2K) \cdot |A|$.

Theorem 2.5 (Balog-Szemerédi-Gowers theorem). Let $A \subseteq \mathbb{F}_2^n$ be a set such that

$$E(A) = |\{(x_1, x_2, x_3, x_4) \in A^4 : x_1 + x_2 = x_3 + x_4\}| \geq |A|^3/K.$$

Then there is a set $A' \subseteq A$ such that

$$|A'| \geq |A|/P_{BSG}^{(1)}(K) \quad \text{and} \quad |A' + A'| \leq P_{BSG}^{(2)}(K)|A|.$$

Lemma 2.6 (Plünnecke’s inequality [TV06]). If $A \subseteq \mathbb{F}_2^n$ satisfies $|2A|/|A| \leq K$, then $|4A|/|A| \leq K^4$.

Lemma 2.7 (Ruzsa’s covering lemma). If $S, T \subseteq \mathbb{F}_2^n$ satisfy $|T + S| \leq K|S|$, then there is a subset $X \subseteq T$ of size $|X| \leq K$ such that $T \subseteq X + 2S$.

2.2 Quantum information

Our quantum algorithms revolve around stabilizer states; we will use the following characterization of stabilizer states.

Theorem 2.8 ([Nes08]). Every k qubit stabilizer state can be written in the following form

$$\frac{1}{\sqrt{|A|}} \sum_{x \in A} i^{\ell(x)} (-1)^{q(x)} |x\rangle,$$

for some affine subspace $A \subseteq \mathbb{F}_2^k$, quadratic polynomial q and linear polynomial ℓ in the variables $(x_1, \dots, x_k) \in \mathbb{F}_2^k$.

Our quantum algorithmic PFR theorems will crucially use the agnostic learnability of stabilizer states. Informally the task here is as follows: supposing an arbitrary quantum state $|\psi\rangle$ was τ -close to an *unknown* stabilizer state $|\phi\rangle$, output the “nearest” stabilizer state $|\phi'\rangle$ that is $\tau - \varepsilon$ close. A recent work of Chen, Gong, Ye, and Zhang [CGYZ25], gave an agnostic learning algorithm that runs in time quasipolynomial in $1/\tau$ and polynomial in the other parameters. Formally, their result is stated in the following theorem.

Theorem 2.9 ([CGYZ25]). Let \mathcal{C} be the class of stabilizer states. Fix any $\varepsilon \leq \tau \in (0, 1)$. There is an algorithm that, given access to copies of ρ with $\max_{|\phi'\rangle \in \mathcal{C}} |\langle \phi' | \rho | \phi' \rangle| \geq \tau$, outputs a $|\phi\rangle \in \mathcal{C}$ such that $|\langle \phi | \rho | \phi \rangle| \geq \tau - \varepsilon$ with high probability. The algorithm performs single-copy and two-copy measurements on at most $n \cdot \text{poly}(1/\varepsilon, (1/\tau)^{\log 1/\tau})$ copies of ρ and runs in time $n^3 \text{poly}(1/\varepsilon, (1/\tau)^{\log 1/\tau})$.

We will also require the following subroutines for estimating the overlap between two states and obtaining unitaries that prepare stabilizer states

Lemma 2.10 (SWAP test). Let $\varepsilon, \delta \in (0, 1)$. Given two arbitrary n -qubit quantum states $|\psi\rangle$ and $|\phi\rangle$, there is a quantum algorithm that estimates $|\langle \psi | \phi \rangle|^2$ up to error ε with probability at least $1 - \delta$ using $O(1/\varepsilon^2 \cdot \log(1/\delta))$ copies of $|\psi\rangle, |\phi\rangle$ and runs in $O(n/\varepsilon^2 \cdot \log(1/\delta))$ time.

Lemma 2.11 (Clifford synthesis [DDM03, PMH03]). *Given the classical description of an n -qubit stabilizer state $|\phi\rangle$, there is a quantum algorithm that outputs a Clifford circuit C that prepares $|\psi\rangle$, using $O(n^2/\log n)$ many single-qubit and two-qubit Clifford gates.*

2.3 Classical learning algorithms

Our classical algorithmic PFR theorems will crucially use the recent result by Briët and Castro-Silva [BC25], which can be thought of as dequantizing the agnostic learning algorithm for stabilizer states of [CGYZ25]. In particular, they showed that if one had *query access* to the amplitude vector of the unknown quantum state $|\psi\rangle$, then one can obtain a quadratic polynomial that most closely approximates the amplitudes. Formally, they proved the following theorem.

Theorem 2.12 ([BC25, Theorem 1.1]). *Let $f : \mathbb{F}_2^n \rightarrow [-1, 1]$ be a 1-bounded function and let $\varepsilon, \delta > 0$. There is a randomized algorithm \mathcal{A} that makes $n^2 \log n \log(1/\delta)(1/\varepsilon)^{O(\log(1/\varepsilon))}$ queries to f and has $O(n^3)$ time complexity such that, with probability at least $1 - \delta$, outputs a quadratic polynomial $p : \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ satisfying*

$$\left| \mathbb{E}_{x \in \mathbb{F}_2^n} [f(x)(-1)^{p(x)}] \right| \geq \max_{q \text{ quadratic}} \left| \mathbb{E}_{x \in \mathbb{F}_2^n} [f(x)(-1)^{q(x)}] \right| - \varepsilon.$$

3 Algorithmic lemmas

In this section, we will prove lemmas that will be useful in establishing our quantum and classical algorithmic PFR theorems.

3.1 Probabilistic dense model and sparse set localization

The following lemmas provide efficient randomized algorithms for finding Freiman isomorphisms and localizing sparse sets by showing that there is an abundance of them, and hence they can be sampled at random efficiently.

Lemma 3.1 (Localizing a sparse set). *Let $\varepsilon, \delta > 0$ and $A \subseteq \mathbb{F}_2^n$. Let $m := \log|\text{span}(A)|$, $k := \lceil 2m/\varepsilon \rceil \cdot \lceil \log(1/\delta) \rceil$. If v_1, \dots, v_k are uniformly random elements of A then, with probability at least $1 - \delta$, we have*

$$|A \cap \text{span}\{v_1, \dots, v_k\}| \geq (1 - \varepsilon)|A|.$$

Proof. Let $\ell \geq 2$ be an integer to be chosen later, and let v_1, \dots, v_ℓ be ℓ independent random elements of A . Let $V_0 = \{0\}$ and, for each $1 \leq i \leq \ell$, denote the linear span of the first i random elements v_1, \dots, v_i by V_i . Suppose first that

$$\Pr_{v_1, \dots, v_\ell \in A} [|A \cap V_\ell| \geq (1 - \varepsilon)|A|] < 1/2. \quad (3)$$

Then $\Pr_{v_1, \dots, v_\ell \in A} [|A \setminus V_\ell| > \varepsilon|A|] \geq 1/2$, and so

$$\Pr_{v_1, \dots, v_i \in A} [|A \setminus V_i| \geq \varepsilon|A|] > 1/2 \quad \text{for all } 0 \leq i \leq \ell. \quad (4)$$

It follows that

$$\begin{aligned}
\mathbb{E}_{v_1, \dots, v_\ell \in A} [\dim(V_\ell)] &= \sum_{i=1}^{\ell} \Pr_{v_1, \dots, v_i \in A} [v_i \notin V_{i-1}] \\
&\geq \sum_{i=1}^{\ell} \Pr_{v_1, \dots, v_{i-1} \in A} [|A \setminus V_{i-1}| \geq \varepsilon |A|] \cdot \Pr_{v_1, \dots, v_i \in A} [v_i \notin V_{i-1} \mid |A \setminus V_{i-1}| \geq \varepsilon |A|] \\
&> \sum_{i=1}^{\ell} 1/2 \cdot \varepsilon = \varepsilon \ell / 2,
\end{aligned}$$

where the final inequality used Eq. (4). Since $V_\ell \subseteq \text{span}(A)$, we must have $\dim(V_\ell) \leq \log |\text{span}(A)| = m$, and thus $\ell < 2m/\varepsilon$ is required for equation (3) to hold. Denoting $t := \lceil 2m/\varepsilon \rceil$, we conclude that

$$\Pr_{v_1, \dots, v_t \in A} [|A \cap \text{span}\{v_1, \dots, v_t\}| \geq (1 - \varepsilon)|A|] \geq 1/2.$$

Repeating this sampling $\lceil \log(1/\delta) \rceil$ times independently at random, the probability that we succeed at least once is at least $1 - \delta$. With $k = t \lceil \log(1/\delta) \rceil$, it follows that

$$\Pr_{v_1, \dots, v_k \in A} [|A \cap \text{span}\{v_1, \dots, v_k\}| \geq (1 - \varepsilon)|A|] \geq 1 - \delta,$$

proving the lemma statement. \square

Lemma 3.2 (Algorithmic dense model). *Let $\delta > 0$, $A \subseteq \mathbb{F}_2^n$ and let $m \geq \log |4A| + \log 1/\delta$ be an integer. Suppose $\pi : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ is a random linear map. Then A is Freiman-isomorphic to $\pi(A)$ with probability at least $1 - \delta$.*

Proof. Recall that π is a Freiman isomorphism between A and $\pi(A)$ if

$$\forall a, b, c, d \in A : a + b + c + d = 0 \iff \pi(a) + \pi(b) + \pi(c) + \pi(d) = 0.$$

Observe that the property above implies that π is bijective.² If $\pi : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ is a linear map, then the forward implication is automatically satisfied, and moreover

$$\pi(a) + \pi(b) + \pi(c) + \pi(d) = \pi(a + b + c + d).$$

It then suffices to check that

$$\forall a, b, c, d \in A : \pi(a + b + c + d) = 0 \implies a + b + c + d = 0,$$

which is equivalent to requiring that $\pi(x) \neq 0$ for all nonzero $x \in 4A$.

