Electronic Colloquium on Computational Complexity, Report No. 178 (2023)

NLTS Hamiltonians and Strongly-Explicit SoS Lower Bounds from Low-Rate Quantum LDPC Codes

Louis Golowich
UC Berkeley
lgolowich@berkeley.edu

Tali Kaufman
Bar-Ilan University
kaufmant@mit.edu

November 15, 2023

Abstract

Recent constructions of the first asymptotically good quantum LDPC (qLDPC) codes led to two breakthroughs in complexity theory: the NLTS (No Low-Energy Trivial States) theorem (Anshu, Breuckmann, and Nirkhe, STOC'23), and explicit lower bounds against a linear number of levels of the Sum-of-Squares (SoS) hierarchy (Hopkins and Lin, FOCS'22).

In this work, we obtain improvements to both of these results using qLDPC codes of *low rate*:

- Whereas Anshu et al. only obtained NLTS Hamiltonians from qLDPC codes of linear dimension, we show the stronger result that qLDPC codes of arbitrarily small positive dimension yield NLTS Hamiltonians.
- The SoS lower bounds of Hopkins and Lin are only weakly explicit because they require running Gaussian elimination to find a nontrivial codeword, which takes polynomial time. We resolve this shortcoming by introducing a new method of planting a strongly explicit nontrivial codeword in linear-distance qLDPC codes, which in turn yields strongly explicit SoS lower bounds.

Our "planted" qLDPC codes may be of independent interest, as they provide a new way of ensuring a qLDPC code has positive dimension without resorting to parity check counting, and therefore provide more flexibility in the code construction.

1 Introduction

Recent breakthrough constructions of asymptotically good quantum LDPC (qLDPC) codes [PK22, LZ22, DHLV23] have led to major advances in complexity theory. Specifically, Anshu et al. [ABN23] applied these codes to prove the NLTS theorem, which provides perhaps the most significant progress to date towards the quantum PCP conjecture. Meanwhile, Hopkins and Lin [HL22] applied the same codes to obtain the first explicit lower bounds against a linear number of levels of the Sum-of-Squares semidefinite programming (SoS SDP) hierarchy, which is one of the most powerful algorithmic frameworks for approximating the satisfiability of constraint satisfaction problems (CSPs).

In this paper, we improve upon both of these complexity theoretic results. Along the way, we introduce a new method for ensuring a qLDPC code has positive dimension, which may be of independent interest. Our contributions are therefore threefold:

ISSN 1433-8092

1. **NLTS Hamiltonians from low-rate codes:** The breakthrough construction of NLTS Hamiltonians of [ABN23] from asymptotically good qLDPC codes relied on both the linear dimension and distance of the codes. A promising approach [Nir23] for further progress towards qPCP is to construct more general NLTS Hamiltonians with additional properties. We make progress in this direction by constructing NLTS Hamiltonians from qLDPC codes of arbitrary positive dimension, thereby removing the linear-dimension requirement in [ABN23]. Our result highlights the usefulness of local Hamiltonians with low-dimensional ground spaces for studying qPCP. Our proof leverages techniques of [EH17], which conjecturally constructed NLTS Hamiltonians from linear-distance quantum locally testable codes of arbitrary positive dimension (which are not known to exist). However, we obtain the NLTS property without assuming local testability nor linear dimension. Instead, the key ingredient ensuring NLTS Hamiltonians is a small-set expansion property of the qLDPC codes.

2. **Planted quantum LDPC codes:** We show how to plant an explicit nontrivial codeword in a linear-distance qLDPC code, which may have otherwise had rate 0. To the best of our knowledge, this construction yields the first linear-distance qLDPC codes for which nontrivial dimension is established without resorting to parity-check counting. It has been an open question in the literature to develop new such techniques for bounding dimension (see for instance Section 1.1 of [DLZ23], and also [DDHRZ20]).

3. **Strongly explicit SoS lower bounds:** We apply our planted qLDPC codes to obtain the first *strongly* explicit family of CSPs that cannot be refuted by a linear number of levels of the SoS hierarchy. This result strengthens the work of [HL22], which provided the first *weakly* explicit construction of such an SoS lower bound using qLDPC codes. Our improvement stems from the fact that our planted codes have planted codeword given by the all-1s vector, which is strongly explicit.

These results together show new ways to both construct and apply qLDPC codes of low rate. In the remainder of this section, after providing some background on qLDPC codes, we describe each of these results in more depth. We then discuss open questions that arise from our results.

1.1 Background on qLDPC Codes

This section provides some definitions we will need to state our results. The quantum codes we consider in this paper are quantum CSS codes. An n-qudit CSS code $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ of alphabet size (i.e. local dimension) q is defined by a pair of classical codes $C_X, C_Z \subseteq \mathbb{F}_q^n$ such that $C_X^\perp \subseteq C_Z$. The associated quantum code is then given by $\mathcal{C} = \mathrm{span}\{\sum_{y' \in C_X^\perp} |y + y'\rangle : y \in C_Z\}$. This code has dimension $k = \dim(C_Z) - \dim(C_X^\perp)$ and distance $d = \min_{y \in (C_Z \setminus C_X^\perp) \cup (C_X \setminus C_Z^\perp)} |y|$, meaning it encodes a k-qudit message into an n-qudit code state, and the message can be recovered from any $n-(d-1)$ code qudits. We assume $C_X = \ker H_X, C_Z = \ker H_Z$ for associated parity check matrices $H_X \in \mathbb{F}_q^{m_X \times n}, H_Z \in \mathbb{F}_q^{m_Z \times n}$. If every row and colum of H_X and H_Z has Hamming weight $\leq \ell$, we say that $\mathcal{C}$ has locality ℓ. A family of qLDPC codes is a family of codes with constant locality ℓ and growing block length n.

It was a longstanding open question to construct linear-distance qLDPC codes. This question was resolved by Panteleev and Kalachev [PK22], who obtained qLDPC codes of linear distance and linear dimension. Subsequent works [LZ22, DHLV23] provided additional related constructions.

These codes in fact possess[1] the following stronger notion of distance, which guarantees that all low-weight errors have syndromes whose weight is linear in the error weight (as opposed to just having nonzero syndromes). Below, for a code C, we denote $|y|_C = \min_{y' \in C} |y + y'|$.

Definition 1 (Small-set (co)boundary expansion; restatement of Definition 16)**.** Let $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ be a CSS code given by parity check matrices $H_X \in \mathbb{F}_q^{m_X \times n}$ and $H_Z \in \mathbb{F}_q^{m_Z \times n}$. For $c_1, c_2 > 0$, we say that $\mathcal{C}$ has (c_1, c_2)**-small-set boundary expansion** if it holds for every $y \in \mathbb{F}_q^n$ with $|y| \leq c_1 n$ that

$$\frac{|H_Z y|}{m_Z} \geq c_2 \frac{|y|_{C_X^\perp}}{n}.$$

Similarly, $\mathcal{C}$ has (c_1, c_2)**-small-set coboundary expansion** if it holds for every $y \in \mathbb{F}_q^n$ with $|y| \leq c_1 n$ that

$$\frac{|H_X y|}{m_X} \geq c_2 \frac{|y|_{C_Z^\perp}}{n}.$$

This notion of small-set (co)boundary expansion underlies both the NLTS Hamiltonians of [ABN23] and the SoS lower bounds of [HL22]. Note that a code with (c_1, c_2)-small set boundary and coboundary expansion by definition has distance $\geq c_1 n$.

1.2 NLTS Hamiltonians from Low-Rate qLDPC Codes

The quantum PCP (qPCP) conjecture, which states that it is QMA-hard to compute a constant-factor approximation to the ground energy of a local Hamiltonian, is a major open question in quantum complexity theory that has remained largely elusive. Perhaps the most significant progress towards this conjecture was the NLTS theorem, which was recently proven by Anshu, Breuckmann, and Nirkhe [ABN23] using an application of asymptotically good qLDPC codes. This result provides a family of local Hamiltonians that have "no low-energy trivial states" (NLTS), where a trivial state is one computed by a constant-depth circuit. The NLTS theorem therefore provides local Hamiltonians exhibiting a weaker form of hardness of approximation than required by qPCP, and is indeed a necessary consequence of the qPCP conjecture under the widely believed assumption that NP $\neq$ QMA.

Anshu et al. [ABN23] constructed their NLTS Hamiltonians using the asymptotically good quantum Tanner codes of [LZ22]. In particular, their proof of NLTS relied on the codes having both linear distance and dimension. It was an open question whether such linear dimension was necessary for NLTS. This question is motivated by the suggestion [Nir23] that constructing more general families of NLTS Hamiltonians may lead to further progress towards the qPCP conjecture. Furthermore, some earlier partial progress towards NLTS used codes of smaller dimension [EH17], which again raises the question of whether linear dimension is necessary. Our main result on NLTS resolves this question, as we obtain NLTS Hamiltonians from qLDPC codes of arbitrary positive dimension.

NLTS Hamiltonians are formally defined as follows. Recall that a family of Hamiltonians is *ℓ-local* if every $\mathbf{H}$ in the family can be expressed as a sum of Hamiltonians, each of which act

[1] [HL22] were the first to consider small-set (co)boundary expansion for linear-distance qLDPC codes, and showed that the codes of [LZ22] possess this property. [DHLV23] later constructed additional good qLDPC codes for which they proved this expansion property. We explain at the end of Section 3.4 why the decoder of [LZ23b, LZ23a] implies that the codes of [PK22] also possess this expansion property.

nontrivially on $\leq \ell$ qubits. If $\ell = O(1)$ we say the family is *local*. We also say that a state ρ is an ϵ-*approximate ground state* of a Hamiltonian $\mathbf{H} \succeq 0$ if $\mathrm{Tr}(\rho \mathbf{H}) \leq \epsilon$.

Definition 2 (NLTS Hamiltonians)**.** A family of local Hamiltonians $(\mathbf{H}_n)_{n\to\infty}$ with $0 \preceq \mathbf{H}_n \preceq I$ is **NLTS** if there exists $\epsilon > 0$ such that the minimum depth of any quantum circuit computing an ϵ-approximate ground state of $\mathbf{H}_n$ approaches ∞ as $n \to \infty$.

Recall that for a CSS code $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$, the associated code Hamiltonian is given by

$$\mathbf{H} = \frac{1}{2}(\mathbf{H}_X + \mathbf{H}_Z)$$

for

$$\mathbf{H}_X = \frac{1}{m_X} \sum_{y \in \mathrm{rows}(H_X)} \frac{I - X^y}{2}$$
$$\mathbf{H}_Z = \frac{1}{m_Z} \sum_{y \in \mathrm{rows}(H_Z)} \frac{I - Z^y}{2},$$

where X and Z denote the respective Pauli operators. Thus in particular the ground space of $\mathbf{H}$ is precisely the code space $\mathcal{C} = \mathrm{span}\{\sum_{y' \in C_X^\perp} |y + y'\rangle : y \in C_Z\}$.

Anshu et al. [ABN23] showed that for every family of qLDPC codes with linear dimension and constant small-set boundary and coboundary expansion, the associated code Hamiltonians are NLTS. Thus for instance the quantum Tanner codes of [LZ22] yield NLTS code Hamiltonians.

Our result below improves upon this result of [ABN23] by removing the linear dimension requirement.

Theorem 3 (NLTS from low-rate codes; informal statement of Corollary 26)**.** *Let $(\mathcal{C}^{(n)})_{n\to\infty}$ be an infinite family of qLDPC codes over the alphabet*[2] *$\mathbb{F}_2$ of block length n and positive dimension that have (c_1, c_2)-small set boundary and coboundary expansion for some constants $c_1, c_2 > 0$. Then the family of associated code Hamiltonians $(\mathbf{H}^{(n)})_{n\to\infty}$ is NLTS.*

Our proof of Theorem 3 follows the general framework of [EH17, ABN23] in showing circuit lower bounds for code Hamiltonians. Specifically, Eldar and Harrow [EH17] showed that in order to show the code Hamiltonians $\mathbf{H}$ are NLTS, it suffices to show that every distribution obtained by measuring an approximate ground state of $\mathbf{H}$ in either the X or Z basis is *well spread*. Here a distribution D over $\mathbb{F}_2^n$ is well spread if there exist sets $S_0, S_1 \subseteq \mathbb{F}_2^n$ separated by a linear Hamming distance $\mathrm{dis}(S_0, S_1) \geq \Omega(n)$ such that D assigns constant probability $D(S_0), D(S_1) \geq \Omega(1)$ to both sets.

Both [EH17, ABN23] show this well-spreadness property for code Hamiltonians by combining a distance/expansion property of the code with an uncertainty principle. However, the two works different use assumptions on the code as well as different uncertainty principles:

- [EH17] assumes the code is locally testable and of linear distance, which implies the approximate ground states have a certain linear structure. They then use an uncertainty principle (see Lemma 21) that is able to leverage this linear structure and prove well-spreadness regardless of the code dimension.