Now let $\pi : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ be a uniformly random linear map. Then, for each $x \in 4A \setminus \{0\}$ individually, $\pi(x)$ is uniformly distributed over \mathbb{F}_2^m . It follows from the union bound that

$$\Pr [\exists x \in 4A \setminus \{0\} : \pi(x) = 0] \leq \frac{|4A| - 1}{2^m},$$

which is less than δ if $2^m \geq |4A|/\delta$. This concludes the proof. \square

²To see this, if $c = d$, then the implication above gives $a = b \iff \pi(a) = \pi(b)$, which implies π is a bijection.

3.2 Algorithmic restricted homomorphism

We will show the following slight strengthening of the homomorphism testing formulation of the PFR theorem, which we prove following the approach of Green [Gre05] and Green-Tao [GT10], and then algorithmize.

Lemma 3.3 (Restricted homomorphism testing). *Suppose $S \subseteq \mathbb{F}_2^m$ and $f : S \rightarrow \mathbb{F}_2^n$ satisfy*

$$|\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4 \text{ and } f(x_1) + f(x_2) = f(x_3) + f(x_4)\}| \geq 2^{3m}/K.$$

Then there exists an affine-linear function $\psi : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ such that $f(x) = \psi(x)$ for at least $2^m/P_4(K)$ values of $x \in S$.

Proof. Consider the “graph” set

$$\Gamma = \{(x, f(x)) : x \in S\} \subseteq \mathbb{F}_2^{m+n}.$$

Then $|\Gamma| = |S| \leq 2^m$ and $E(\Gamma) \geq 2^{3m}/K \geq |\Gamma|^3/K$. By the Balog-Szemerédi-Gowers Theorem (Theorem 2.5), there exists a set $\Gamma' \subseteq \Gamma$ such that

$$|\Gamma'| \geq |\Gamma|/P_{BSG}^{(1)}(K) \quad \text{and} \quad |\Gamma' + \Gamma'| \leq P_{BSG}^{(2)}(K) \cdot |\Gamma|.$$

By the combinatorial PFR theorem (Theorem 1.1), it then follows that Γ' can be covered by

$$K' := P_1(P_{BSG}^{(1)}(K)P_{BSG}^{(2)}(K))$$

translates of a subspace $H \leq \mathbb{F}_2^{m+n}$ of size $|H| \leq |\Gamma'|$, say

$$\Gamma' \subseteq \bigcup_{i=1}^{K'} (u_i + H).$$

Let $\pi : \mathbb{F}_2^{m+n} \rightarrow \mathbb{F}_2^m$ denote the projection map onto the first m coordinates: $\pi(x, y) = x$ for $x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n$. Let $\ker_H(\pi) = H \cap (\{0^m\} \times \mathbb{F}_2^n)$ be the kernel of π restricted to H and let H' be a complemented subspace of $\ker_H(\pi)$ in H , so that $H = H' \oplus \ker_H(\pi)$. By linearity and the injectivity of π on H' , there exists a matrix $M \in \mathbb{F}_2^{n \times m}$ such that

$$H' = \{(x, Mx) : x \in \pi(H)\}, \tag{5}$$

and by the rank-nullity theorem we have that

$$|H| = |\ker_H(\pi)| \cdot |\pi(H)|.$$

Moreover, since Γ' is a graph, we have

$$|\Gamma' \cap (u_i + H)| = |\pi(\Gamma' \cap (u_i + H))| \leq |\pi(u_i + H)| = |\pi(H)|,$$

and thus

$$|\Gamma'| = \left| \Gamma' \cap \bigcup_{i=1}^{K'} (u_i + H) \right| \leq \sum_{i=1}^{K'} |\Gamma' \cap (u_i + H)| \leq K' |\pi(H)|,$$

from which we conclude that $|\pi(H)| \geq |\Gamma'|/K'$. Finally, since $H = H' \oplus \ker_H(\pi)$, we have

$$|\Gamma'| = \left| \Gamma' \cap \bigcup_{i=1}^{K'} \bigcup_{v \in \ker_H(\pi)} (u_i + v + H') \right| \leq \sum_{i=1}^{K'} \sum_{v \in \ker_H(\pi)} |\Gamma' \cap (u_i + v + H')|;$$

there must then exist some translate u_i and some $v \in \ker_H(\pi)$ such that

$$|\Gamma' \cap (u_i + v + H')| \geq \frac{|\Gamma'|}{K'|\ker_H(\pi)|}.$$

Using the assumption $|H| \leq |\Gamma'|$, the identity $|H| = |\ker_H(\pi)| \cdot |\pi(H)|$ and the bound $|\pi(H)| \geq |\Gamma'|/K'$, we conclude from the last inequality that

$$|\Gamma' \cap (u_i + v + H')| \geq \frac{|H|}{K'|\ker_H(\pi)|} = \frac{|\pi(H)|}{K'} \geq \frac{|\Gamma'|}{(K')^2}.$$

We can now easily conclude. Fixing $u_i = (x_1, y_1)$, $v = (x_2, y_2) \in \mathbb{F}_2^{m+n}$ such that the above inequality holds, we obtain from the description of H' (equation (5)) that

$$\begin{aligned} \Gamma' \cap (u_i + v + H') &= \Gamma' \cap \{(x + x_1 + x_2, Mx + y_1 + y_2) : x \in \pi(H)\} \\ &= \Gamma' \cap \{(x, Mx - Mx_1 - Mx_2 + y_1 + y_2) : x \in \pi(H) + x_1 + x_2\}. \end{aligned}$$

There must then be at least $|\Gamma'|/(K')^2$ values of $x \in S$ such that $(x, f(x)) \in \Gamma'$ and

$$f(x) = Mx - Mx_1 - Mx_2 + y_1 + y_2.$$

Denote $\psi(x) = Mx - Mx_1 - Mx_2 + y_1 + y_2$. Recalling that $|\Gamma'| \geq |\Gamma|/P_{BSG}^{(1)}(K)$ and

$$2^{3m}/K \leq E(\Gamma) \leq |\Gamma|^3,$$

we obtain that $f(x) = \psi(x)$ for at least

$$\frac{|\Gamma'|}{(K')^2} = \frac{|\Gamma'|}{P_1(P_{BSG}^{(1)}(K)P_{BSG}^{(2)}(K))^2} \geq \frac{2^m}{KP_{BSG}^{(1)}(K)P_1(P_{BSG}^{(1)}(K)P_{BSG}^{(2)}(K))^2}$$

values of $x \in S$. The theorem follows with $P_4(K) = KP_{BSG}^{(1)}(K)P_1(P_{BSG}^{(1)}(K)P_{BSG}^{(2)}(K))^2$. \square

The main technical lemma for our proofs is an algorithmic version of this last result, which relies on the quadratic Goldreich-Levin theorem (see Theorem 2.12).

Lemma 3.4 (Algorithmic restricted homomorphism). *Suppose $S \subseteq \mathbb{F}_2^m$ and $f : S \rightarrow \mathbb{F}_2^n$ satisfy*

$$|\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4 \text{ and } f(x_1) + f(x_2) = f(x_3) + f(x_4)\}| \geq 2^{3m}/K.$$

There is a randomized algorithm that makes $K^{O(\log K)}(m+n)^2 \log(m+n)$ queries to S and to f , runs in $K^{O(\log K)}(m+n)^3 \log(m+n)$ time and, with probability at least 0.7, returns $M \in \mathbb{F}_2^{n \times m}$, $v \in \mathbb{F}_2^n$ such that

$$|\{x \in S : f(x) = Mx + v\}| \geq 2^m/P_4'(K).$$

Proof. Define the function $g : \mathbb{F}_2^{m+n} \rightarrow \{-1, 0, 1\}$ by

$$g(x, y) = \mathbf{1}_S(x) \cdot (-1)^{f(x) \cdot y}.$$

Note that one query to g can be made using one query to S , one query to f and $O(n)$ time. We first show that g correlates well with a quadratic function:

Claim 3.5. *There exists a quadratic polynomial $p : \mathbb{F}_2^{m+n} \rightarrow \mathbb{F}_2$ such that*

$$\left| \mathbb{E}_{\substack{x \in \mathbb{F}_2^m, \\ y \in \mathbb{F}_2^n}} g(x, y) (-1)^{p(x, y)} \right| \geq \frac{1}{P_4(K)},$$

where $P_4(\cdot)$ is the polynomial promised by Lemma 3.3.