[2]For simplicity we restrict attention to the binary alphabet $\mathbb{F}_2$ in our proof of Theorem 3, though we suspect the result should extend to arbitrary alphabets $\mathbb{F}_q$.

- [ABN23] assumes the code has small-set boundary and coboundary expansion, which is weaker than local testability and therefore yields less structure in the approximate ground states. They then use a different uncertainty principle with which they are still able to prove well-spreadness, but only for codes of linear dimension.

Because linear-distance quantum locally testable codes are not known to exist, the NLTS Hamiltonians of [EH17] remain conjectural.

We prove Theorem 3 by combining these two approaches: we make the weaker assumption that our code has small-set boundary and coboundary expansion, but show that the approximate ground states still have enough linear structure to apply the uncertainty principle in Lemma 21. We then conclude that the code Hamiltonians are NLTS regardless of the code dimension.

At the core of our argument is the application of a "decoding" procedure for approximate ground states of codes with small-set (co)boundary expansion, which is unintuitive in the sense that far-apart approximate ground states may decode to the the same true ground state. However, we are able to show that in some sense, the low-energy space of the code Hamiltonian acts similarly enough to a true code space that the argument still goes through.

1.3 Planted Quantum LDPC Codes

This section presents our result on planting a nontrivial codeword in qLDPC codes.

The recent breakthrough constructions of linear-distance qLDPC codes ([PK22], followed by [LZ22, DHLV23]) all bound the code dimension by counting parity checks. Specifically, these works use the fact that if $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ for $H_X \in \mathbb{F}_q^{m_X \times n}, H_Z \in \mathbb{F}_q^{m_Z \times n}$, then $\mathcal{C}$ has dimension $k \geq n - m_X - m_Z$. However, this bound may not be tight if there are redundant parity checks in H_X, H_Z. Indeed, it has been an open question in the coding theory literature to provide new ways of ensuring that LDPC codes have positive dimension; for instance, this question was of central importance in the code constructions of [DDHRZ20, DLZ23].

Our result below makes progress on this question, by showing how to plant a nontrivial codeword in the linear-distance quantum Tanner codes of [LZ22]. In fact, we show that like the codes of [LZ22] our planted codes possess small-set (co)boundary expansion.

Theorem 4 (Planted quantum Tanner codes; restatement of Theorem 32)**.** *For every finite field $\mathbb{F}_q$, there exist constants $c_1, c_2 > 0$ such that there is a strongly explicit infinite family $(\mathcal{C}^{(n)})_{n\to\infty}$ of quantum LDPC CSS codes for which every $\mathcal{C}^{(n)} = \mathrm{CSS}(C_X^{(n)}, C_Z^{(n)})$ with $C_X^{(n)}, C_Z^{(n)} \subseteq \mathbb{F}_q^n$ has the following properties:*

1. *$\mathcal{C}^{(n)}$ has (c_1, c_2)-small-set boundary and coboundary expansion, and therefore has distance $\geq c_1 n$.*
2. *The all-1s vector $\mathbf{1} \in \mathbb{F}_q^n$ lies in $C_X^{(n)} \setminus C_Z^{(n)\perp}$ and in $C_Z^{(n)} \setminus C_X^{(n)\perp}$.*

Theorem 3 implies that the code Hamiltonians of our planted quantum Tanner codes over the binary alphabet $\mathbb{F}_2$ in Theorem 4 are NLTS. In Section 1.4 below, we present another complexity-theoretic application of these codes, namely to SoS lower bounds, which crucially relies on the their planted nature.

Our construction of planted quantum Tanner codes is motivated by a more basic classical analogue. Recall that a classical Tanner code is specified by a Δ-regular graph Γ and an inner code

$C_{\text{in}} \subseteq \mathbb{F}_q^\Delta$, where the code components correspond to edges of the graph, and the parity checks impose the constraint that the local view of each vertex is a codeword in C_{in}.

The standard method for ensuring a classical Tanner code C has positive rate is to require C_{in} to have sufficiently large rate $> 1/2$, and then to bound the number of resulting linear constraints on C from the parity checks. However, we may alternatively simply require that C_{in} contain the all-1s vector $\mathbf{1} \in \mathbb{F}_q^\Delta$, so that C then must contain the global all-1s vector $\mathbf{1} \in \mathbb{F}_q^n$. If C contains no other nontrivial codewords, then it is a repetition code, which is typically unintersting classically.

However, we construct a quantum analogue of this construction, which is more nuanced, and has interesting complexity theoretic implications. Indeed, whereas classically it is easy to achieve linear distance and positive dimension by taking a repetition code, to the best of our knowledge the only known quantum LDPC codes of linear distance and positive dimension are the recent constructions of [PK22, LZ22, DHLV23], which can in fact achieve linear dimension.

Recall that a quantum Tanner code $\mathcal{C} = \text{CSS}(C_X, C_Z)$ [LZ22] is constructed by imposing constraints from a *pair* of classical codes $C_A, C_B \subseteq \mathbb{F}_q^\Delta$ on a *square Cayley complex* (V, E, Q), which is a graph (V, E) with the additional high-dimensional structure of faces, or squares, in Q; the qudits of the code correspond to the $n = |Q|$ faces in Q.

To prove Theorem 4, we show that if we require the local all-1s vector $\mathbf{1} \in \mathbb{F}_q^\Delta$ to lie in C_A and in $C_B^\perp$, and q is relatively prime with n, then the global all-1s vector $\mathbf{1} \in \mathbb{F}_q^n$ lies in $C_Z \setminus C_X^\perp$ and $C_X \setminus C_Z^\perp$, so in particular $\mathcal{C} = \text{CSS}(C_X, C_Z)$ has dimension ≥ 1.

The proof that $\mathbf{1} \in C_A, C_B^\perp$ implies $\mathbf{1} \in C_X, C_Z$ is immediate, as in the classical case. However, we prove that $\mathbf{1} \notin C_X^\perp, C_Z^\perp$ using a parity (or more generally, arity) mismatch: we argue that $C_X^\perp$ and $C_Z^\perp$ are spanned by vectors whose components sum to $0 \in \mathbb{F}_q$, whereas the components of $\mathbf{1} \in \mathbb{F}_q^n$ do not sum to 0 by the assumption that q, n are relatively prime, so that the characteristic p of $\mathbb{F}_q$ does not divide n.

The requirement that q, n are relatively prime requires some care to enforce. As $n = |Q|$ equals the number of faces in an expanding square Cayley complex, it must be a multiple of the order of a group on which there exist constant-degree Cayley expanders (see Section 3, and in particular Section 3.3, for background on Cayley graphs and expansion). Therefore if we for instance focus on the $q = 2$ case, we need families of Cayley expanders over groups of odd order. However, many well-known Cayley expanders, such as the Ramanujan graphs of [LPS88] and [Mor94], exclusively use groups of even order. We therefore instead use the Cayley expanders given in Example 3.4 of [LW93], for which the number of vertices is a power of any desired prime. While these graphs have constant degree and constant expansion, we amplify the expansion to be almost-Ramanujan using the techniques of [JMRW22].[3]

We still must show that the resulting planted quantum Tanner codes have good small-set (co)boundary expansion and therefore good distance. By the results of [LZ22], it suffices to show that the inner codes (C_A, C_B) can be chosen to possess a property called *product-expansion* (Definition 8). This property was shown for random inner codes by [KP23, DHLV23]; we extend the proof of [KP23] for our case of planted inner codes where $\mathbf{1} \in C_A, C_B^\perp$. As these inner codes are constant-sized as $n \to \infty$, the randomized construction can be made strongly explicit by a brute force search.

An interesting consequence of our result is that we can construct planted quantum Tanner codes $\mathcal{C}$ of positive dimension $k > 0$ with inner codes C_A, C_B of any desired respective rates

[3] This expansion amplification may be stronger than necessary, but for consistency with prior works and simplicity of presentation, it is convenient for us to have almost-Ramanujan expansion.

$R_A, R_B \in (0,1)$; for instance, we can take $R_A = R_B$. In contrast, the prior technique of bounding k by counting parity checks only implies that $k \geq -(1-2R_A)(1-2R_B) \cdot n$, which never gives a meaningful bound when $R_A = R_B$. Thus our construction allows instantiations in new parameter regimes.

We also remark that while we only show how to plant a nontrivial codeword in the qLDPC codes of [LZ22], our techniques also apply to the codes of [PK22]; to avoid redundancy we do not spell out the details.

1.4 Strongly Explicit SoS Lower Bounds

The Sum-of-Squares semidefinite programming hierarchy is one of the most powerful algorithmic frameworks for approximating the satisfiability of CSPs (see [FKP19] for a survey). However, almost all of the known hard instances (i.e. lower bounds) for this hierachy are given by randomized constructions. Hopkins and Lin [HL22], building on the techniques of Dinur et al. [DFHT21], constructed the first explicit unsatisfiable CSPs that cannot be refuted by a linear number of levels of the SoS SDP hierarchy. In contrast, explicit lower bounds prior to their work applied to at best a logarithmic number of levels of the SoS hierarchy.

Hopkins and Lin [HL22] proved their result by showing that hard instances for SoS can be obtained from a family of qLDPC codes with small-set boundary and coboundary expansion. Explicit such qLDPC codes, such as the quantum Tanner codes of [LZ22], then yield the desired explicit hard CSPs.

Remark 5. The SoS lower bounds of [HL22] marked the first complexity theoretic application of linear-distance qLDPC codes; the subsequent proof of the NLTS theorem [ABN23] provided a second notable application. Such applications were perhaps surprising given that the construction of asymptotically good qLDPC codes, first obtained by [PK22] and subsequently extended and modified by [LZ22, DHLV23], was originally motivated in large part by applications to quantum error correction.

However, the explicitness of the CSP construction in [HL22] was weak in the sense of Definition 6 below. One of the major questions left open by their work was to make this construction strongly explicit [Hop23]. We apply our construction of planted quantum Tanner codes in Theorem 4 to resolve this problem.

Definition 6 (Weak vs. strong explicitness)**.** Let $X = (x_n)_{n\in\mathbb{N}}$ be an infinite family of objects such that each x_n can be represented by a bitstring $x_n \in \{0,1\}^{a_n}$ of length a_n, where $a_n \to \infty$ as $n \to \infty$. We say that X is:

- **weakly explicit** (or simply "explicit") if there exist a $\mathrm{poly}(a_n)$-time algorithm $A(n)$ that outputs x_n
- **strongly explicit** if there exists a $\mathrm{poly}(\log n, \log a_n)$-time algorithm $A(n,i)$ that outputs the ith bit of x_n for $i \in [a_n]$.

We specifically say that a family of matrices is weakly (resp. strongly) explicit if for each $n \times m$ matrix in the family, there is a $\mathrm{poly}(n,m)$ (resp. $\mathrm{poly}(\log n, \log m)$) time algorithm to compute the jth nonzero entry of the ith row, as well as the jth nonzero entry of the ith column.

Then a family of graphs is weakly (resp. strongly) explicit if the associated adjacency matrices are weakly (resp. strongly) explicit. Similarly, a CSS code $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ is weakly (resp. strongly) explicit if the matrices H_X, H_Z are weakly (resp. strongly) explicit.

As another relevant example, consider a family of CSPs given by ℓ-LIN instances, which are defined by n linear constraints on m variables over a fixed finite field $\mathbb{F}_q$, such that each linear equation has $\leq \ell = O(1)$ nonzero coefficients. A family of such ℓ-LIN instances is weakly (resp. strongly) explicit if the ith linear equation can be computed in time $\mathrm{poly}(n, m)$ (resp. $\mathrm{poly}(\log n, \log m)$).

Given a qLDPC code $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ of locality ℓ and an arbitrary element $\beta \in C_X \setminus C_Z^\perp$, Hopkins and Lin [HL22] considered the associated ℓ-LIN instance $\mathcal{I}_{\mathcal{C},\beta}$ with $m = m_Z$ variables $y_1, \ldots, y_m$ and n linear constraints over $\mathbb{F}_q$ given by the system of equations $H_Z^\top y = \beta$ for $y = (y_1, \ldots, y_m)$. They showed that if $\mathcal{C}$ has $(\Omega(1), \Omega(1))$-small-set boundary and coboundary expansion, then at most $1 - \Omega(1)$ fraction of the constraints in $\mathcal{I}_{\mathcal{C},\beta}$ can be satisfied, but $\mathcal{I}_{\mathcal{C},\beta}$ is hard to refute for $\Omega(n)$ levels of SoS. Furthermore, they presented a reduction to reduce the size of the constraints, thereby providing similarly unsatisfiable and hard instances of 3-LIN over $\mathbb{F}_2$, that is, of 3-XOR.

However, even if $\mathcal{C}$ comes from a strongly explicit family of qLDPC codes, the associated ℓ-LIN instance $\mathcal{I}_{\mathcal{C},\beta}$ is only weakly explicit in general, as one must perform Gaussian elimination to compute some $\beta \in C_X \setminus C_Z^\perp$, which takes $\mathrm{poly}(n, m)$ time.

Because our planted quantum Tanner codes in Theorem 4 by construction have $\mathbf{1} \in C_X \setminus C_Z^\perp$, they resolve this issue, and hence yield the following result.

Theorem 7 (Strongly explicit SoS lower bounds; restatement of Corollary 51 and Corollary 53)**.** *The ℓ-LIN instances $\mathcal{I}_{\mathcal{C},\mathbf{1}}$ for planted quantum Tanner codes $\mathcal{C}$ over any fixed prime-sized alphabet $\mathbb{F}_p$ provide a family of strongly explicit instances such that each $\mathcal{I}_{\mathcal{C},\mathbf{1}}$:*

1. *has $\Theta(n)$ variables and constraints,*
2. *has satisfiability $\leq (1 - \Omega(1))$,*
3. *cannot be refuted by cn levels of the SoS hierarchy for a sufficiently small constant $c > 0$.*

Furthermore, there exists a strongly explicit family of 3-XOR (i.e. 3-LIN over $\mathbb{F}_2$) instances that also satisfies the three properties above.

We remark that [HL22] actually restricted attention to the binary alphabet $q = 2$ case, though their SoS lower bounds for ℓ-LIN extend to larger prime alphabets. We suspect that their reduction to 3-LIN similarly extends to larger alphabets, though for conciseness we do not check the details.

1.5 Open Questions

Our results raise the following open questions:

- Can our construction of NLTS Hamiltonians from low-rate qLDPC codes lead to more progress towards qPCP or hardness of approximation? For instance, perhaps the fact that low-rate codes, which correspond to Hamiltonians with low-dimensional ground spaces, suffice for NLTS will be helpful in constructing Hamiltonians with stronger hardness of approximation guarantees.

- Our results highlight the usefulness of low-rate qLDPC codes, and suggest that for complexity theoretic applications there is often little benefit to having high rate. However, to the best of our knowledge, our planted quantum Tanner codes provide the only known "inherently" low-rate qLDPC codes, and they still have high rate in some parameter regimes. In contrast, there are many interesting classical low-rate LDPC codes such as Hadamard and Reed-Muller codes, which have properties not shared by any high-rate codes. In the quantum case, can similar stronger properties be obtained by allowing for low rate in qLDPC codes?

2 Notation

For a string $y \in \mathbb{F}_q^n$, we denote the Hamming weight by $|y| = |\{i \in [n] : y_i \neq 0\}|$. For subsets $S, T \subseteq \mathbb{F}_q^n$, we denote the Hamming distance by $\mathrm{dis}(S,T) = \min_{s\in S, t\in T} |s-t|$.

Unless explicitly stated otherwise, by a "code" we mean a linear subspace $C \subseteq \mathbb{F}_q^n$. The code C has block length n, dimension $k = \dim_{\mathbb{F}_q}(C)$, and distance $d = \min_{y\in C\setminus\{0\}} |y|$, which can be summarized by saying it is a $[n,k,d]_q$ code. The dual code is $C^\perp = \{x \in \mathbb{F}_q^n : x \cdot y = 0 \ \forall y \in C\}$.

For codes $C_i \subseteq \mathbb{F}_q^{n_i}$ for $i = 1, 2$, the tensor code $C_1 \otimes C_2 \subseteq \mathbb{F}_q^{n_1 \times n_2}$ consists of all $n_1 \times n_2$ matrices where every column lies in C_1 and every row lies in C_2. The dual of the tensor code is $(C_1 \otimes C_2)^\perp = C_1^\perp \otimes \mathbb{F}_q^{n_2} + \mathbb{F}_q^{n_1} \otimes C_2^\perp$.

Given a Δ-regular graph Γ with n edges and an inner code $C_{\mathrm{in}} \subseteq \mathbb{F}_q^\Delta$, we denote the associated classical Tanner code by $C = \mathrm{Tan}(\Gamma, C_{\mathrm{in}}) \subseteq \mathbb{F}_q^n$, which is constructed as follows. We associate the set of all edges in Γ with the set $[n]$, and we associate the set of edges incident to each vertex v in Γ with the set $[\Delta]$. Then we define C to be the set of all edge labelings $y \in \mathbb{F}_q^n = \mathbb{F}_q^{E(\Gamma)}$ such that the labels of edges incident to each $v \in \Gamma$ form a codeword in C_{in}.

For a pure quantum state $|\psi\rangle$, we denote the density matrix by $\psi = |\psi\rangle\langle\psi|$. For a set $S \subseteq \mathbb{F}_q^n$, we let $|S\rangle = |S|^{-1/2} \sum_{s\in S} |s\rangle$ denote the uniform superposition over elements of S.

The quantum codes we consider in this paper are CSS codes, which are defined as follows. For classical codes $C_X, C_Z \subseteq \mathbb{F}_q^n$ such that $C_X^\perp \subseteq C_Z$, the associated quantum CSS code $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ is defined by $\mathcal{C} = \mathrm{span}\{|y + C_X^\perp\rangle : y \in C_Z\} \subseteq (\mathbb{C}^q)^{\otimes n}$. This code has block length n, dimension $k = \log_q \dim_{\mathbb{C}}(\mathcal{C}) = \dim_{\mathbb{F}_q}(C_Z) - \dim_{\mathbb{F}_q}(C_X^\perp)$, and distance $d = \min_{y\in(C_Z\setminus C_X^\perp)\cup(C_X\setminus C_Z^\perp)} |y|$, which can be summarized by saying that $\mathcal{C}$ is a $[[n,k,d]]_q$ code.

If $C_X = \ker H_X$ and $C_Z = \ker H_Z$ for parity check matrices H_X, H_Z in which each row and column has Hamming weight $\leq \ell$, we say that $\mathcal{C}$ is a CSS code with check weight, or locality, $\leq \ell$. A family of codes with constant locality $\ell = O(1)$ as $n \to \infty$ is said to be LDPC. The family of codes is (strongly) explicit if the associated families of parity check matrices H_X, H_Z are (strongly) explicit.

3 Review of Quantum Tanner Codes

In this section we review the construction and relevant properties of the asymptotically good quantum LDPC codes of Leverrier and Zémor [LZ22, LZ23a], which are called quantum Tanner codes. Although [LZ22, LZ23a] present the construction over binary alphabets, we consider arbitrary finite field alphabets; all their results and proofs extend to this more general case with just some '+' signs changed to '−' signs for fields of characteristic $\neq 2$.

Recall that a classical Tanner code is constructed by imposing constraints from an inner code on a graph (see Section 2). In contrast, a quantum Tanner code $\mathcal{C}$ is constructed by imposing constraints from *two* inner codes on a higher-dimensional object called a *square Cayley complex*. In particular, $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$, where both C_X, C_Z are classical Tanner codes on graphs obtained from a square Cayley complex, with distinct inner codes.

3.1 Construction

We now desribe the construction of a quantum Tanner code $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$. We first need to define a square Cayley complex. Recall that for a group G and a subset $A \subseteq G$, the Cayley graph $\mathrm{Cay}(G, A)$ has vertex set G and edge set $\{(g, ag) : g \in G,\ a \in A\}$. As described below, a square Cayley complex is a sort of 2-dimensional generalization of a Cayley graph.

A square Cayley complex consists a tuple (V, E, Q) of vertices, edges, and faces (or "squares") that is specified by a group G and two generating sets $A, B \subseteq G$ as follows. We typically take $|A| = |B| = \Delta = O(1)$ as $|G| = \Theta(n) \to \infty$, and assume that $A = A^{-1}$ and $B = B^{-1}$ are closed under inversion. The complex then has vertex set $V = G \times \{0,1\}^2$, edge set $E = E_A \sqcup E_B$ for

$$E_A = \{(g, i0), (ag, i1) : g \in G,\ i \in \{0,1\},\ a \in A\}$$
$$E_B = \{(g, 0j), (gb, 1j) : g \in G,\ j \in \{0,1\},\ b \in B\},$$

and face set

$$Q = \{(g, 00), (ag, 01), (gb, 10), (agb, 11) : g \in G,\ a \in A,\ b \in B\}.$$

For $i, j \in \{0,1\}$, let $V_{ij} = G \times (i, j)$. Define bipartite graphs $\Gamma_0 = (V_{00} \sqcup V_{11}, Q)$ and $\Gamma_1 = (V_{01} \sqcup V_{10}, Q)$ whose edges are given by pairs of vertices that form a diagonal in a square in Q; for instance, Γ_0 has an edge between $v \in V_{00}$ and $v' \in V_{11}$ if v, v' share a face in Q. Observe that both Γ_0 and Γ_1 have a unique edge associated to each square in Q. Furthermore, the Γ_i-edges incident to a vertex v correspond to the squares in Q that contain v; we let $Q(v)$ denote the set of these squares. But by definition $Q(v)$ consists of an $A \times B$ grid of squares. For instance, for $v = (g, 00) \in V_{00}$ then

$$Q(v) = \{(g, 00), (ag, 01), (gb, 10), (agb, 11) : a \in A,\ b \in B\}.$$

Therefore given a square Cayley complex (V, E, Q) of degree $\Delta = |A| = |B|$ along with classical codes $C_A \subseteq \mathbb{F}_q^A = \mathbb{F}_q^\Delta$ and $C_B \subseteq \mathbb{F}_q^B = \mathbb{F}_q^\Delta$, we may define a quantum Tanner code $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ by

$$C_X = \mathrm{Tan}(\Gamma_0, (C_A \otimes C_B)^\perp)$$
$$C_Z = \mathrm{Tan}(\Gamma_1, (C_A^\perp \otimes C_B^\perp)^\perp).$$

That is, C_X and C_Z are classical Tanner codes on the graphs Γ_0 and Γ_1 respectively, where the inner codes are given by dual tensor codes. Because $E(\Gamma_0) \cong E(\Gamma_1) \cong Q$, both C_X and C_Z are subspaces of $\mathbb{F}_q^Q$.

3.2 Locality and Dimension

We now describe some basic properties of $\mathcal{C}$. By definition $\mathcal{C}$ has block length $n = |Q|$. Parity checks for C_X are given by tensor codewords in $C_A \otimes C_B$ supported in the neighborhood $Q(v_0)$ of any $v_0 \in V_{00} \sqcup V_{11}$. Similarly, parity checks for C_Z are given by tensor codewords in $C_A^\perp \otimes C_B^\perp$ supported in the neighborhood $Q(v_1)$ of any $v_1 \in V_{01} \sqcup V_{10}$. Because any such $Q(v_0), Q(v_1)$ are either disjoint or intersect in a single row or column, the parity checks for C_X and C_Z are orthogonal, so $C_X^\perp \subseteq C_Z$. Furthermore, as $|Q(v_0)| = |Q(v_1)| = \Delta^2$, the quantum Tanner code $\mathcal{C}$ is LDPC with locality $\Delta^2 = O(1)$ as $n = |Q| \to \infty$.

Counting parity checks to bound the number of linear constraints on C_X, C_Z implies that $\mathcal{C}$ has dimension $k \geq -(1 - 2R_A)(1 - 2R_B) \cdot n$, where $R_A = \dim(C_A)/\Delta$ and $R_B = \dim(C_B)/\Delta$ denote the rate of C_A and C_B respectively.

3.3 Distance

To present the distance bound for quantum Tanner codes, we need the following definition.

Definition 8 (Product-expansion)**.** A pair of codes $C_1, C_2 \subseteq \mathbb{F}_q^n$ is **ρ-product-expanding** if every $x \in (C_1^\perp \otimes C_2^\perp)^\perp = C_1 \otimes \mathbb{F}_q^n + \mathbb{F}_q^n \otimes C_2$ can be decomposed as $x = c + r$ for some $c \in C_1 \otimes \mathbb{F}_q^n$ and $r \in \mathbb{F}_q^n \otimes C_2$ satisfying

$$|x| \geq \rho n(|c|_{\text{col}} + |r|_{\text{row}}),$$

where $|c|_{\text{col}}$ denotes the number of nonzero columns in c and $|r|_{\text{row}}$ denotes the number of nonzero rows in r.

It is immediate that product-expansion yields a bound on the distances of the associated codes:

Lemma 9 (Well known)**.** *If the pair $C_1, C_2 \subseteq \mathbb{F}_q^n$ is ρ-product expanding, then C_1 and C_2 have distance $\geq \rho n$.*

Proof. Let $x \in C_1 \otimes \mathbb{F}_q^n$ have its first column be a minimum-weight nonzero codeword of C_1, and have all other columns be 0. Then ρ-product-expansion implies that C_1 has distance $|x| \geq \rho n$. A similar argument holds for C_2. □

The following result bounding the product expansion of random pairs of codes was shown independently by [KP23] and [DHLV23], though only the former explicitly considered non-binary alphabets.