Proof. From Lemma 3.3, we know there exists an affine-linear function $\psi : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ such that

$$\Pr_{x \in \mathbb{F}_2^m} [x \in S \text{ and } f(x) = \psi(x)] \geq \frac{1}{P_4(K)}.$$

Let E be the set where f and ψ agree:

$$E = \{x \in S : f(x) = \psi(x)\}.$$

Note that $g(x, y) = (-1)^{\psi(x) \cdot y}$ for all $x \in E$, $y \in \mathbb{F}_2^n$, and so by Cauchy-Schwarz

$$\begin{aligned} \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbf{1}_E(x) &= \mathbb{E}_{x \in \mathbb{F}_2^m} \left(\mathbf{1}_E(x) \cdot \mathbb{E}_{y \in \mathbb{F}_2^n} g(x, y) (-1)^{\psi(x) \cdot y} \right) \\ &\leq \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \mathbf{1}_E(x)^2 \right)^{1/2} \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \left(\mathbb{E}_{y \in \mathbb{F}_2^n} g(x, y) (-1)^{\psi(x) \cdot y} \right)^2 \right)^{1/2} \\ &= \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \mathbf{1}_E(x) \right)^{1/2} \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{y, y' \in \mathbb{F}_2^n} g(x, y) g(x, y') (-1)^{\psi(x) \cdot (y+y')} \right)^{1/2} \\ &= \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \mathbf{1}_E(x) \right)^{1/2} \left(\mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{z \in \mathbb{F}_2^n} \mathbf{1}_S(x) (-1)^{f(x) \cdot z} (-1)^{\psi(x) \cdot z} \right)^{1/2}. \end{aligned}$$

We conclude that

$$\left| \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{z \in \mathbb{F}_2^n} g(x, z) (-1)^{\psi(x) \cdot z} \right| \geq \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbf{1}_E(x) = \Pr_{x \in \mathbb{F}_2^m} [x \in E] \geq \frac{1}{P_4(K)}.$$

The quadratic function $p : (x, z) \mapsto \psi(x) \cdot z$ thus satisfies the claim. \square

We now use the quadratic Goldreich-Levin theorem (Theorem 2.12) with f replaced by g and $\varepsilon := 1/(2P_4(K))$. We conclude that, in $O((m+n)^3)$ time and using $(m+n)^2 \log(m+n) \cdot K^{O(\log(K))}$ queries to g , we can obtain a quadratic function $q : \mathbb{F}_2^{m+n} \rightarrow \mathbb{F}_2$ which satisfies the following with probability at least 0.9:

$$\left| \mathbb{E}_{x \in \mathbb{F}_2^m, y \in \mathbb{F}_2^n} \mathbf{1}_S(x) (-1)^{f(x) \cdot y} (-1)^{q(x, y)} \right| \geq \frac{1}{2P_4(K)}. \quad (6)$$

Assume that this inequality holds, and write

$$q(x, y) = (x, y)^T A(x, y) + u \cdot x + u' \cdot y + b,$$

where $A \in \mathbb{F}_2^{(m+n) \times (m+n)}$, $u \in \mathbb{F}_2^m$, $u' \in \mathbb{F}_2^n$ and $b \in \mathbb{F}_2$. Denote the $(m \times n)$ -submatrix of A defined by its first m rows and last n columns by A_{12} , and the $(n \times m)$ -submatrix of A defined by its last n rows and first m columns by A_{21} . We claim that f agrees often with an affine-linear function whose linear part equals $(A_{12}^T + A_{21})x$:

Claim 3.6. *If equation (6) holds, then there exists some $z_0 \in \mathbb{F}_2^n$ such that*

$$|\{x \in S : f(x) = (A_{12}^T + A_{21})x + z_0\}| \geq \frac{2^m}{64P_4(K)^3}. \quad (7)$$

Proof. Define the bilinear form $B : \mathbb{F}_2^m \times \mathbb{F}_2^n \rightarrow \mathbb{F}_2$ by

$$B(x, y) = q(x, y) - q(x, 0) - q(0, y) + q(0, 0).$$

We have that

$$\begin{aligned} B(x, y) &= [x \ y] A \begin{bmatrix} x \\ y \end{bmatrix} + [u \ u'] \begin{bmatrix} x \\ y \end{bmatrix} + b - [x \ 0^n] A \begin{bmatrix} x \\ 0^n \end{bmatrix} - [u \ u'] \begin{bmatrix} x \\ 0^n \end{bmatrix} - b \\ &\quad - [0^m \ y] A \begin{bmatrix} 0^m \\ y \end{bmatrix} - [u \ u'] \begin{bmatrix} 0^m \\ y \end{bmatrix} - b + b \\ &= ([x \ 0^n] + [0^m \ y]) A \begin{bmatrix} x \\ y \end{bmatrix} - [x \ 0^n] A \begin{bmatrix} x \\ 0^n \end{bmatrix} - [0^m \ y] A \begin{bmatrix} 0^m \\ y \end{bmatrix} \\ &= [x \ 0^n] A \begin{bmatrix} 0^m \\ y \end{bmatrix} + [0^m \ y] A \begin{bmatrix} x \\ 0^n \end{bmatrix} \\ &= x^T A_{12} y + y^T A_{21} x \\ &= y^T (A_{12}^T + A_{21}) x. \end{aligned}$$

Write $\sigma := 1/(2P_4(K))$ and $M := A_{12}^T + A_{21}$ for convenience, so that $B(x, y) = Mx \cdot y$. Plugging in

$$q(x, y) = Mx \cdot y + q(x, 0) + q(0, y) - q(0, 0)$$

into Eq. (6), we obtain

$$\left| \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{y \in \mathbb{F}_2^n} \mathbf{1}_S(x) (-1)^{f(x) \cdot y} (-1)^{Mx \cdot y} (-1)^{q(x, 0) + q(0, y) - q(0, 0)} \right| \geq \sigma.$$

By the triangle inequality, we conclude that

$$\sum_{x \in S} \left| \mathbb{E}_{y \in \mathbb{F}_2^n} (-1)^{f(x) \cdot y} (-1)^{Mx \cdot y} (-1)^{q(0, y)} \right| \geq \sigma \cdot 2^m, \quad (8)$$

where we used the fact that for a fixed $x \in S$, the quantity $(-1)^{q(x, 0) - q(0, 0)}$ is constant and has absolute value 1.

Defining the function $h : \mathbb{F}_2^n \rightarrow \{-1, 1\}$ by $h(y) = (-1)^{q(0, y)}$, one can rewrite Eq. (8) as

$$\sum_{x \in S} |\widehat{h}(f(x) + Mx)| \geq \sigma \cdot 2^m.$$

Since $|\widehat{h}(z)| \leq 1$ for all $z \in \mathbb{F}_2^n$, this implies that there exist at least $(\sigma/2) \cdot 2^m$ many $x \in S$ such that $|\widehat{h}(f(x) + Mx)| \geq \sigma/2$. Let us define the set $T = \{z \in \mathbb{F}_2^n : |\widehat{h}(z)| \geq \sigma/2\}$, so that

$$|\{x \in S : f(x) + Mx \in T\}| \geq \frac{\sigma 2^m}{2}.$$

Then

$$\frac{\sigma 2^m}{2} \leq \sum_{z \in T} |\{x \in S : f(x) + Mx = z\}| \leq |T| \cdot \max_{z_0 \in T} |\{x \in S : f(x) + Mx = z_0\}|.$$

Since h is a Boolean function, by Parseval we have that

$$1 = \sum_{z \in \mathbb{F}_2^n} |\widehat{h}(z)|^2 \geq \sum_{z \in T} (\sigma/2)^2 \implies |T| \leq 4/\sigma^2.$$

From these last two inequalities, we conclude there exists some $z_0 \in T$ such that

$$|\{x \in S : f(x) + Mx = z_0\}| \geq \frac{1}{|T|} \frac{\sigma 2^m}{2} \geq \frac{\sigma^3 2^m}{8},$$

which proves the claim. \square

It now suffices to find such a vector $z_0 \in \mathbb{F}_2^n$ such that Eq. (7) holds. We do this by sampling x_1, x_2, \dots, x_t uniformly at random from \mathbb{F}_2^m , passing each point x_i through the indicator function $\mathbf{1}_S$ and then computing the difference $d(x_i) = f(x_i) - (A_{12}^T + A_{21})x_i$. For each $z \in \{d(x_i)\}_{i \in [t]}$, we then estimate $\Pr_{x \in S} [f(x) = (A_{12}^T + A_{21})x + z]$ and output the value z^* which maximizes the agreement. To complete the argument, let us now comment on the value of t required to determine a good value of z^* . First, note that Eq. (7) implies

$$\Pr_{x \in \mathbb{F}_2^m} [d(x) = z_0] \geq \frac{1}{64P_4(K)^3}.$$

Thus, by sampling $t = O(P_4(K)^3)$ times, we ensure that $v_0 \in \{d(x_i)\}_{i \in [t]}$ with probability at least 0.9. Finally, we determine z^* as mentioned before by estimating $\Pr_{x \in S} [f(x) = (A_{12}^T + A_{21})x + z]$ for each $z \in \{d(x_i)\}_{i \in [t]}$, which can be done up to error $1/(128P_4(K)^3)$ with probability at least $1 - 0.1/t$ using an empirical estimator³ that uses $O(\log(K)P_4(K)^3)$ samples from \mathbb{F}_2^n and queries to S and f for each $i \in [t]$. In total, this procedure consumes $O(\log(K)P_4(K)^6)$ queries to S and to f , and succeeds with probability at least 0.8 (after taking the union bound).