Proposition 10 ([KP23])**.** *Fix any finite field $\mathbb{F}_q$. For every fixed $\epsilon > 0$, there exists a constant $\rho = \rho(\epsilon) > 0$ and a function $\delta(n) = \delta(n; \epsilon) \to 0$ as $n \to \infty$ such that the following holds. For every pair of integers $k_1, k_2 \in (\epsilon n, (1 - \epsilon)n)$, if $C_i \subseteq \mathbb{F}_q^n$ for $i = 1, 2$ is drawn uniformly at random from the set of linear codes of dimension k_i, then with probability $\geq 1 - \delta(n)$ the pair (C_1, C_2) will be ρ-product-expanding.*

Applying Proposition 10 with a union bound over (C_1, C_2) and $(C_1^\perp, C_2^\perp)$ immediately yields the following corollary.

Corollary 11 ([KP23])**.** *Defining all variables as in Proposition 10, then with probability $\geq 1 - 2\delta(n)$ both (C_1, C_2) and $(C_1^\perp, C_2^\perp)$ will be ρ-product-expanding.*

The distance bound for quantum Tanner codes will also rely on the Cayley graphs $\mathrm{Cay}(G, A)$ and $\mathrm{Cay}(G, B)$ having sufficiently good expansion.

Definition 12. For a regular graph Γ of degree $\Delta(\Gamma)$, the **(unnormalized) spectral expansion** $\lambda(\Gamma)$ is the second largest absolute value of an eigenvalue of the adjacency matrix of Γ. If $\lambda(\Gamma) \leq 2\sqrt{\Delta(\Gamma) - 1}$, then Γ is **Ramanujan**. Meanwhile, if an infinite family of regular graphs Γ all satisfy $\lambda(\Gamma) \leq \Delta(\Gamma)^{1/2+o(1)}$, then the family is **almost Ramanujan**. Here $o(1)$ denotes any function of Δ that approaches 0 as $\Delta \to \infty$.

Constructions of Ramanujan Cayley graphs have for instance been given by [LPS88] and [Mor94]; the latter construction in particular is strongly explicit:

Theorem 13 ([Mor94])**.** *For every prime power $q \geq 3$, there exists a strongly explicit family of $(q+1)$-regular Ramanujan Cayley graphs $(\Gamma_m)_{m\in\mathbb{N}}$ with the number of vertices given by*

$$|V(\Gamma_m)| = \begin{cases} q^{2m}(q^{4m}-1), & q \equiv 0 \pmod 2 \\ q^{2m}(q^{4m}-1)/2, & q \equiv 1 \pmod 2. \end{cases}$$

The graphs in Theorem 13 can be used to instantiate strongly explicit linear-distance quantum Tanner codes by [LZ22, LZ23a]. However, they are not quite sufficient for our purposes. Specifically, using these graphs, our planted quantum Tanner codes and the resulting strongly explicit SoS lower bounds described in Section 5 would hold only for alphabets $\mathbb{F}_q$ of characteristic $p \geq 7$. This restricton on the field size arises because the graphs in Theorem 13 have $|V(\Gamma_m)|$ divisible by 2, 3, and 5, but our results require Cayley expanders Γ for which q is relatively prime with $|V(\Gamma)|$.

We therefore instead use the Cayley expanders given by Example 3.4 in [LW93], for which the number of vertices is guaranteed to be a power of any desired prime. [LW93] showed that these graphs have constant degree Δ and constant spectral expansion $< \Delta$. By amplifying the expansion to be near-Ramanujan by applying Theorem 1.2 in [JMRW22], we obtain the following result, which we formally prove in Section 5.4.

Theorem 14 (Follows from [LW93, JMRW22])**.** *For every prime p, there is an infinite set $\boldsymbol{\Delta} \subseteq \mathbb{N}$ for which there exists a strongly explicit family of almost-Ramanujan Cayley graphs $(\Gamma_{m,\Delta})_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}}$, where $\Gamma_{m,\Delta}$ has $|V(\Gamma_{m,\Delta})| = p^{3m}$ vertices and has degree Δ. Furthermore, we may choose $\boldsymbol{\Delta}$ such that for every $\Delta \in \boldsymbol{\Delta}$, either $\Delta + 1 \in \boldsymbol{\Delta}$ or $\Delta - 1 \in \boldsymbol{\Delta}$.*

We are now ready the present the distance bound for quantum Tanner codes.

Theorem 15 ([LZ22, LZ23a])**.** *For every fixed $\rho > 0$, the following holds for all sufficiently large Δ. Let $\mathcal{C}$ be a quantum Tanner code for which:*

1. *$\mathrm{Cay}(G, A), \mathrm{Cay}(G, B)$ are almost-Ramanujan graphs of degree Δ.*
2. *$(C_A, C_B), (C_A^\perp, C_B^\perp)$ are ρ-product-expanding.*

Then $\mathcal{C}$ has distance $d \geq cn$ for a constant $c > 0$ depending only on ρ, Δ.

Recall that by Lemma 9, Condition 2 in Theorem 15 implies that $C_A, C_B, C_A^\perp, C_B^\perp$ have distance $\geq \rho\Delta$.

Condition 1 in Theorem 15 can be met using the strongly explicit Ramanujan graphs in Theorem 13, or using the strongly explicit almost-Ramanujan graphs in Theorem 14. If $C_A, C_B \subseteq \mathbb{F}_q^\Delta$

are chosen to be random codes of some fixed rates $0 < R_A, R_B < 1$ for any sufficiently large constant Δ, then Condition 2 is met by Corollary 11. Because Δ is a constant as $n \to \infty$, we may find C_A, C_B in constant time by a brute force search, so the overall construction of $\mathcal{C}$ is strongly explicit.

While [LZ22, LZ23a] only proved Theorem 15 in the case where $\mathrm{Cay}(G, A), \mathrm{Cay}(G, A)$ are Ramanujan graphs of degree Δ, their proof generalizes flawlessly to allow for almost-Ramanujan graphs. Specifically, the proof of linear distance in Section 3 of [LZ23a] defines "exceptional vertices" using a parameter $\alpha = \delta^2/256$, where δ denotes the relative distance of the inner codes. If we for instance redefine $\alpha = \delta^2/\Delta^{1/100}$, we can essentially carry through their exact same analysis assuming $\mathrm{Cay}(G, A), \mathrm{Cay}(G, B)$ are almost-Ramanujan. An additional factor of $\Delta^{1/50+o(1)}$ arises in the bound in their Lemma 10, and an additional $\Delta^{o(1)}$ factor arises in the bound in their Lemma 11, but the proof of linear still goes through with these slightly worse parameters.

We suspect that almost-Ramanujan expansion $\lambda \leq \Delta^{1/2+o(1)}$ is still stronger than necessary, and it may in fact be sufficient to have spectral expansion λ equal to some small constant fraction of Δ. However, it is less transparent how to modify the proof of [LZ23a] for this case, so for conciseness we do not pursue this direction.

3.4 Small-Set (Co)boundary Expansion

For our applications of quantum Tanner codes, we will need a stronger notion than distance, called *small-set (co)boundary expansion*, which was first formally stated in the context of quantum codes by Hopkins and Lin [HL22]. Below, for a code C, we denote $|y|_C = \min_{y' \in C} |y + y'|$.

Definition 16 (Small-set (co)boundary expansion)**.** Let $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ be a CSS code given by parity check matrices $H_X \in \mathbb{F}_q^{m_X \times n}$ and $H_Z \in \mathbb{F}_q^{m_Z \times n}$. For $c_1, c_2 > 0$, we say that $\mathcal{C}$ has (c_1, c_2)**-small-set boundary expansion** if it holds for every $y \in \mathbb{F}_q^n$ with $|y| \leq c_1 n$ that

$$\frac{|H_Z y|}{m_Z} \geq c_2 \frac{|y|_{C_X^\perp}}{n}.$$

Similarly, $\mathcal{C}$ has (c_1, c_2)**-small-set coboundary expansion** if it holds for every $y \in \mathbb{F}_q^n$ with $|y| \leq c_1 n$ that

$$\frac{|H_X y|}{m_X} \geq c_2 \frac{|y|_{C_Z^\perp}}{n}.$$

Small-set (co)boundary expansion immediately implies a bound on the distance of the code:

Lemma 17 (Well known)**.** *If $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ of block length n has (c_1, c_2)-small set boundary and coboundary expansion for $c_1, c_2 > 0$, then $\mathcal{C}$ has distance $\geq c_1 n$.*

Proof. For every $y \in \mathbb{F}_q^n \setminus C_X^\perp$ with $|y| \leq c_1 n$, then $|y|_{C_X^\perp} > 0$, so small-set boundary expansion implies that $|H_Z y| \geq c_2 m_Z |y|_{C_X^\perp}/n > 0$ and thus $y \notin C_Z$. An analogous argument shows that $C_X \setminus C_Z^\perp$ has no elements of weight $\leq c_1 n$. □

It was originally observed that quantum Tanner codes have small set (co)boundary expansion in [HL22]. We remark that another proof is given implicitly by the decoder of Leverrier and Zémor [LZ23a]. Specifically, there exists a constant $c_1 > 0$ such that for any errors $e_X, e_Z \in \mathbb{F}_q^n$ of sufficiently low weight $|e_X|, |e_Z| \leq c_1 n$, the decoder of [LZ23a] takes as input the syndromes

$s_X = H_X e_X$ and $s_Z = H_Z e_Z$, and outputs some $e'_X \in e_X + C_Z^\perp$ and $e'_Z \in e_Z + C_X^\perp$ such that $|e'_X| \leq O(|s_X|)$, $|e'_Z| \leq O(|s_Z|)$. It follows that $|e_X|_{C_Z^\perp} \leq |e'_X| \leq O(|s_X|)$ and $|e_Z|_{C_X^\perp} \leq |e'_Z| \leq O(|s_Z|)$, which are precisely the conditions required by small-set coboundary and boundary expansion, respectively. The result below formally summarizes this small-set (co)boundary expansion.

Theorem 18 ([HL22, LZ23a]). *For every fixed $\rho > 0$, the following holds for all sufficiently large Δ. Let $\mathcal{C}$ be a quantum Tanner code that satisfies Conditions 1, 2 in the statement of Theorem 15. Then $\mathcal{C}$ has (c_1, c_2)-small-set boundary and coboundary expansion for constants $c_1, c_2 > 0$ depending only on the values of ρ, Δ.*

Note that Theorem 18 implies Theorem 15 by Lemma 17.

As described in Section 3.3, [LZ23a] assume the Cayley graphs $\mathrm{Cay}(G, A), \mathrm{Cay}(G, B)$ are Ramanujan, whereas we make the slightly weaker assumption that they are almost-Ramanujan. However, the decoding proof in [LZ23a] is similar to the distance proof, and the extension to almost-Ramanujanness is nearly identical. Specifically, while Section 5 of [LZ23a] defines a parameter $\alpha = \delta^2\epsilon^2/2^{10}$, we redefine this parameter to be $\delta^2\epsilon^2/\Delta^{1/100}$, and carry through the rest of the proof essentially as before. An additional factor of $\Delta^{1/50+o(1)}$ arises in the bound in their Lemma 15, and an additional factor of $\Delta^{o(1)}$ arises in the bound in their Lemma 16, but otherwise the decoding analysis goes through as before, with these slightly worse parameters.

Leverrier and Zémor [LZ23b] show how their decoder can also be used to decode the asymptotically good qLDPC codes of Panteleev and Kalachev [PK22]; again in this case the decoder outputs an error whose weight is linear in the syndrome weight. Therefore a similar result as Theorem 18 holds for the codes of [PK22] as well.

4 NLTS Hamiltonians from Codes of Arbitrary Dimension

In this section, we show that quantum LDPC codes with linear distance and an appropriate clustering property yield NLTS Hamiltonians, regardless of the code dimension. This result improves upon the prior construction of NLTS Hamiltonians of [ABN23], which required the stronger assumption that the code dimension be linearly large. For simplicity in this section, we restrict attention to binary alphabets, though we expect the results to generalize naturally to qudits for more general alphabet sizes.

4.1 Setup of the Local Hamiltonian

We let $\mathcal{C} = \mathrm{CSS}(C_X, C_Z) = \mathrm{span}\{|y + C_X^\perp\rangle : y \in C_Z\}$ be an $[[n, k, d]]_2$ quantum LDPC CSS code, where all parity checks have weight $\leq \ell$. We assume $\mathcal{C}$ belongs to a family of such codes with constant relative distance $d/n = \Omega(1)$ and constant locality $\ell = O(1)$ as the block length $n \to \infty$. We will also assume that $\mathcal{C}$ satisfies the clustering property described in Definition 19 below. The main novel aspect of our proof is that it holds for any nonzero code dimension $k > 0$. In contrast, the prior NLTS proof [ABN23] assumed the rate k/n is constant as $n \to \infty$.

Recent good qLDPC codes, such as the quantum Tanner codes of [LZ22], satisfy all of the conditions above, and have constant rate. Our planted quantum Tanner codes in Theorem 32 also satisfy these conditions, but in some parameter regimes have dimension 1 (i.e. inverse linear rate). Hence our planted quantum Tanner codes provide examples of qLDPC codes of subconstant rate where our NLTS result applies.

Denote the parity check matrices of C_X and C_Z by $H_X \in \mathbb{F}_2^{m_X \times n}$ and $H_Z \in \mathbb{F}_2^{m_Z \times n}$ respectively, so that $C_X = \ker H_X$ and $C_Z = \ker H_Z$. By assumption all rows of H_X, H_Z have $\leq \ell$ nonzero entries. Also define $G_X^\epsilon = \{y \in \mathbb{F}_2^n : |H_X y| \leq \epsilon m_Z\}$, and define G_Z^ϵ analogously. We assume $\mathcal{C}$ satisfies the following clustering property for G_X^ϵ and G_Z^ϵ, which is stated as Property 1 in [ABN23]. Below, we denote $|y|_C = \min_{y' \in C} |y + y'|$.

Definition 19 (Clustering property [ABN23])**.** For constants $c_1, c_2, \epsilon_0 > 0$, we say that $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ exhibits (c_1, c_2, ϵ_0)**-clustering** if for all $0 < \epsilon < \epsilon_0$, the following hold:

1. Every $y \in G_X^\epsilon$ satisfies either $|y|_{C_Z^\perp} \leq c_1 \epsilon n$ or $|y|_{C_Z^\perp} \geq c_2 n$.
2. Every $y \in G_Z^\epsilon$ satisfies either $|y|_{C_X^\perp} \leq c_1 \epsilon n$ or $|y|_{C_X^\perp} \geq c_2 n$.

This clustering property follows from small-set (co)boundary expansion (Definition 16), as is shown below.

Lemma 20. *If $\mathcal{C}$ has (c_1', c_2')-small-set boundary and coboundary expansion, then $\mathcal{C}$ has (c_1, c_2, ϵ_0)-clustering for $c_1 = 1/c_2'$, $c_2 = c_1'$, $\epsilon_0 = 1$.*

Proof. Assume that $y \in G_Z^\epsilon$ satisfies $|y|_{C_X^\perp} \leq c_2 n = c_1' n$. Let y' be the minimum-weight element of $y + C_X^\perp$, so that $H_Z y' = H_Z y$ and $|y'| = |y|_{C_X^\perp}$. Then the small-set boundary expansion implies that $|H_Z y'|/m_Z \geq c_2' \cdot |y'|/n$, so $|y|_{C_X^\perp} = |y'| \leq (n/c_2' m_Z)|H_Z y'| \leq \epsilon n / c_2' = c_1 \epsilon n$. Thus we have shown the desired clustering for G_Z^ϵ; an analogous argument applies to G_X^ϵ. □

For the remainder of Section 4, we assume that $\mathcal{C}$ satisfies (c_1, c_2, ϵ_0)-clustering for some constants $c_1, c_2, \epsilon_0 > 0$ as $n \to \infty$.

Following [ABN23], we define our ℓ-local Hamiltonian $\mathbf{H}$ to be the code Hamiltonian

$$\mathbf{H} = \frac{1}{2}(\mathbf{H}_X + \mathbf{H}_Z)$$

for

$$\mathbf{H}_X = \frac{1}{m_X} \sum_{y \in \mathrm{rows}(H_X)} \frac{I - X^y}{2}$$
$$\mathbf{H}_Z = \frac{1}{m_Z} \sum_{y \in \mathrm{rows}(H_Z)} \frac{I - Z^y}{2}.$$

Thus in particular the ground space of $\mathbf{H}$ is precisely the code space $\mathcal{C} = \mathrm{span}\{|y + C_X^\perp\rangle : y \in C_Z\}$

While our general proof will follow that of [ABN23], our use of an uncertainty principle instead follows the earlier work of [EH17]. Below we state the uncertainty principle we will use, which appears implicitly in [EH17].

Lemma 21 (Uncertainty principle [HT03, EH17])**.** *Let A, B be Hermitian observables with $AB + BA = 0$ and $A^2 = B^2 = I$. Then for every (possibly mixed) state ρ, at least one of the inequalities $|\operatorname{Tr}(A\rho)| \leq 1/2 + 1/2\sqrt{2}$ or $|\operatorname{Tr}(B\rho)| \leq 1/2 + 1/2\sqrt{2}$ holds.*

For completeness, we present a proof of Lemma 21 from the following result for pure states, which is given as Lemma 37 in [EH17], but previously shown by [HT03].

Lemma 22 ([HT03])**.** *Let A, B be Hermitian observables with $AB + BA = 0$ and $A^2 = B^2 = I$, and let $|\psi\rangle$ be a pure state. Letting $\Delta A^2 = \langle\psi| A^2 |\psi\rangle - \langle\psi| A |\psi\rangle^2$, then $\Delta A^2 + \Delta B^2 \geq 1$.*

Proof of Lemma 21. Let $\rho = \sum_\psi p_\psi |\psi\rangle$ be a decomposition of ρ into an ensemble of pure states $|\psi\rangle$. By Lemma 22, it holds for each ψ that $\langle\psi| A |\psi\rangle^2 + \langle\psi| B |\psi\rangle^2 \leq 1$, so either $|\langle\psi| A |\psi\rangle| \leq 1/\sqrt{2}$ or $|\langle\psi| B |\psi\rangle| < 1/\sqrt{2}$. Partition the pure states into sets $\mathcal{A}, \mathcal{B}$ so that the former inequality holds for $\psi \in \mathcal{A}$ and the latter for $\psi \in \mathcal{B}$. Then either $p_\mathcal{A} := \sum_{\psi\in\mathcal{A}} p_\psi \geq 1/2$ or $1 - p_\mathcal{A} = p_\mathcal{B} := \sum_{\psi\in\mathcal{B}} p_\psi \geq 1/2$. Assume that $p_\mathcal{A} \geq 1/2$; the proof for $p_\mathcal{B} \geq 1/2$ is identical. Then the desired result follows by applying the triangle inequality:

$$|\operatorname{Tr}(A\rho)| = \left| \sum_{\psi\in\mathcal{A}} p_\psi \langle\psi| A |\psi\rangle + \sum_{\psi\in\mathcal{B}} p_\psi \langle\psi| A |\psi\rangle \right| \leq p_\mathcal{A} \cdot \frac{1}{\sqrt{2}} + p_\mathcal{B} \cdot 1 \leq \frac{1}{2\sqrt{2}} + \frac{1}{2}.$$

□

We follow prior works such as [EH17, ABN23] in estabilishing circuit lower bounds for approximate ground states of $\mathbf{H}$ by showing that the measurement distributions of these states are well-spread, in the following sense.

Definition 23. For $\mu, \delta > 0$, A probability distribution D over $\mathbb{F}_2^n$ is (μ, δ)**-spread** if there exist $S_0, S_1 \subseteq \mathbb{F}_2^n$ such that $D(S_0) \geq \mu$, $D(S_1) \geq \mu$, and $\operatorname{dis}(S_0, S_1) \geq \delta n$.

We specifically use following result, which appears as Fact 4 in [ABN23] but is similar to an earlier result of [EH17]. Below, for an n-qubit state ψ, we let D_X^ψ and D_Z^ψ denote the distributions over $\mathbb{F}_2^n$ obtained by measuring ψ in the X and Z bases respectively.

Lemma 24 (Circuit lower bound [EH17, ABN23])**.** *Let ψ be a (possibly mixed) quantum state on n qubits such that the Z-measurement distribution D_Z^ψ is (μ, δ)-spread. Then any circuit (on $\geq n$ qubits) that constructs ψ must have depth at least*

$$\frac{1}{3} \log \left(\frac{\delta^2 n}{400 \log(1/\mu)} \right).$$

4.2 Statement of Main Result on NLTS Hamiltonians

In this section, we state our main technical result, which implies that the code Hamiltonian $\mathbf{H}$ for a CSS code with linear distance that exhibits the clustering property is NLTS, that is, its approximate ground states cannot be constructed by constant-depth circuits. Crucially, we only assume that the dimension of the code is positive.

Specifically, our main technical result below shows that the measurement distribution of every approximate ground state of $\mathbf{H}$ is well-spread in either the X or Z basis.

Theorem 25. *Let $\mathbf{H}$ be the code Hamiltonian for a $[[n, k, d]]_2$ CSS code $\mathcal{C} = \operatorname{CSS}(C_X, C_Z)$ of positive dimension $k > 0$ that exhibits (c_1, c_2, ϵ_0)-clustering. For any*

$$\epsilon < \frac{1}{1000} \cdot \min \left\{ \frac{\epsilon_0}{2}, \frac{c_2}{4c_1}, \frac{d}{2c_1 n} \right\}, \tag{1}$$

let ρ be an ϵ-approximate ground state of $\mathbf{H}$, so that $\operatorname{Tr}(\mathbf{H}\rho) \leq \epsilon$. Then at least one of D_X^ρ or D_Z^ρ is (μ, δ)-spread for $\mu = .02$ and $\delta = c_2$.

Lemma 24 immediately yields the following corollary.

Corollary 26 (**H** is NLTS)**.** *Define ϵ, μ, δ, and* **H** *as in Theorem 25. Then no ϵ-approximate ground state of* **H** *can be constructed by a circuit of depth less than*

$$\frac{1}{3}\log\left(\frac{\delta^2 n}{400\log(1/\mu)}\right) + 1.$$

Our proof of Theorem 25 is similar to [EH17] in that we combine an uncertainty principle with a decoding procedure to obtain uncertainty for approximate ground states. We furthermore use the clustering property of $\mathcal{C}$ similarly to [ABN23]. As such, the key novel aspect of our proof is the use of a "decoding" procedure that handles clusters of approximate ground states which do not correspond to any true codeword.

4.3 Proof of Well-Spreadness for Approximate Ground States

In this section, we prove Theorem 25. Throughout this section, we maintain the notation in the statement of Theorem 25, so that $\mathcal{C} = \mathrm{CSS}(C_X, C_Z)$ is a $[[n,k,d]]_2$ CSS code exhibiting (c_1, c_2, ϵ_0)-clustering, ρ is an ϵ-approximate ground state of **H** for ϵ as in (1), and $\mu = .02$, $\delta = c_2$.

4.3.1 Reducing to Well-Spreadness of Pure States with Small Syndrome

We will first show that D_X^ρ and D_Z^ρ are mostly supported inside $G_X^{O(\epsilon)}$ and $G_Z^{O(\epsilon)}$ respectively, so that up to a small loss in parameters we may assume they are entirely supported inside these sets. We will also show that it suffices to consider pure states $\psi' = |\psi'\rangle\langle\psi'|$, rather than arbitrary mixed states ρ.

Formally, we may decompose our Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ into orthogonal subspaces as

$$\mathcal{H} = \bigoplus_{e_X + C_X \in \mathbb{F}_2^n/C_X, e_Z + C_Z \in \mathbb{F}_2^n/C_Z} X^{e_Z} Z^{e_X} \mathcal{C},$$

where the choices of coset representatives in the above sum does not matter because by definition $X^{c_Z} Z^{c_X} \mathcal{C} = \mathcal{C}$ for $c_X \in C_X$, $c_Z \in C_Z$. Observe furthermore that each subspace $X^{e_Z} Z^{e_X} \mathcal{C}$ is by definition an eigenspace of the code Hamiltonian **H** with eigenvalue $|H_X e_X|/2m_X + |H_Z e_Z|/2m_Z$.

Set

$$\epsilon' = 1000\epsilon,$$

and let

$$\mathcal{C}^{\leq \epsilon'} = \bigoplus_{e_X + C_X : |H_X e_X| \leq \epsilon' m_X, e_Z + C_Z : |H_Z e_Z| \leq \epsilon' m_Z} X^{e_Z} Z^{e_X} \mathcal{C}.$$

Therefore $\mathcal{C}^{\leq \epsilon'}$ is the span of some of the eigenspaces of energy $\leq \epsilon'$, and contains all of the eigenspaces of energy $\leq \epsilon'/2$. Let $\Pi_{\mathcal{C}^{\leq \epsilon'}}$ denote projection onto this subspace. Note that by definition, every $|\psi'\rangle \in \mathcal{C}^{\leq \epsilon'}$ has $\mathrm{supp}(D_X^{\psi'}) \subseteq G_X^{\epsilon'}$ and $\mathrm{supp}(D_Z^{\psi'}) \subseteq G_Z^{\epsilon'}$.

We now reduce the task of proving Theorem 25 to the following proposition. Below, recall that we carry the definitions of $\mathbf{H}, \rho, \epsilon, \mu, \delta$ from Theorem 25.

Proposition 27. *There exist sets* $S_X^0, S_X^1, S_Z^0, S_Z^1 \subseteq \mathbb{F}_2^n$ *such that* $\mathrm{dis}(S_X^0, S_X^1), \mathrm{dis}(S_Z^0, S_Z^1) \geq \delta n$, *and such that for every pure state* $|\psi'\rangle \in \mathcal{C}^{\leq \epsilon'}$, *either*

$$D_X^{\psi'}(S_X^0), D_X^{\psi'}(S_X^1) \geq \mu' \quad \text{or} \quad D_Z^{\psi'}(S_Z^0), D_Z^{\psi'}(S_Z^1) \geq \mu', \tag{2}$$

where $\mu' = 1/4 - 1/4\sqrt{2}$.