We then return $M = A_{12}^T + A_{21}$ and $v = z^*$ as given above. With probability at least 0.7, the guarantee of the statement is satisfied with $P'_4(K) = 128P_4(K)^3$. The overall query and time complexities of the algorithm are mostly due to the use of Theorem 2.12, and match the complexities stated in the lemma. \square

4 Proof of the algorithmic PFR theorems

In this section, we provide the proof of our main results, relying on the algorithmic lemmas established in Section 3. We begin with classical algorithms in Section 4.1 and then proceed to quantum algorithms in Section 4.2.

4.1 Classical algorithmic PFR theorems

We first prove our main classical results, which we restate more precisely below.

Theorem 4.1 (Algorithmic PFR). *Suppose $A \subseteq \mathbb{F}_2^n$ satisfies $|A+A| \leq K|A|$. There is a randomized algorithm that takes $O(\log|A|+K)$ random samples from A , makes $2^{O(K)} \log^2|A| \cdot \log \log|A|$ queries to A , runs in time $K^{O(\log K)} n^4 \log n$ and has the following guarantee: with probability at least $2/3$, it outputs a basis for a subspace $V \leq \mathbb{F}_2^n$ of size $|V| \leq |A|$ such that A can be covered by $P'_1(K)$ translates of V .*

³In particular, for Boolean functions f, g , one can estimate $\Pr_x[f(x) = g(x)]$ up to error ε with probability at least $1 - \delta$ using the empirical estimate $\text{Est}_m := \frac{1}{m} \sum_{j=1}^m f(x_j)g(x_j)$ which can be computed by querying f, g at uniformly random $x_1, \dots, x_m \in \mathbb{F}_2^n$ and for $m = \text{poly}(1/\varepsilon \log(1/\delta))$.

Proof. We first describe the algorithm to find V :

1. Sample $t = 28 \log|A| + 56K$ many uniformly random elements from A , and denote their linear span by U . Let $A' := A \cap U$.
2. Take a random linear map $\pi : U \rightarrow \mathbb{F}_2^m$ where $m = \log|A| + 4 \log K + 10$. Let $S = \pi(A')$ denote the image of A' under π , and let $f : S \rightarrow U$ be the inverse of π when restricted to S .⁴
3. Apply Lemma 3.4 to obtain an affine-linear map $\psi : \mathbb{F}_2^m \rightarrow U$ such that $f(x) = \psi(x)$ for at least $|A|/P'_4(2^{33}K^{13})$ values $x \in S$.
4. Take a subspace V of $\text{Im}(\psi)$ having size at most $|A|$, and output a basis for V .

We now analyze the correctness and complexity of this algorithm. For Step (1), note that Theorem 2.4 directly implies that $|\text{span}(A)| \leq 2^{2K} \cdot |A|$. Now, by our choice of t , Lemma 3.1 implies that $|A'| \geq |A|/2$ with probability at least 0.99. Supposing this is the case, we have that

$$|A' + A'| \leq |A + A| \leq K|A| \leq 2K|A'|.$$

Moreover, by Lemma 2.6 we conclude that $|4A'| \leq |4A| \leq K^4|A| \leq 2K^4|A'|$.

For Step (2), note that Lemma 3.2 shows that (with probability at least 0.99) π is a Freiman isomorphism from A' to $S = \pi(A')$. In this case, the inverse map $f : S \rightarrow A'$ is a Freiman isomorphism and $|S| = |A'|$.⁵

In Step (3) we wish to apply Lemma 3.4, which requires us to bound from below the quantity

$$|\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4 \text{ and } f(x_1) + f(x_2) = f(x_3) + f(x_4)\}|.$$

We claim that this is at least $|A'|^3/(2K)$:

Claim 4.2. *If $f : S \rightarrow A'$ is a Freiman isomorphism and $|A' + A'| \leq 2K|A'|$, then*

$$|\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4 \text{ and } f(x_1) + f(x_2) = f(x_3) + f(x_4)\}| \geq \frac{|A'|^3}{2K}.$$

Proof. If f is a Freiman isomorphism, then the quantity above equals

$$\begin{aligned} & |\{(x_1, x_2, x_3, x_4) \in S^4 : f(x_1) + f(x_2) = f(x_3) + f(x_4)\}| \\ &= |\{(y_1, y_2, y_3, y_4) \in (A')^4 : y_1 + y_2 = y_3 + y_4\}| \\ &= E(A'). \end{aligned}$$

Now note that

$$\sum_{z \in 2A'} |\{(y_1, y_2) \in (A')^2 : y_1 + y_2 = z\}| = |A'|^2$$

and

$$\begin{aligned} & \sum_{z \in 2A'} |\{(y_1, y_2) \in (A')^2 : y_1 + y_2 = z\}|^2 \\ &= \sum_{z \in 2A'} |\{(y_1, y_2, y_3, y_4) \in (A')^4 : y_1 + y_2 = z = y_3 + y_4\}|^2 \\ &= E(A'), \end{aligned}$$

⁴In our analysis we show that this inverse is well-defined with high probability.

⁵We remark that the definition of Freiman isomorphism gives these two properties immediately.

so by Cauchy-Schwarz

$$|A'|^2 \leq |2A'|^{1/2} E(A')^{1/2} \implies E(A') \geq |A'|^4 / |2A'|.$$

The claim now follows from the assumption $|2A'| \leq 2K|A'|$. \square

Next we note that, by assumption and by our choice for m , we have that

$$|A'| \geq \frac{|A|}{2} \geq \frac{2^m}{2^{11} K^4}.$$

From the claim, we then conclude that S and f satisfy the hypothesis of Lemma 3.4 with K substituted by $K' := 2^{33} K^{13}$. We can then obtain an affine-linear function $\psi : \mathbb{F}_2^m \rightarrow U$ such that, with probability at least 0.7,

$$|\{x \in S : f(x) = \psi(x)\}| \geq \frac{2^m}{P'_4(2^{33} K^{13})}. \quad (9)$$

It remains to argue how one can simulate queries to S and f , as required by the statement of Lemma 3.4. To this end, observe that we have a full description of the linear map $\pi : U \rightarrow \mathbb{F}_2^m$, so in time $O(m^2 n)$ we can find its kernel $\ker(\pi) = \{v \in U : \pi(v) = 0\}$. We first make three observations about this: (a) $\ker(\pi)$ is a subspace of size

$$\frac{|U|}{|\text{Im}(\pi)|} \leq \frac{|\text{span}(A)|}{|S|} \leq \frac{2|\text{span}(A)|}{|A|} \leq 2^{2K},$$

where we used Theorem 2.4 in the final inequality; (b) for every $x \in \text{Im}(\pi)$, we have that $\pi^{-1}(x)$ is a translate of $\ker(\pi)$; (c) in $O(m^2 n)$ time we can find the inverse map $\pi^{-1} : \text{Im}(\pi) \rightarrow U/\ker(\pi)$. Using item (b), we can check whether $x \in S$ (i.e., $\pi^{-1}(x) \cap A \neq \emptyset$) by enumerating over all $y \in \pi^{-1}(x)$ and checking if $y \in A$ or not. By item (a), this takes at most 2^{2K} queries to A . Hence, after computing $\ker(\pi)$ and π^{-1} , one can make one query to S and to f using 2^{2K} queries to A and $O(mn)$ time.

Now define the affine subspace $V' = \text{Im}(\psi)$. By definition we have that $|V'| \leq 2^m = O(K^4)|A'|$, and by Eq. (9) we have

$$|A \cap V'| = |\text{Im}(f) \cap \text{Im}(\psi)| \geq |S| \Pr_{x \in S}[f(x) = \psi(x)] \geq \frac{2^m}{P'_4(2^{33} K^{13})}.$$

It follows that

$$|A + (A \cap V')| \leq |A + A| \leq K|A| \leq 2^m \leq P'_4(2^{33} K^{13})|A \cap V'|.$$

Applying Ruzsa's covering lemma (Lemma 2.7), we obtain that A can be covered by $P'_4(2^{33} K^{13})$ translates of $2(A \cap V') \subseteq V' + V' = \psi(0) + V'$.