Below, we first prove Theorem 25 assuming Proposition 27; this proof uses relatively standard techniques, though it is slightly tedious. We will subsequently prove the proposition, which contains the key ideas for our result.

Proof of Theorem 25. Fix any mixed state ρ with $\mathrm{Tr}(\mathbf{H}\rho) \leq \epsilon$. Our goal is to use Proposition 27 to show that either

$$D_X^{\rho}(S_X^0), D_X^{\rho}(S_X^1) \geq \mu \quad \text{or} \quad D_Z^{\rho}(S_Z^0), D_Z^{\rho}(S_Z^1) \geq \mu.$$

For this purpose, we first decompose $\rho = \sum_\psi p_\psi |\psi\rangle\langle\psi|$ into a classical ensemble of pure states ψ. Then

$$\begin{aligned}
\epsilon &\geq \mathrm{Tr}(\mathbf{H}\rho) \\
&= \sum_\psi p_\psi \langle\psi| \mathbf{H} |\psi\rangle \\
&\geq \sum_\psi p_\psi \langle\psi| \frac{\epsilon'}{2}(I - \Pi_{\mathcal{C}^{\leq \epsilon'}}) |\psi\rangle ,
\end{aligned}$$

where we have used the fact that $\mathcal{C}^{\leq \epsilon'}$ contains all eigenspaces of $\mathbf{H}$ of eigenvalue $\leq \epsilon'/2$. Therefore

$$\sum_\psi p_\psi \langle\psi| \Pi_{\mathcal{C}^{\leq \epsilon'}} |\psi\rangle = 1 - \sum_\psi p_\psi \langle\psi| (I - \Pi_{\mathcal{C}^{\leq \epsilon'}}) |\psi\rangle \geq 1 - \frac{2\epsilon}{\epsilon'}. \tag{3}$$

Proposition 27 implies that (2) holds for every $|\psi'\rangle = \Pi_{\mathcal{C}^{\leq \epsilon'}} |\psi\rangle / \|\Pi_{\mathcal{C}^{\leq \epsilon'}} |\psi\rangle\|$. Therefore if we let Ψ_X denote the set of ψ for which the first inequality in (2) holds for $|\psi'\rangle$ and Ψ_Z the set of ψ for which the second inequality in (2) holds for $|\psi'\rangle$, then either

$$\sum_{\psi \in \Psi_X} p_\psi \geq \frac{1}{2} \quad \text{or} \quad \sum_{\psi \in \Psi_Z} p_\psi \geq \frac{1}{2}.$$

Assume the latter of the two inequalities above holds; the proof for the former is analogous. For

$b = 0, 1$, let Π_Z^b denote orthogonal projection onto $\text{span}\{|y\rangle : y \in S_Z^b\}$. Then by definition

$$\begin{aligned} D_Z^\rho(S_Z^b) &= \text{Tr}\Big(\Pi_Z^b \rho\Big) \\ &= \sum_\psi p_\psi \|\Pi_Z^b |\psi\rangle\|^2 \\ &\geq \sum_\psi p_\psi (\|\Pi_Z^b |\psi'\rangle\| - \|\Pi_Z^b (|\psi\rangle - |\psi'\rangle)\|)^2 \\ &\geq \sum_\psi p_\psi (\|\Pi_Z^b |\psi'\rangle\| - \||\psi\rangle - |\psi'\rangle\|)^2 \\ &\geq \sum_\psi p_\psi \|\Pi_Z^b |\psi'\rangle\|^2 - 2\sum_\psi p_\psi \||\psi\rangle - |\psi'\rangle\|^2 \end{aligned}$$

Now by defininition we can bound the first term on the RHS above by restricting to the sum over $\psi \in \Psi_Z$ and applying (2) with $D_Z^{\psi'}(S_Z^b) = \|\Pi_Z^b |\psi'\rangle\|^2$ to obtain

$$\sum_\psi p_\psi \|\Pi_Z^b |\psi'\rangle\|^2 \geq \sum_{\psi \in \Psi_Z} p_\psi \|\Pi_Z^b |\psi'\rangle\|^2 \geq \frac{1}{2} \cdot \mu'.$$

Meanwhile, we can bound the second term by expanding $\||\psi\rangle - |\psi'\rangle\|^2$ as in inner product to obtain

$$\begin{aligned} 2\sum_\psi p_\psi \||\psi\rangle - |\psi'\rangle\|^2 &= 4\sum_\psi p_\psi (1 - \langle\psi'|\psi\rangle) \\ &= 4\sum_\psi p_\psi \left(1 - \sqrt{\langle\psi| \Pi_{\mathcal{C}^{\leq\epsilon'}} |\psi\rangle}\right) \\ &\leq 4\sum_\psi p_\psi \left(\frac{1 - \langle\psi| \Pi_{\mathcal{C}^{\leq\epsilon'}} |\psi\rangle}{2}\right) \\ &\leq 2 \cdot \frac{2\epsilon}{\epsilon'} \\ &= \frac{4\epsilon}{\epsilon'}, \end{aligned}$$

where the first two equalities above apply the definition of $|\psi'\rangle = \Pi_{\mathcal{C}^{\leq\epsilon'}} |\psi\rangle / \|\Pi_{\mathcal{C}^{\leq\epsilon'}} |\psi\rangle\|$, and the second inequality holds by (3). Thus we have shown that

$$D_Z^\rho(S_Z^b) \geq \frac{\mu'}{2} - \frac{4\epsilon}{\epsilon'} = \frac{1}{8} - \frac{1}{8\sqrt{2}} - \frac{4}{1000} > \mu,$$

as desired. □

4.3.2 Decoding Clusters of Small-Syndrome States

To prove Proposition 27, we begin by using the clustering property of $\mathcal{C}$ to partition $G_X^{\epsilon'}$ and $G_Z^{\epsilon'}$ into clusters, for which we will subsequently choose representative elements that we will use to

define a decoding map for states $|\psi'\rangle$ with small syndrome. Below, we first analyze the clustering of $G_Z^{\epsilon'}$; the case of $G_X^{\epsilon'}$ will be exactly analogous.

We consider clusters defined similarly as in [ABN23]. However, to obtain our improvement over [ABN23], we will leverage an additional linear structure in the set of clusters (Property 2 in Lemma 2 below), which ultimately allows us to use the uncertainty principle in Lemma 21.

Given $y \in G_Z^{\epsilon'}$, define a cluster $Y_Z^y \subseteq G_Z^{\epsilon'}$ by

$$Y_Z^y = \{y' \in G_Z^{\epsilon'} : |y + y'|_{C_X^\perp} \leq 2c_1\epsilon' n\}.$$

The following lemma follows directly from our definitions.

Lemma 28. *The clusters Y_Z^y for $y \in G_Z^{\epsilon'}$ form a partition of $G_Z^{\epsilon'}$ satisfying the following properties:*

1. *Every pair of distinct clusters $Y_Z^y \neq Y_Z^{y'}$ satisfies $\mathrm{dis}(Y_Z^y, Y_Z^{y'}) \geq c_2 n$.*
2. *For $c \in C_Z$, then $Y_Z^{y+c} = Y_Z^y + c$, and in particular $Y_Z^{y+c} = Y_Z^y$ if and only if $c \in C_X^\perp$.*

Proof. We first show that the cluters form a parition of $G_Z^{\epsilon'}$. Fix some $y \in G_Z^{\epsilon'}$. Then for every $y' \in Y_Z^y$, it follows that every $y'' \in Y_Z^{y'}$ has $|y'' + y|_{C_X^\perp} \leq |y'' + y'|_{C_X^\perp} + |y' + y|_{C_X^\perp} \leq 4c_1\epsilon' n$. But by assumption (see the statement of Theorem 25) $2\epsilon' < \epsilon_0$ and $4c_1\epsilon' < c_2$, so because $y'' + y \in G_Z^{2\epsilon'}$, the clustering property implies that $|y'' + y|_{C_X^\perp} \leq 2c_1\epsilon' n$, so that $y'' \in Y_Z^y$. Thus we have shown that every $y' \in Y_Z^y$ has $Y_Z^{y'} \subseteq Y_Z^y$, and by the same reasoning $Y_Z^y \subseteq Y_Z^{y'}$, so $Y_Z^{y'} = Y_Z^y$. Thus every pair of clusters is either equal or disjoint, so the clusters Y_Z^y form a partition of $G_Z^{\epsilon'}$.

Now every pair of distinct clusters $Y_Z^y \neq Y_Z^{y'}$ satisfies $\mathrm{dis}(Y_Z^y, Y_Z^{y'}) \geq c_2 n$, as if this distance was $< c_2 n$, the clustering property would imply that it is $\leq 2c_1\epsilon' n$, which then implies that $Y_Z^y = Y_Z^{y'}$.

It remains to show Property 2 in the lemma statement. For every $y \in G_Z^{\epsilon'}$ and $c \in C_Z$, by definition $H_Z(y+c) = H_Z y$ and thus Y_Z^{y+c} is also a cluster in $G_Z^{\epsilon'}$, which is isomorphic to Y_Z^y under the isomorphism $y' \mapsto y' + c$; that is, $Y_Z^{y+c} = Y_Z^y + c$. If $c \in C_X^\perp$, then $|y + (y + c)|_{C_X^\perp} = 0$ so that $y + c \in Y_Z^y$ and therefore $Y_Z^{y+c} = Y_Z^y$. Meanwhile, if $c \in C_Z \setminus C_X^\perp$, then $Y_Z^{y+c} \neq Y_Z^y$, as otherwise it would follow that $|y + (y + c)|_{C_X^\perp} = |c|_{C_X^\perp} \leq 2c_1\epsilon' n$. But by assumption (see Theorem 25) $\mathcal{C}$ has distance $d > 2c_1\epsilon' n$, so $|c|_{C_X^\perp} > 2c_1\epsilon' n$. □

Lemma 28 implies that Y_Z^y has distinct translates by all $c + C_X^\perp \in C_Z/C_X^\perp$, where all representatives of a given coset of $C_X^\perp$ yield the same translate. We denote the collection of these translates for a given cluster Y_Z^y by

$$\mathcal{Y}_Z^{Y_Z^y} = \{Y_Z^{y+c} : c \in C_Z\}.$$

For each such collection $\mathcal{Y}_Z = \mathcal{Y}_Z^{Y_Z^y}$ of clusters, we fix an arbitrary representative $Y_Z(\mathcal{Y}_Z) \in \mathcal{Y}_Z$.

Now for a given syndrome $s = H_Z y \in \mathbb{F}_2^{m_Z}$ of some $y \in G_Z^{\epsilon'}$, so that $|s| \leq \epsilon' m_Z$, then the set of bit strings with syndrome s is precisely the coset $y + C_Z$. By Lemma 28, $(y + C_Z) \cap Y_Z(\mathcal{Y}_Z^{Y_Z^y})$ is a coset of $C_X^\perp$, and is in particular therefore nonempty. Thus we may associate to s an arbitrary representative $e_Z(s) \in (y + C_Z) \cap Y_Z(\mathcal{Y}_Z^{Y_Z^y})$.

We now let Dec_Z be a unitary acting on $n + m_Z$ qubits with the following "decoding" property: for every $y \in G_Z^{\epsilon'}$, it holds that

$$\mathrm{Dec}_Z |y\rangle \otimes |0\rangle = |y + e_Z(H_Z y)\rangle \otimes |H_Z y\rangle .$$

We let Dec_Z^1 be the channel acting on n qubits that simply applies Dec_Z and traces out the syndrome register. Formally, for every $y \in G_Z^{\epsilon'}$, then

$$\mathrm{Dec}_Z^1(|y\rangle\langle y|) = |y + e_Z(H_Z y)\rangle\langle y + e_Z(H_Z y)|\,.$$

Equivalently, Dec_Z^1 is the channel that performs a Z-syndrome measurement on its input $|y\rangle$, and then adds $e_Z(s)$ to the post-measurement state, where $s = H_Z y$ was the measurement outcome.

The channel Dec_Z^1 performs a weak form of decoding in the following sense. Recall that an ordinary decoder in the Z basis for a CSS code maps all bit strings near a given codeword $c \in C_Z$ to some element of the coset $c + C_X^\perp$. In contrast, as shown below, the key property of our decoding channel Dec_Z^1 is that it sends all bit strings in a given cluster Y_Z^y to elements of the same coset $c + C_X^\perp \in C_Z/C_X^\perp$, though this coset may be far away from the cluster Y_Z^y.