Overall, the complexity of the algorithm is as follows: we need $O(K + \log|A|)$ samples from A , and the number of queries to A is as given by Lemma 3.4, i.e.,

$$2^{2K} \cdot K^{O(\log K)} (m + \log|U|)^2 \log(m + \log|U|) = 2^{O(K)} (\log|A|)^2 \log \log|A|,$$

where we used that $m = \log|A| + O(\log K)$ and $\log|U| = O(\log|A| + K)$. The total runtime is the cost of Lemma 3.4, the cost of inverting π , and the cost for making the queries to f and S , i.e.,

$$K^{O(\log K)} (m+n)^3 \log(m+n) + O(m^2 n) + K^{O(\log K)} (m+n)^2 \log(m+n) \cdot O(mn) = K^{O(\log K)} n^4 \log n,$$

proving the theorem statement. \square

Theorem 4.3 (Homomorphism testing). *Suppose $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ satisfies*

$$\Pr_{x_1+x_2=x_3+x_4} [f(x_1) + f(x_2) = f(x_3) + f(x_4)] \geq 1/K.$$

There is a randomized algorithm that makes $K^{O(\log K)}(m+n)^2 \log(m+n)$ queries to f , runs in $K^{O(\log K)}(m+n)^3 \log(m+n)$ time and, with probability at least $2/3$, outputs a matrix $M \in \mathbb{F}_2^{n \times m}$ and a vector $v \in \mathbb{F}_2^n$ such that

$$\Pr_{x \in \mathbb{F}_2^m} [f(x) = Mx + v] \geq 1/P'_2(K).$$

Proof. This follows immediately from Lemma 3.4 with $S = \mathbb{F}_2^m$ and $P'_2(K) = P'_4(K)$. \square

Theorem 4.4 (Structured approximate homomorphism). *Suppose $f : \mathbb{F}_2^m \rightarrow \mathbb{F}_2^n$ satisfies*

$$|\{f(x) + f(y) - f(x+y) : x, y \in \mathbb{F}_2^m\}| \leq K.$$

There is a randomized algorithm that makes $K^{O(\log K)}(m+n)^2 \log(m+n)$ queries to f , runs in $K^{O(\log K)}(m+n)^3 \log(m+n)$ time and, with probability at least $2/3$, outputs a matrix $M \in \mathbb{F}_2^{n \times m}$ such that

$$|\{f(x) - Mx : x \in \mathbb{F}_2^m\}| \leq P'_3(K).$$

Proof. We first show that the property in the statement implies that

$$\Pr_{x_1+x_2=x_3+x_4} [f(x_1) + f(x_2) = f(x_3) + f(x_4)] \geq \frac{1}{K}. \quad (10)$$

Indeed, denote $\Delta f := \{f(x) + f(y) - f(x+y) : x, y \in \mathbb{F}_2^m\}$, so that $|\Delta f| \leq K$ by assumption. Then

$$\begin{aligned} \mathbb{E}_{b \in \Delta f} \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{y \in \mathbb{F}_2^m} \mathbf{1}[f(x) + f(y) - f(x+y) = b] &= \frac{1}{|\Delta f|} \mathbb{E}_{x, y \in \mathbb{F}_2^m} \sum_{b \in \Delta f} \mathbf{1}[f(x) + f(y) - f(x+y) = b] \\ &= \frac{1}{|\Delta f|} \mathbb{E}_{x, y \in \mathbb{F}_2^m} \mathbf{1} \\ &\geq \frac{1}{K}, \end{aligned}$$

and so by Cauchy-Schwarz

$$\begin{aligned} \frac{1}{K^2} &\leq \mathbb{E}_{b \in \Delta f} \mathbb{E}_{x \in \mathbb{F}_2^m} \left(\mathbb{E}_{y \in \mathbb{F}_2^m} \mathbf{1}[f(x) + f(y) - f(x+y) = b] \right)^2 \\ &= \mathbb{E}_{b \in \Delta f} \mathbb{E}_{x \in \mathbb{F}_2^m} \mathbb{E}_{y, z \in \mathbb{F}_2^m} \mathbf{1}[f(x) + f(y) - f(x+y) = b = f(x) + f(z) - f(x+z)] \\ &= \frac{1}{K} \mathbb{E}_{x, y, z \in \mathbb{F}_2^m} \sum_{b \in \Delta f} \mathbf{1}[f(y) - f(x+y) = b - f(x) = f(z) - f(x+z)] \\ &= \frac{1}{K} \mathbb{E}_{x, y, z \in \mathbb{F}_2^m} \mathbf{1}[f(y) - f(x+y) = f(z) - f(x+z)] \\ &= \frac{1}{K} \mathbb{E}_{x_1+x_2=x_3+x_4} \mathbf{1}[f(x_1) + f(x_2) = f(x_3) + f(x_4)], \end{aligned}$$

which gives inequality (10) as desired.

We may then apply Lemma 3.4 (with $S = \mathbb{F}_2^m$) to obtain a matrix $M \in \mathbb{F}_2^{n \times m}$ and a vector $v \in \mathbb{F}_2^n$ such that, with probability at least 0.7, we have

$$\Pr_{x \in \mathbb{F}_2^m} [f(x) = Mx + v] \geq 1/P'_4(K). \quad (11)$$

We claim that, if this inequality holds (and $|\Delta f| \leq K$), then

$$|\{f(x) - Mx : x \in \mathbb{F}_2^m\}| \leq K^2 P'_4(K), \quad (12)$$

which is the property we want with $P'_3(K) = K^2 P'_4(K)$. It then suffices to prove (12).

Denote $E := \{x \in \mathbb{F}_2^m : f(x) = Mx + v\}$, so that $|E| \geq 2^m/P'_4(K)$ by Eq. (11). Then

$$|\mathbb{F}_2^m + E| = 2^m \leq P'_4(K) \cdot |E|,$$

so we may use Ruzsa's covering lemma (Lemma 2.7, with $S = E$ and $T = \mathbb{F}_2^m$) to conclude there exists a set $X \subseteq \mathbb{F}_2^m$ of size $P'_4(K)$ such that $\mathbb{F}_2^m \subseteq X + 2E$. In other words, every element of \mathbb{F}_2^m can be written as $x + y + z$ with $x \in X$ (where $|X| \leq P'_4(K)$) and $y, z \in E$.

Now, for every $x \in X, y, z \in E$, by definition of the set Δf there exist $b, b' \in \Delta f$ such that

$$f(x + y) - f(x) - f(y) = b \quad \text{and} \quad f(x + y + z) - f(x + y) - f(z) = b'.$$

Summing these two identities, we conclude that

$$\begin{aligned} f(x + y + z) &= f(x) + f(y) + f(z) + b + b' \\ &= f(x) + My + Mz + b + b' \\ &= f(x) + M(x + y + z) - Mx + b + b', \end{aligned}$$

and so

$$f(x + y + z) - M(x + y + z) = f(x) - Mx + b + b' \in \{f(x') - Mx' : x' \in X\} + \Delta f + \Delta f$$

is contained in a set of size at most $|X| \cdot |\Delta f|^2 \leq K^2 P'_4(K)$. This gives Eq. (12) and concludes the proof of the theorem. \square

4.2 Quantum algorithmic PFR theorem

We now give a quantum algorithm that is quadratically better than in the query complexity compared to the classical algorithm shown in the section above. We restate the statement of the quantum result in more detail below.

Theorem 4.5 (Quantum algorithmic PFR). *Suppose $A \subseteq \mathbb{F}_2^n$ satisfies $|A + A| \leq K|A|$. There is a quantum algorithm that takes $O(\log|A| + K)$ random samples from A , makes $2^{O(K)} \log|A|$ quantum queries to A , runs in time $K^{O(\log K)} n^3$ and has the following guarantee: with probability at least $2/3$, it outputs a basis for a subspace $V \subseteq \mathbb{F}_2^n$ of size $|V| \leq |A|$ such that A can be covered by $P'_1(K)$ translates of V .*

To prove the above theorem, we will reprove Lemma 3.4 in the quantum setting, but now taking advantage of the main result (Theorem 2.9) of Chen, Gong, Ye, and Zhang [CGYZ25], which allows us to find the closest stabilizer state to a given unknown n -qubit quantum state. Formally, the quantum version of Lemma 3.4 is as follows.

Lemma 4.6. Suppose $S \subseteq \mathbb{F}_2^m$ and $f : S \rightarrow \mathbb{F}_2^n$ satisfy

$$|\{(x_1, x_2, x_3, x_4) \in S^4 : x_1 + x_2 = x_3 + x_4 \text{ and } f(x_1) + f(x_2) = f(x_3) + f(x_4)\}| \geq 2^{3m}/K.$$

There is a quantum algorithm that makes $K^{O(\log K)}(m+n)$ quantum queries to S and to f , runs in $K^{O(\log K)}(m+n)^3$ time and, with probability at least 0.7, returns $M \in \mathbb{F}_2^{n \times m}$, $v \in \mathbb{F}_2^n$ such that

$$|\{x \in S : f(x) = Mx + v\}| \geq 2^m/P'_4(K).$$

To prove Lemma 4.6 and describe its corresponding algorithm, we need a quantum protocol to prepare the quantum state that encodes the function

$$g_S(x, y) = \mathbf{1}_S(x)(-1)^{f(x) \cdot y},$$

which from Claim 3.5, we know has high correlation with a quadratic function.

Claim 4.7. Consider the context of Lemma 4.6. Let $\delta \in (0, 1)$. Suppose we have quantum query access to S via the oracle O_S and query access to $f : S \rightarrow \mathbb{F}_2^n$ via the oracle O_f as follows

$$|x, 0\rangle \xrightarrow{O_S} |x, \mathbf{1}_S(x)\rangle, \quad |x, 0^n\rangle \xrightarrow{O_f} |x, f(x)\rangle.$$

There is a quantum algorithm that makes $O(K \log(1/\delta))$ queries to O_S, O_f and prepares an $(m+n)$ -qubit state $|\psi\rangle$ encoding $g_S(x, y)$ for $(x, y) \in \mathbb{F}_2^{m+n}$

$$|\psi\rangle = \frac{1}{\sqrt{2^n|S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} (-1)^{f(x) \cdot y} |x, y\rangle,$$

using $O(Kn \log(1/\delta))$ time to prepare one copy of $|\psi\rangle$ with probability at least $1 - \delta$.