Lemma 29. *For every cluster Y_Z^y and every pair of elements $y, y' \in Y_Z^y$, then $y' + e_Z(H_Z y') \in y + e_Z(H_Z y) + C_X^\perp$.*

Proof. By definition $e_Z(H_Z y) \in y + C_Z$ and $e_Z(H_Z y') \in y' + C_Z$ both lie in the cluster $Y_Z(\mathcal{Y}_Z^{Y_Z^y})$. Therefore by Lemma 28, both $e_Z(H_Z y')$ and $y' + (y + e_Z(H_Z y))$ belong to both $y' + C_Z$ and $Y_Z(\mathcal{Y}_Z^{Y_Z^y})$, and thus $e_Z(H_Z y') \in y' + (y + e_Z(H_Z y)) + C_X^\perp$, or equivalently, $y' + e_Z(H_Z y') \in y + e_Z(H_Z y) + C_X^\perp$. □

To conclude this section, we extend all of the clustering terminology and results above for $G_Z^{\epsilon'}$ to their analogues for $G_X^{\epsilon'}$. Specifically, we similarly obtain a partition of $G_X^{\epsilon'}$ into clusters Y_X^y for $y \in G_X^{\epsilon'}$. We again conclude that each cluster Y_X^y in $G_X^{\epsilon'}$ has a set $\mathcal{Y}_X^{Y_X^y}$ of distinct translates by all $c + C_Z^\perp \in C_X/C_Z^\perp$. We fix arbitrary representative clusters $Y_X(\mathcal{Y}_X) \in \mathcal{Y}_X$, and assign to each syndrome $s = H_X y$ for $y \in G_X^{\epsilon'}$ an element $e_X(s) \in (y + C_X) \cap Y_X(\mathcal{Y}_X^{Y_X^y})$. We then obtain an X decoding unitary Dec_X and channel Dec_X^1, which are defined analogously to their Z analogues, except the syndrome measurement and error correction steps are performed in the X basis instead of the Z basis. Observe that Dec_X^1 and Dec_Z^1 commute, so we can define $\mathrm{Dec}^1 = \mathrm{Dec}_X^1\,\mathrm{Dec}_Z^1 = \mathrm{Dec}_Z^1\,\mathrm{Dec}_X^1$.

4.3.3 Applying Decoding to Prove Well-Spreadness

We now complete the proof of Proposition 27, which as shown above in turn implies Theorem 25, by applying the uncertainty principle in Lemma 21 to the decodings of small-syndrome states for $\mathbf{H}$.

Proof of Proposition 27. Because $\mathcal{C}$ has dimension $k > 0$, the space $C_Z/C_X^\perp = (C_X/C_Z^\perp)^\perp$ is nonzero, so there exist $\bar{c}_X \in C_X \setminus C_Z^\perp$, $\bar{c}_Z \in C_Z \setminus C_X^\perp$ such that $\bar{c}_X \cdot \bar{c}_Z = 1$. Fix an arbitrary pair of such elements $\bar{c}_X, \bar{c}_Z$, so that $\bar{X} := X^{\bar{c}_Z}$ and $\bar{Z} := Z^{\bar{c}_X}$ are anticommuting logical operators for the code $\mathcal{C}$.

For $b = 0, 1$, define

$$S_X^b = \{y \in G_X^{\epsilon'} : \bar{c}_Z \cdot (y + e_X(H_X y)) = b\}$$
$$S_Z^b = \{y \in G_Z^{\epsilon'} : \bar{c}_X \cdot (y + e_Z(H_Z y)) = b\}.$$

Then by Lemma 29, for a given cluster Y_Z^y in $G_Z^{\epsilon'}$, all $y' \in Y_Z^y$ have $y' + e_Z(H_Z y')$ lying in the same coset $y + e_Z(H_Z y) + C_X^\perp$, and thus all $y' \in Y_Z^y$ have the same value of $\bar{c}_X \cdot (y' + e_Z(H_Z y')) = \bar{c}_X \cdot (y +$

$e_Z(H_Z y)$). Therefore all $y' \in Y_Z^y$ lie in the same set S_Z^b, where $b = \bar{c}_X \cdot (y + e_Z(H_Z y))$. It follows from Lemma 28 that $\mathrm{dis}(S_Z^0, S_Z^1) \geq c_2 n = \delta n$. Analogous reasoning implies that $\mathrm{dis}(S_X^0, S_X^1) \geq c_2 n = \delta n$.

It remains to be shown that (2) holds for every $|\psi'\rangle \in \mathcal{C}^{\epsilon'}$. By Lemma 21, either $|\operatorname{Tr}(\bar{X} \operatorname{Dec}^1(\psi'))| \leq 1/2 + 1/2\sqrt{2}$ or $|\operatorname{Tr}(\bar{Z} \operatorname{Dec}^1(\psi'))| \leq 1/2 + 1/2\sqrt{2}$. Assume the latter; the proof for the former is analogous. Now because $\bar{Z}$ by definition commutes with Dec_X, the distribution from measuring $\bar{Z}$ on $\operatorname{Dec}^1(\psi') = \operatorname{Dec}_X^1 \operatorname{Dec}_Z^1(\psi')$, or equivalently on $\operatorname{Dec}_X(\operatorname{Dec}_Z^1(\psi' \otimes |0\rangle\langle 0|)) \operatorname{Dec}_X^\dagger$, is the same as the distribution from measuring $\bar{Z}$ on $\operatorname{Dec}_Z^1(\psi')$. Therefore $|\operatorname{Tr}(\bar{Z} \operatorname{Dec}_Z^1(\psi'))| \leq 1/2 + 1/2\sqrt{2}$. But by definition if we expand $|\psi'\rangle = \sum_{y \in G_Z^{\epsilon'}} \psi'_y |y\rangle$, then it follows that

$$\begin{aligned}
\frac{1}{2} + \frac{1}{2\sqrt{2}} &\geq \operatorname{Tr}\left(\bar{Z} \operatorname{Dec}_Z^1(\psi')\right) \\
&= \operatorname{Tr}\left((\bar{Z} \otimes I) \operatorname{Dec}_Z(|\psi'\rangle\langle\psi'| \otimes |0\rangle\langle 0|) \operatorname{Dec}_Z^\dagger\right) \\
&= \left(\langle\psi'| \otimes \langle 0| \operatorname{Dec}_Z^\dagger\right) \left((\bar{Z} \otimes I) \operatorname{Dec}_Z |\psi'\rangle \otimes |0\rangle\right) \\
&= \left(\sum_{y \in G_Z^{\epsilon'}} \langle y + e_Z(H_Z y)| \otimes \langle H_Z y| (\psi'_y)^\dagger\right) \\
&\quad \cdot \left(\sum_{y \in G_Z^{\epsilon'}} \psi'_y (-1)^{\bar{c}_X \cdot (y + e_Z(H_Z y))} |y + e_Z(H_Z y)\rangle \otimes |H_Z y\rangle\right) \\
&= \sum_{y \in G_Z^{\epsilon'}} (-1)^{\bar{c}_X \cdot (y + e_Z(H_Z y))} |\psi'_y|^2 \\
&= \left|\sum_{y \in S_Z^0} |\psi'_y|^2 - \sum_{y \in S_Z^1} |\psi'_y|^2\right| \\
&= |D_Z^{\psi'}(S_Z^0) - D_Z^{\psi'}(S_Z^1)|.
\end{aligned}$$

Then because $D_Z^{\psi'}$ is supported inside $G_Z^{\epsilon'} = S_Z^0 \sqcup S_Z^1$ by the definition of ψ', it follows that $D_Z^{\psi'}(S_Z^0) + D_Z^{\psi'}(S_Z^1) = 1$, so we must have

$$D_Z^{\psi'}(S_Z^0), D_Z^{\psi'}(S_Z^1) \geq \frac{1}{4} - \frac{1}{4\sqrt{2}} = \mu',$$

as desired. □

5 Planting Codewords in QLDPC Codes

In this section, we show how to plant a nontrivial codeword in the quantum Tanner codes of [LZ22], thereby ensuring the code has positive dimension regardless of other parameters in the instantiation. For instance, when the inner codes C_A, C_B are chosen to be of rate $1/2$ in the quantum Tanner code construction, the only prior method for bounding dimension, namely by counting parity checks, fails to ensure the dimension of the global code is positive (see Section 3). However, our planted construction of quantum Tanner codes has positive dimension regardless of the rates of the inner

codes, and thus provides a new way to ensure positive dimension, that works in previously unfeasible parameter regimes. We remark that a similar technique also works for the codes of [PK22], though we do not present the details to avoid redundancy.

Using the strongly explicit nature of the planted codeword, we apply our construction to improve upon the explicit SoS lower bounds of [HL22] to obtain *strongly* explicit SoS lower bounds.

5.1 Intuition: Planted Classical Tanner Codes

In this section, we present the simpler case of how to plant a codeword in a classical Tanner code, which motivates our construction in the quantum case.

Recall that a classical Tanner code $C = \mathrm{Tan}(\Gamma, C_{\mathrm{in}}) \subseteq \mathbb{F}_q^n$ is constructed from a Δ-regular graph Γ with n edges and an inner code $C_{\mathrm{in}} \subseteq \mathbb{F}_q^\Delta$ as follows. We associate the set of all edges in Γ with the set $[n]$, and we associate the set of edges incident to each vertex $v \in \Gamma$ with the set $[\Delta]$. Then we define C to be the set of all edge labelings $y \in \mathbb{F}_q^n = \mathbb{F}_q^{E(\Gamma)}$ such that the labels of edges incident to each $v \in \Gamma$ form a codeword in C_{in}.

The standard method for ensuring that the rate R of C is positive (and in fact linear in n) is to require that C_{in} be a linear code of rate $R_{\mathrm{in}} > 1/2$, so that by counting linear constraints it follows that $R \geq 1 - 2(1 - R_{\mathrm{in}})$.

However, if we only care about ensuring that $R > 0$, we may instead simply require that C_{in} contains the all-1s vector $\mathbf{1} \in \mathbb{F}_q^\Delta$, as then by definition the global all-1s vector $\mathbf{1} \in \mathbb{F}_q^n$ must lie in C. If the resulting "planted" classical code has no other nontrivial codewords, it is simply a repetition code, which is typically uninteresting classically.

However, below we construct a quantum analogue of such planted codes, which are more difficult to construct than their classical counterparts, and yield interesting complexity theoretic applications regardless of their rate. For instance, because the planted codeword is trivial to describe and therefore strongly explicit, we improve the explicit SoS lower bounds of [HL22] to be strongly explicit. Furthermore, in Corollary 26 of Section 4, we showed that such qLDPC codes of arbitrarily small rate also yield NLTS Hamiltonians.

5.2 Construction of Planted Quantum Tanner Codes

In this section, we present our construction of planted quantum Tanner codes. This construction can be viewed as a quantum analogue of the planted classical Tanner codes described in Section 5.1. The quantum case requires significantly more care, as desribed below.

The following proposition presents our paradigm for planting a nontrivial codeword in a quantum Tanner code

Proposition 30. *Let $\mathcal{C}$ be a quantum Tanner code as defined in Section 3.1 such that the following hold:*

1. *The all-1s vector $\mathbf{1} \in \mathbb{F}_q^\Delta$ lies in C_A and in $C_B^\perp$.*
2. *$n = |Q| = |G||A||B|$ is relatively prime with q.*

Then the all-1s vector $\mathbf{1} \in \mathbb{F}_q^Q$ lies in $C_Z \setminus C_X^\perp$ and in $C_X \setminus C_Z^\perp$.

Proof. Because $\mathbf{1} \in C_A$, the components in every given codeword of $C_A^\perp$ sum to 0. Therefore every codeword in $C_A^\perp \otimes C_B^\perp$, and thus also in $C_Z^\perp$, has components summing to 0, as $C_Z^\perp$ is by definition

spanned by codewords in $C_A^\perp \otimes C_B^\perp$ supported in neighborhoods of vertices in the square Cayley complex. Thus as the components of $\mathbf{1} \in \mathbb{F}_q^Q$ sum to $n \neq 0$ in $\mathbb{F}_q$ because n is relatively prime with q, it follows that $\mathbf{1} \notin C_Z^\perp$.

However, as $\mathbf{1} \in \mathbb{F}_q^\Delta$ lies in $C_B^\perp$, it follows that $\mathbf{1} \in \mathbb{F}_q^{\Delta \times \Delta}$ lies in $(C_A \otimes C_B)^\perp$, and thus $\mathbf{1} \in \mathbb{F}_q^Q$ lies in $C_X = \mathrm{Tan}(\Gamma_0, (C_A \otimes C_B)^\perp)$.

Thus we have shown that $\mathbf{1} \in C_X \setminus C_Z^\perp$. Analogous reasoning shows that $\mathbf{1} \in C_Z \setminus C_X^\perp$. □

To instantiate the construction in Proposition 30 such that Condition 1 is satisfied, we will choose C_A and $C_B^\perp$ at random from the set of codes of some constant rate that contain $\mathbf{1}$. The following result, which we prove in Section 5.3, shows that such random "planted" codes are still product-expanding, thereby providing a planted analogue of Corollary 11.