Proof. First, given quantum query access to S , the algorithm prepares

$$\frac{1}{\sqrt{2^m}} \sum_{x \in \mathbb{F}_2^m} |x, 0\rangle \xrightarrow{O_S} \frac{1}{\sqrt{2^m}} \sum_{x \in \mathbb{F}_2^m} |x, \mathbf{1}_S(x)\rangle,$$

and measures the second register. With probability $|S|/2^m \geq 1/K$, the algorithm obtains 1 in which case the resulting state is $|S\rangle = \frac{1}{\sqrt{|S|}} \sum_{x \in S} |x\rangle$. So making $O(K \log(1/\delta))$ quantum queries, one can prepare $|S\rangle$ with probability $\geq 1 - \delta$. Then algorithm then simply performs the following

$$\frac{1}{\sqrt{|S|}} \sum_{x \in S} |x\rangle \otimes \frac{1}{\sqrt{2^n}} \sum_{y \in \mathbb{F}_2^n} |y\rangle \xrightarrow{O_f} \frac{1}{\sqrt{2^n|S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} |x, y, f(x)\rangle \tag{13}$$

$$\rightarrow \frac{1}{\sqrt{2^n|S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} |x, y, f(x)\rangle \otimes_{i=1}^n |f(x)_i \cdot y_i\rangle \tag{14}$$

$$\rightarrow \frac{1}{\sqrt{2^n|S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} |x, y\rangle |f(x)\rangle \otimes_{i=1}^n |f(x)_i \cdot y_i\rangle |f(x) \cdot y\rangle. \tag{15}$$

where the second operation is by applying n many CCNOT gates with the control qubits being $y_i, f(x)_i$ applied onto the target qubit $|0\rangle_i$, and the third operation is by applying n CNOT gate

between the control qubit $|f(x)_i \cdot y_i\rangle$ and target qubit $|0\rangle$. After obtaining the state in Eq. (15), the algorithm applies a single-qubit Hadamard on the last qubit and measures in the computational basis. If it is 1, it continues. First note that if the last qubit was 1, then the resulting quantum state is

$$\frac{1}{\sqrt{2^n |S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} (-1)^{f(x) \cdot y} |x, y\rangle |f(x)\rangle \otimes_{i=1}^n |f(x)_i \cdot y_i\rangle |1\rangle. \quad (16)$$

Furthermore the probability of obtaining 1 is exactly $1/2$. Upon succeeding, the algorithm inverts the n many CCNOT gates and the query operator O_f to obtain the state

$$|\psi\rangle = \frac{1}{\sqrt{2^n |S|}} \sum_{\substack{x \in S, \\ y \in \mathbb{F}_2^n}} (-1)^{f(x) \cdot y} |x, y\rangle.$$

Overall, the algorithm used $O(Kn \log(1/\delta))$ time and $O(K \log(1/\delta))$ queries to prepare $|\psi\rangle$. \square

We are now ready to prove Lemma 4.6.

Proof of Lemma 4.6. The proof will be similar to the classical proof in Lemma 3.4. Similar to the proof there, we are guaranteed by Claim 3.5 that there exists a *quadratic* polynomial $q : \mathbb{F}_2^n \times \mathbb{F}_2^n \rightarrow \{0, 1\}$ which has high correlation with $g_S(x, y) = \mathbf{1}_S(x) (-1)^{f(x) \cdot y}$, i.e.,

$$\left| \mathbb{E}_{x, y \in \mathbb{F}_2^n \times \mathbb{F}_2^n} [\mathbf{1}_S(x) (-1)^{f(x) \cdot y} (-1)^{q(x, y)}] \right| \geq \frac{1}{P_4(K)},$$

where $P_4(\cdot)$ is the polynomial promised by Lemma 3.3. For simplicity in notation, let us denote $\sigma := 1/P_4(K)$. In particular, defining the quantum states

$$|\psi\rangle = \frac{1}{\sqrt{2^n |S|}} \sum_{x \in S, y \in \mathbb{F}_2^n} (-1)^{f(x) \cdot y} |x, y\rangle, \quad |\phi_q\rangle = \frac{1}{\sqrt{2^{n+m}}} \sum_{x \in \mathbb{F}_2^n, y \in \mathbb{F}_2^m} (-1)^{q(x, y)} |x, y\rangle,$$

we have that $|\langle \psi | \phi_q \rangle|^2 \geq \sigma^2$. Moreover by Theorem 2.8, we note that the quantum state $|\phi_q\rangle$ is a *stabilizer state*⁶, and thus the stabilizer fidelity of $|\psi\rangle$ is also at least σ^2 . We now use Theorem 2.9 on copies of $|\psi\rangle$ prepared using Claim 4.7, with the error instantiated as $\varepsilon = \sigma^2/2$ to learn a stabilizer state $|s\rangle$ such that $|\langle s | \psi \rangle|^2 \geq \sigma^2/2$. By Theorem 2.8, we can write this stabilizer state as

$$|s\rangle = \frac{1}{\sqrt{|A_s|}} \sum_{z \in A_s} i^{\ell_s(z)} (-1)^{q_s(z)} |z\rangle, \quad (17)$$

where $A_s \subseteq \mathbb{F}_2^{n+m}$ is an affine subspace, ℓ_s is a linear polynomial and q_s is a quadratic form. To

⁶We remark that $|\phi_q\rangle$ is in fact a degree-2 phase state (i.e., the subspace is \mathbb{F}_2^{m+n} and there are no complex phases), but we will not use that here.

lower bound the size of $|A_s|$, we will lower bound the size of $|A_s \cap T|$ (denoting $T := S \times \mathbb{F}_2^n$):

$$\begin{aligned} \frac{\sigma}{\sqrt{2}} &\leq |\langle \psi | s \rangle| = \left| \frac{1}{\sqrt{|A_s| \cdot |S| \cdot 2^n}} \sum_{\substack{x \in S \\ y \in \mathbb{F}_2^n \\ (x,y) \in A_s}} i^{\ell(x,y)} (-1)^{q(x,y)+f(x) \cdot y} \right| \\ &\leq \frac{1}{\sqrt{|A_s| \cdot |S| \cdot 2^n}} \sum_{(x,y) \in A_s \cap T} \left| i^{\ell(x,y)} (-1)^{q(x,y)+f(x) \cdot y} \right| \\ &\leq \frac{\sqrt{|A_s \cap T|}}{\sqrt{|S| \cdot 2^n}}, \end{aligned}$$

where we have used the triangle inequality in the second line and noted that each internal term is at most 1 in the final inequality along with using $|A_s| \geq |A_s \cap T|$. The above result implies that $|A_s|$ is large i.e.,

$$|A_s| \geq |A_s \cap T| \geq (\sigma^3/2)2^{m+n}, \quad (18)$$

as $|S| \geq \sigma \cdot 2^m$. Writing $A_s = a + H_s$ where H_s is a linear subspace, we then have $\text{codim}(H_s) \leq \log(2/\sigma^3)$. To obtain a quadratic phase state $|\phi_p\rangle$ corresponding to a quadratic phase polynomial $p: \mathbb{F}_2^{n+m} \rightarrow \mathbb{F}_2$ that has high fidelity with $|\psi\rangle$ from the description of $|s\rangle$, we have the following observations (similar to that in [BC25, Proof of Theorem 1.1]) and which will inform our approach.

Let us denote the orthogonal complement of H_s as $H_s^\perp = \{x \in \mathbb{F}_2^{n+m} : x \cdot h = 0, \forall h \in H_s\}$. The Fourier decomposition of $\mathbf{1}_{A_s}(x)$ is given by

$$\mathbf{1}_{A_s}(x) = \frac{|H_s|}{|2^{n+m}|} \sum_{\lambda \in H_s^\perp} (-1)^{\lambda \cdot (a+x)}, \quad (19)$$

which follows from the observation that

$$\begin{aligned} \mathbb{E}_x[\mathbf{1}_{A_s}(x)(-1)^{\lambda \cdot x}] &= 2^{-(n+m)} \sum_{x \in H_s} (-1)^{\lambda \cdot (a+x)} = |H_s| 2^{-(n+m)} (-1)^{\lambda \cdot a} \mathbb{E}_{x \in H_s} [(-1)^{\lambda \cdot x}] \\ &= |H_s| 2^{-(n+m)} (-1)^{\lambda \cdot a} [\lambda \in H_s^\perp]. \end{aligned}$$

We then observe that

$$\sigma/\sqrt{2} \leq |\langle \psi | s \rangle| = \frac{1}{\sqrt{|A_s| \cdot |S| \cdot 2^n}} \left| \sum_{z \in \mathbb{F}_2^{n+m}} \mathbf{1}_{A_s}(z) g_S(z) (-1)^{q_s(z)} i^{\ell_s(z)} \right| \quad (20)$$

$$= \frac{|H_s|}{\sqrt{|A_s| \cdot |S| \cdot 2^n}} \left| \sum_{\lambda \in H_s^\perp} \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} (-1)^{\lambda \cdot (a+z)} g_S(z) (-1)^{q_s(z)} i^{\ell_s(z)} \right| \quad (21)$$

$$\leq \frac{|H_s| \cdot |H_s^\perp|}{\sqrt{|A_s| \cdot |S| \cdot 2^n}} \max_{\lambda \in H_s^\perp} \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} (-1)^{\lambda \cdot z} g_S(z) (-1)^{q_s(z)} i^{\ell_s(z)} \right| \quad (22)$$

$$\leq \frac{\sqrt{2}}{\sigma^2} \max_{\lambda \in H_s^\perp} \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} (-1)^{\lambda \cdot z} g_S(z) (-1)^{q_s(z)} i^{\ell_s(z)} \right|, \quad (23)$$

where we used the Fourier decomposition of $\mathbf{1}_{A_s}(z)$ from Eq. (19) in the second line, applied the triangle inequality along with considering the $\lambda \in H_s^\perp$ which maximizes the expectation in the third

line, and finally used Eq. (18) as well as noting $|H_s| \cdot |H_s^\perp| = 2^{n+m}$ and $|S| \geq \sigma \cdot 2^m$. From Eq. (23), we have that $\exists \lambda^* \in H_s^\perp$ such that

$$\left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} g_S(z) (-1)^{q_s(z) + \lambda^* \cdot z} i^{\ell_s(z)} \right| \geq \sigma^3/2.$$

Define the function $h(z) := g_S(z) (-1)^{q_s(z) + \lambda^* \cdot z}$. Additionally, we denote $R_h = \text{Re}(\mathbb{E}_z[h(z)])$ to be the real part of the above expectation and $I_h = \text{Im}(\mathbb{E}_z[h(z)])$ to be the imaginary part of the expectation. Now, we consider the two candidate quadratic polynomials $p_0(z) := q_s(z) + \lambda^* \cdot z$ and $p_1(z) := q_s(z) + \lambda^* \cdot z + \ell_s(z)$, where q_s and ℓ_s are the quadratic and linear polynomials corresponding to the stabilizer state $|s\rangle$ in hand (Eq. (17)). We observe that the quadratic phase states $|\phi_{p_0}\rangle$ and $|\phi_{p_1}\rangle$ satisfy

$$\begin{aligned} |\langle \psi | \phi_{p_0} \rangle| &= \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z)] \right| = \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z) \mathbf{1}_{\ell_s(z)=0}] + \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z) \mathbf{1}_{\ell_s(z)=1}] \right| = |R_h + I_h| \\ |\langle \psi | \phi_{p_1} \rangle| &= \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z) (-1)^{\ell_s(z)}] \right| = \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z) \mathbf{1}_{\ell_s(z)=0}] - \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} [h(z) \mathbf{1}_{\ell_s(z)=1}] \right| = |R_h - I_h| \end{aligned}$$

Noting that $\max\{|a+b|, |a-b|\} = |a|+|b| \geq \sqrt{|a|^2+|b|^2}$, we then have

$$\max\{|\langle \psi | \phi_{p_0} \rangle|, |\langle \psi | \phi_{p_1} \rangle|\} \geq \left| \mathbb{E}_{z \in \mathbb{F}_2^{n+m}} g_S(z) (-1)^{q_s(z) + \lambda^* \cdot z} i^{\ell_s(z)} \right| \geq \sigma^3/2. \quad (24)$$

In other words, one of the quadratic polynomials p_0 or p_1 has high correlation with $g_S(z)$.

To determine this quadratic polynomial, we now use the following approach. We create the list of candidate quadratic polynomials L where we add the polynomials $p_0^\lambda(z) := q_s(z) + \lambda \cdot z$ and $p_1^\lambda(z) := q_s(z) + \lambda \cdot z + \ell_s(z)$ for each $\lambda \in H_s^\perp$. This list will be of size $|L| = 2|H_s^\perp| \leq 4/\sigma^3$, where we have used $\text{codim}(H_s) \leq \log(2/\sigma^3)$. For each $p \in L$, we prepare copies of the quadratic phase state $|\phi_p\rangle$ (which is also a stabilizer state) using Lemma 2.11 and then measure $|\langle \psi | \phi_p \rangle|^2$ using the SWAP test (Lemma 2.10) up to error $\sigma^3/4$ and output the quadratic polynomial p^* that maximizes the fidelity. This consumes $\text{poly}(1/\sigma)$ sample complexity and $O(n^2/\log n \cdot \text{poly}(1/\sigma))$ time complexity. We are guaranteed by Eq. (24) that $(-1)^{p^*}$ satisfies

$$\left| \mathbb{E}_{(x,y) \in \mathbb{F}_2^m \times \mathbb{F}_2^n} \mathbf{1}_S(x) (-1)^{p^*(x,y) + f(x) \cdot y} \right| \geq \frac{\sigma^3}{2} = \frac{1}{2P_4(K)^3}, \quad (25)$$

where we have substituted back $\sigma = 1/P_4(K)$ set earlier. Having determined the polynomial p^* , we now proceed as in Lemma 3.4 to determine the affine linear function φ that agrees with ϕ on many values $x \in S$. This completes the proof of the lemma. The main contribution to query complexity and time complexity is utilizing Theorem 2.9. \square

With this lemma, we are finally ready to prove the main theorem in this section.

Proof of Theorem 4.5. The proof of this theorem is very similar to the classical proof. We write the details below:

1. Sample $t = 28 \log|A| + 56K$ many uniformly random elements from A , and denote their linear span by U . Let $A' := A \cap U$.
2. Take a random linear map $\pi : U \rightarrow \mathbb{F}_2^m$ where $m = \log|A| + 4 \log K + 10$. Let $S = \pi(A')$ denote the image of A' under π , and let $f : S \rightarrow U$ be the inverse of π when restricted to S .

3. Apply Lemma 4.6 to obtain an affine-linear map $\psi : \mathbb{F}_2^m \rightarrow U$ such that $f(x) = \psi(x)$ for at least $|A|/P'_4(2^{33}K^{13})$ values $x \in S$.
4. Take a subspace V of $\text{Im}(\psi)$ having size at most $|A|$, and output a basis for V .

The only difference between the classical and quantum algorithm is in Step (3). So, we do not reproduce the correctness analysis and refer the reader to the classical proof of Theorem 1.2.

Overall, the complexity of the algorithm is as follows. The sample complexity to the set A is $O(K + \log|A|)$, as given in step (1). Computing $\ker(\pi) \leq U$ and $\pi^{-1} : \text{Im}(\pi) \rightarrow U/\ker(\pi)$ takes $O(m^2n)$ time and, after this is done, each query to S and f takes 2^{2K} queries to A and $O(mn)$ time. The total number of queries to A needed to apply Lemma 4.6 is then

$$2^{2K} \cdot K^{O(\log K)}(m + \log|U|) = 2^{O(K)} \log|A|,$$

where we used that $m = \log|A| + O(\log K)$ and $\log|U| = O(\log|A| + K)$. The total runtime is the cost of Lemma 4.6, the cost of inverting π and the cost of making queries to S and f , i.e.

$$K^{O(\log K)}(m + \log|U|)^3 + O(m^2n) + K^{O(\log K)}(m + \log|U|) \cdot O(mn) = K^{O(\log K)}n^3,$$

concluding the proof of the theorem. □

5 Lower bounds

In this section, we show that the query complexities of the classical and quantum algorithms presented in Section 4 are essentially optimal in terms of their dependence on n . Our lower bounds are a simple information-theoretic argument (similar to the argument in [Mon12, Proposition 1]). We remark that both our lower bounds will consider the *hard instances* to be *dense* sets $A \subseteq \mathbb{F}_2^n$ (i.e., $|A|/2^n \geq \Omega(1)$). In this case the classical and quantum algorithm need not use random samples from A at all: observe that randomly sampled points will lie in A (which can be verified by queries) with constant probability, so with a constant overhead in complexity one can remove the need for samples in the classical algorithms.

5.1 Classical lower bound

We first show that the classical algorithm is essentially optimal in terms of the dependence in n .

Theorem 5.1. *Suppose $A \subseteq \mathbb{F}_2^n$ satisfies $|A + A| \leq K|A|$. Suppose an algorithm makes t queries to O_A and, with non-negligible probability, identifies a subspace $H \leq \mathbb{F}_2^n$ that covers A by $(2K)^C$ cosets for some constant $C > 0$. Then $t = \Omega_K(n^2)$.*

Proof. The proof follows from letting A itself be a subspace of \mathbb{F}_2^n (clearly subspaces have doubling constant 1). Let \mathbf{H} be a uniformly random subspace in \mathbb{F}_2^n . Given query access to $H \leq \mathbb{F}_2^n$, the algorithm needs to output a subspace H' such that $(2K)^C$ translates of it cover H . In particular, this implies that one learns a subspace $H \cap H'$ whose size is at least $2^n/(2K)^C$ (i.e., has dimension at least $n - C \log(2K)$).

Now, observe that every classical query provides one bit of information, i.e., on input $x \in \mathbb{F}_2^n$, outputs if $x \in H$ or not. Let the random variable \mathbf{H}' be the random variable corresponding to the output of the algorithm after making t queries to O_H . Then we have that

$$I(\mathbf{H} : \mathbf{H}') \leq t,$$

since the algorithm makes t queries each giving one bit of information. Furthermore since the goal of the algorithm is to obtain $\mathbf{H} \cap \mathbf{H}'$ which is a subspace of dimension $n - C \log(2K)$, by Fano's inequality [Cov99] we have that

$$\Pr[\text{identification}] \geq 1 - \frac{I(\mathbf{H} : \mathbf{H}')}{\log|\text{supp}(\mathbf{H})|} \geq 1 - \frac{t}{\log(2^{\Theta((n-C \log(2K))^2)})} = 1 - O\left(\frac{t}{(n - C \log(2K))^2}\right),$$

where we used that the number of subspaces in \mathbb{F}_2^n is $2^{\Theta(n^2)}$. The above implies that $t = \Omega_K(n^2)$ for an algorithm that identifies \mathbf{H} with non-negligible probability. \square

5.2 Quantum lower bound

We now show that the quantum algorithm is optimal in terms of the dependence in n .

Theorem 5.2. *Suppose $A \subseteq \mathbb{F}_2^n$ satisfies $|A + A| \leq K|A|$. Suppose an algorithm makes t quantum queries to O_A and, with non-negligible probability, identifies a subspace $H \leq \mathbb{F}_2^n$ that covers A by $(2K)^C$ cosets for some constant $C > 0$. Then $t = \Omega_K(n)$.*

Proof. Like in the classical proof, we let A itself be a subspace of \mathbb{F}_2^n . Let \mathbf{H} be a uniformly random subspace in \mathbb{F}_2^n . Recall from the previous proof that one needs to learn a subspace $H \cap H'$ whose size is at least $2^n/(2K)^C$ (i.e., has dimension at least $n - C \log(2K)$). More formally, let \mathbf{H} be a uniformly random subspace in \mathbb{F}_2^n . Given quantum query access to $H \subseteq \mathbb{F}_2^n$, suppose the algorithm outputs \mathbf{H}' . A quantum query algorithm can be viewed as a communication protocol, that on input $|x, 0\rangle$ outputs $|x, \mathbf{1}_H(x)\rangle$, which uses $n + 1$ qubits of communication in total. Let the random variable \mathbf{H}' be the random variable corresponding to the output of the algorithm after making t queries to O_H . Then by Holevo's theorem, we have that

$$I(\mathbf{H} : \mathbf{H}') \leq 2t(n + 1),$$

since the communication involves sending $n + 1$ qubits back and forth, in total t many times. Furthermore since the goal of the algorithm is to *identify* \mathbf{H} , by Fano's inequality [Cov99] we have that

$$\Pr[\text{identification}] \geq 1 - \frac{I(\mathbf{H} : \mathbf{H}')}{\log|\text{supp}(\mathbf{H})|} \geq 1 - \frac{2tn}{\log(2^{\Theta(n^2)})} = 1 - O(t/n),$$

where we used that the number of subspaces in \mathbb{F}_2^n is $2^{\Theta(n^2)}$. The above implies that $t = \Omega(n)$ for an algorithm that identifies \mathbf{H} with non-negligible probability. \square

References

- [AD25a] Srinivasan Arunachalam and Arkopal Dutt. Learning stabilizer structure of quantum states, 2025. Manuscript.
- [AD25b] Srinivasan Arunachalam and Arkopal Dutt. Polynomial-time tolerant testing stabilizer states. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1234–1241. Association for Computing Machinery, 2025.
- [ADL14] Divesh Aggarwal, Yevgeniy Dodis, and Shachar Lovett. Non-malleable codes from additive combinatorics. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, page 774–783. Association for Computing Machinery, 2014.

- [AGG⁺24] Vahid R Asadi, Alexander Golovnev, Tom Gur, Igor Shinkar, and Sathyawageeswar Subramanian. Quantum worst-case to average-case reductions for all linear problems. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2535–2567. SIAM, 2024.
- [AGGS22] Vahid R. Asadi, Alexander Golovnev, Tom Gur, and Igor Shinkar. Worst-case to average-case reductions via additive combinatorics. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 1566–1574. Association for Computing Machinery, 2022.
- [BC25] Jop Briët and Davi Castro-Silva. A near-optimal Quadratic Goldreich-Levin algorithm. *arXiv preprint arXiv:2505.13134*, 2025.
- [BDL13] Abhishek Bhowmick, Zeev Dvir, and Shachar Lovett. New bounds for matching vector families. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, page 823–832. Association for Computing Machinery, 2013.
- [BNOŽ25] Benjamin Bedert, Tamio-Vesa Nakajima, Karolina Okrasa, and Stanislav Živný. Strong sparsification for 1-in-3-SAT via Polynomial Freiman-Ruzsa. *arXiv preprint arXiv:2507.17878*, 2025.
- [BSLRZ14] Eli Ben-Sasson, Shachar Lovett, and Noga Ron-Zewi. An additive combinatorics approach relating rank to communication complexity. *Journal of the ACM (JACM)*, 61(4):1–18, 2014.
- [BSRZTW14] Eli Ben-Sasson, Noga Ron-Zewi, Madhur Tulsiani, and Julia Wolf. Sampling-based proofs of almost-periodicity results and algorithmic applications. In *International Colloquium on Automata, Languages, and Programming*, pages 955–966. Springer, 2014.
- [BvdDH25] Zongbo Bao, Philippe van Dordrecht, and Jonas Helsen. Tolerant testing of stabilizer states with a polynomial gap via a generalized uncertainty relation. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 1254–1262. Association for Computing Machinery, 2025.
- [CGYZ25] Sitan Chen, Weiyuan Gong, Qi Ye, and Zhihan Zhang. Stabilizer bootstrapping: A recipe for efficient agnostic tomography and magic estimation. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 429–438. Association for Computing Machinery, 2025.
- [Cov99] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.
- [DDM03] Jeroen Dehaene and Bart De Moor. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Phys. Rev. A*, 68:042318, Oct 2003.
- [DdW09] Andrew Drucker and Ronald de Wolf. Quantum proofs for classical theorems. *arXiv preprint arXiv:0910.3376*, 2009.
- [EZ12] Chaim Even-Zohar. On sums of generating sets in \mathbb{Z}_2^n . *Combinatorics, probability and computing*, 21(6):916–941, Nov 2012.
- [Fre87] Gregory A Freiman. What is the structure of k if $k + k$ is small? *Number Theory*, page 109, 1987.

- [GGMT25] William Timothy Gowers, Ben Green, Freddie Manners, and Terence Tao. On a conjecture of Marton. *Annals of Mathematics*, 201(2):515–549, 2025.
- [Gre04a] Ben Green. Finite field models in additive combinatorics. *arXiv preprint math/0409420*, 2004.
- [Gre04b] Ben Green. Notes on the polynomial Freiman-Ruzsa conjecture. *Available at <https://people.maths.ox.ac.uk/greenbj/papers/PFR.pdf>*, 2004.
- [Gre05] Ben Green. Notes on the polynomial Freiman–Ruzsa conjecture. *preprint*, 2005.
- [GT10] Ben Green and Terence Tao. An equivalence between inverse sumset theorems and inverse conjectures for the u_3 norm. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 149, pages 1–19. Cambridge University Press, 2010.
- [KLT23] Dain Kim, Anqi Li, and Jonathan Tidor. Cubic Goldreich-Levin. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4846–4892. SIAM, 2023.
- [Lov12] Shachar Lovett. Equivalence of polynomial conjectures in additive combinatorics. *Combinatorica*, 32(5):607–618, 2012.
- [Lov15] Shachar Lovett. An exposition of Sanders’ quasi-polynomial Freiman-Ruzsa theorem. *Theory of Computing*, pages 1–14, 2015.
- [Mon12] Ashley Montanaro. The quantum query complexity of learning multilinear polynomials. *Information Processing Letters*, 112(11):438–442, 2012.
- [MT25] Saeed Mehraban and Mehrdad Tahmasbi. Improved bounds for testing low stabilizer complexity states. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC ’25, page 1222–1233. Association for Computing Machinery, 2025.
- [Nes08] M Nest. Classical simulation of quantum computation, the Gottesman-Knill theorem, and slightly beyond. *arXiv preprint arXiv:0811.0898*, 2008.
- [PMH03] Ketan N Patel, Igor L Markov, and John P Hayes. Efficient synthesis of linear reversible circuits. *arXiv preprint quant-ph/0302002*, 2003.
- [Ruz99] Imre Ruzsa. An analog of Freiman’s theorem in groups. *Astérisque*, 258(199):323–326, 1999.
- [Sam07] Alex Samorodnitsky. Low-degree tests at large distances. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC ’07, page 506–515. Association for Computing Machinery, 2007.
- [TV06] Terence Tao and Van H Vu. *Additive combinatorics*, volume 105. Cambridge University Press, 2006.
- [TW14] Madhur Tulsiani and Julia Wolf. Quadratic Goldreich–Levin theorems. *SIAM Journal on Computing*, 43(2):730–766, 2014.
- [ZBS11] Noga Zewi and Eli Ben-Sasson. From affine to two-source extractors via approximate duality. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*, STOC ’11, page 177–186. Association for Computing Machinery, 2011.