Proposition 31. *Fix any finite field $\mathbb{F}_q$. For every fixed $\epsilon > 0$, there exists a constant $\rho = \rho(\epsilon) > 0$ and a function $\delta(n) = \delta(n; \epsilon) \to 0$ as $n \to \infty$ such that the following holds. For every pair of integers $k_1, k_2 \in (\epsilon n, (1-\epsilon)n)$, if $C_i \subseteq \mathbb{F}_q^n$ for $i = 1, 2$ is drawn uniformly at random from the set of linear codes of dimension k_i that contain $\mathbf{1} \in \mathbb{F}_q^n$, then with probability $\geq 1 - \delta(n)$ both $(C_1, C_2^\perp)$ and $(C_1^\perp, C_2)$ will be ρ-product-expanding.*

Meanwhile, to ensure that Condition 2 in Proposition 30 is satisfied, we will choose the graphs $\mathrm{Cay}(G, A), \mathrm{Cay}(G, B)$ to be almost-Ramanujan graphs from the family given by Theorem 14, which we restate below and prove in Section 5.4:

Theorem 14 (Follows from [LW93, JMRW22])**.** *For every prime p, there is an infinite set $\boldsymbol{\Delta} \subseteq \mathbb{N}$ for which there exists a strongly explicit family of almost-Ramanujan Cayley graphs $(\Gamma_{m,\Delta})_{m \in \mathbb{N}, \Delta \in \boldsymbol{\Delta}}$, where $\Gamma_{m,\Delta}$ has $|V(\Gamma_{m,\Delta})| = p^{3m}$ vertices and has degree Δ. Furthermore, we may choose $\boldsymbol{\Delta}$ such that for every $\Delta \in \boldsymbol{\Delta}$, either $\Delta + 1 \in \boldsymbol{\Delta}$ or $\Delta - 1 \in \boldsymbol{\Delta}$.*

Combining the results above, we immediately obtain the following strongly explicit construction of quantum Tanner codes with a planted all-1s vector.

Theorem 32 (Planted quantum Tanner codes)**.** *For every finite field $\mathbb{F}_q$, there exist constants $c_1, c_2 > 0$ such that there is a strongly explicit infinite family $(\mathcal{C}^{(n)})_{n \to \infty}$ of quantum LDPC CSS codes for which every $\mathcal{C}^{(n)} = \mathrm{CSS}(C_X^{(n)}, C_Z^{(n)})$ with $C_X^{(n)}, C_Z^{(n)} \subseteq \mathbb{F}_q^n$ has the following properties:*

1. *$\mathcal{C}^{(n)}$ has (c_1, c_2)-small-set boundary and coboundary expansion (and therefore has distance $\geq c_1 n$ by Lemma 17).*
2. *The all-1s vector $\mathbf{1} \in \mathbb{F}_q^n$ lies in $C_X^{(n)} \setminus C_Z^{(n)\perp}$ and in $C_Z^{(n)} \setminus C_X^{(n)\perp}$.*

In particular, for a sufficiently large constant Δ and a sufficiently small constant $\rho > 0$, such a family $(\mathcal{C}^{(n)})_{n \to \infty}$ is given by quantum Tanner codes, where we choose $\mathrm{Cay}(G, A) = \mathrm{Cay}(G, B)$ from a strongly explicit family of Δ-regular almost-Ramanujan graphs given by Theorem 14, and the inner codes $C_A, C_B \subseteq \mathbb{F}_q^\Delta$ are found by a brute force search to ensure that $\mathbf{1} \in C_A, C_B^\perp$ and that $(C_A, C_B), (C_A^\perp, C_B^\perp)$ are ρ-product expanding.

Proof. By Proposition 31, if $\rho > 0$ is sufficiently small and $\Delta > 0$ is sufficiently large then we can find codes $C_A, C_B \subseteq \mathbb{F}_q^\Delta$ satisfying the criteria in the theorem statement, namely that $\mathbf{1} \in C_A, C_B^\perp$ and that $(C_A, C_B), (C_A^\perp, C_B^\perp)$ are ρ-product expanding.

Furthermore, choose any fixed prime p relatively prime with q, and let Δ_0 be sufficiently large and $c_1, c_2 > 0$ be sufficiently small constants such that Theorem 18 implies that all $\mathcal{C}^{(n)}$ have (c_1, c_2)-small-set boundary and coboundary expansion as long as $\mathrm{Cay}(G, A) = \mathrm{Cay}(G, B)$ is chosen to be an almost-Ramanujan graph of degree $\Delta \geq \Delta_0$ from the family $(\Gamma_{m,\Delta})_{m\in\mathbb{N}}$ given by Theorem 14, where $|V(\Gamma_{m,\Delta})| = p^{3m}$. Because Theorem 14 guarantees that $\mathbf{\Delta} \subseteq \mathbb{N}$ is infinitely large and consists of the union of pairs of consecutive integers, we can always find some sufficiently large $\Delta \in \mathbf{\Delta}$ that is relatively prime with q and that satisfies $\Delta \geq \Delta_0$. Thus we indeed obtain the desired Cayley expanders $\mathrm{Cay}(G, A) = \mathrm{Cay}(G, B) = \Gamma_{m,\Delta}$ for $m \in \mathbb{N}$, where $|G||A||B| = p^{3m}\Delta^2$ is relatively prime with q.

Now we have shown that the instantiation of quantum Tanner codes above satisfies Conditions 1, 2 in Proposition 30, so this proposition implies that $\mathbf{1} \in \mathbb{F}_q^n$ lies in $C_X^{(n)} \setminus C_Z^{(n)\perp}$ and in $C_Z^{(n)} \setminus C_X^{(n)\perp}$. Meanwhile, as described above, Theorem 18 implies that $\mathcal{C}^{(n)}$ has (c_1, c_2)-small-set boundary and coboundary expansion.

Because the almost-Ramanujan graphs in Theorem 14 are strongly explicit, and the inner codes C_A, C_B have constant size beacuse Δ is constant as $n \to \infty$, the parity check matrices $H_X^{(n)}, H_Z^{(n)}$ for $C_X^{(n)}, C_Z^{(n)}$ respectively are strongly explicit, which by definition means that $\mathcal{C}^{(n)}$ is strongly explicit. □

In Theorem 32, we may choose C_A, C_B to have any fixed rates $0 < R_A, R_B < 1$ for sufficiently large Δ. Because $\mathcal{C}$ has rate $R \geq -(1 - 2R_A)(1 - 2R_B)$ (see Section 3.2), it follows that we can in fact ensure that the the codes in Theorem 32 have any desired constant rate $0 < R < 1$.

However, our construction alternatively allows us to obtain quantum Tanner codes of positive dimension for R_A, R_B in previously impossible parameter regimes. Because Theorem 32 ensures that $\mathbf{1} \in C_Z \setminus C_X^\perp$, it follows that $\mathcal{C}$ always has dimension $\dim(\mathcal{C}) = \dim(C_Z) - \dim(C_X^\perp) \geq 1$, even when we choose R_A, R_B to be constants for which the bound $R \geq -(1 - 2R_A)(1 - 2R_B)$ is meaningless. For instance, we can choose $R_A = R_B$, or take both $R_A, R_B < 1/2$, in which case counting parity checks fails to show that the resulting quantum Tanner code $\mathcal{C}$ has positive dimension. Nevertheless the planted all-1s vector ensures that even in this case $\mathcal{C}$ must have dimension ≥ 1.

5.3 Proof of Product-Expansion for Planted Inner Codes

In this section we prove Proposition 31. We begin with the following lemma.

Lemma 33. *If (C_1, C_2) is ρ-product expanding and $C_1' \subseteq C_1$ is a codimension-1 subcode, then (C_1', C_2) is $\rho^2/2$-product expanding.*

Proof. Fix an arbitrary $x \in C_1' \otimes \mathbb{F}_q^n + \mathbb{F}_q^n \otimes C_2$. Our goal is to show that there exists a decomposition $x = c + r$ for some $c \in C_1' \otimes \mathbb{F}_q^n$ and $r \in \mathbb{F}_q^n \otimes C_2$ satisfying

$$|x| \geq \frac{\rho^2 n}{2}(|c|_{\mathrm{col}} + |r|_{\mathrm{row}}). \tag{4}$$

If $|x| \geq \rho^2 n^2$, then any decomposition $x = c + r$ suffices, as the right hand side above is always at most $\rho^2 n/2 \cdot 2n = \rho^2 n^2$.

Therefore assume that $|x| < \rho^2 n$. By the ρ-product expansion of (C_1, C_2), there exists some decomposition $x = c + r$ for $c \in C_1 \otimes \mathbb{F}_q^n$ and $r \in \mathbb{F}_q^n \otimes C_2$ such that

$$\rho^2 n > |x| \geq \rho n(|c|_{\text{col}} + |r|_{\text{row}}). \tag{5}$$

Meanwhile, by the definition of x, there exists some decomposition $x = c' + r'$ for $c' \in C_1' \otimes \mathbb{F}_q^n$ and $r' \in \mathbb{F}_q^n \otimes C_2$. Letting $y = c - c' = r' - r$, then $y \in C_1 \otimes \mathbb{F}_q^n$ and $y \in \mathbb{F}_q^n \otimes C_2$, so $y \in C_1 \otimes C_2 = C_1' \otimes C_2 + \text{span}\{a\} \otimes C_2$. Therefore we can decompose $y = w + z$ for $w \in C_1' \otimes C_2$ and $z \in \text{span}\{a\} \otimes C_2$. It follows that $c = c' + y = (c' + w) + z$, where $c' + w \in C_1' \otimes \mathbb{F}_q^n$ and $z \in \text{span}\{a\} \otimes C_2$. That is, every column of $c' + w$ lies in C_1', and every nonzero column of z is a scalar multiple of $a \notin C_1'$. Thus the ith column of c is not in C_1' if and only if the ith column of z is nonzero. But by Lemma 9, C_2 has distance $\geq \rho n$, and therefore if $z \in \text{span}\{a\} \otimes C_2$ is nonzero then z has $\geq \rho n$ nonzero columns, which implies that c has $\geq \rho n$ columns that are not in C_1', so in particular c has $\geq \rho n$ nonzero columns. But this assertion that $|c|_{\text{col}} \geq \rho n$ contradicts (5), so our assumption that z is nonzero must have been false. Therefore $z = 0$, so $c = c' + w \in C_1' \otimes \mathbb{F}_q^n$. Thus $x = c + r$ provides our desired decomposition of x satisfying (4), as we have shown that $c \in C_1' \otimes \mathbb{F}_q^n$ and $r \in \mathbb{F}_q^n \otimes C_2$ in fact satisfy the stronger inequality (5). □

In light of Lemma 33, we can reduce the problem of proving Propositive 31 to proving the following result.

Proposition 34. *For every fixed $\epsilon > 0$, there exists a constant $\rho = \rho(\epsilon) > 0$ and a function $\delta(n) = \delta(n; \epsilon) \to 0$ as $n \to \infty$ such that the following holds. For every pair of integers $k_1, k_2 \in (\epsilon n, (1-\epsilon)n)$, if $C_1 \subseteq \mathbb{F}_q^n$ is drawn uniformly at random from the set of all linear codes of dimension k_1, and $C_2 \subseteq \mathbb{F}_q^n$ is drawn uniformly at random from the set of linear codes of dimension k_2 that contain $\mathbf{1} \in \mathbb{F}_q^n$, then with probability $\geq 1 - \delta(n)$ the pair (C_1, C_2) will be ρ-product-expanding.*

We first show how Proposition 34 implies Proposition 31, and then we will prove Proposition 34

Proof of Proposition 31. Let $C_1, C_2 \subseteq \mathbb{F}_q^n$ be random codes as in the statement of Proposition 31. Let p be the probability that a random code in $\mathbb{F}_q^n$ of dimension $k_1' = k_1 - 1$ contains $\mathbf{1} \in \mathbb{F}_q^n$. We draw a k_1'-dimensional subcode $C_1' \subseteq C_1$ as follows:

- With probability $1 - p$, C_1' is drawn uniformly at random from the codimension-1 subcodes of C_1 that do not contain $\mathbf{1}$.
- With probability p, C_1' is drawn uniformly at random from the codimension-1 subcodes of C_1 that contain $\mathbf{1}$.

Then because C_1 is a random k_1-dimensional code containing $\mathbf{1}$, by construction C_1' is a uniformly random k_1'-dimensional code (correlated with C_1), as can be seen by conditioning on the events that $\mathbf{1} \in C_1'$ and $\mathbf{1} \notin C_1'$. Thus Proposition 34 implies that there is a sufficiently small constant $\rho' = \rho'(\epsilon) > 0$ such that $(C_1'^{\perp}, C_2)$ is ρ'-product-expanding with probability $\to 1$ as $n \to \infty$. Lemma 33 then implies that $(C_1^{\perp}, C_2)$ is $\rho := \rho'^2/2$-product-expanding with probability $\to 1$ as $n \to \infty$.

By similar reasoning as above, $(C_1, C_2^{\perp})$ is also $\rho = \rho'^2/2$-product-expanding with probability $\to 1$ as $n \to \infty$. Then the desired result follows by union bounding over the events that both $(C_1^{\perp}, C_2)$ and $(C_1, C_2^{\perp})$ are ρ-product expanding. □

It remains to prove Proposition 34. For this purpose, we adapt the proof of Proposition 10 given by Kalachev and Panteleev [KP23]. Specifically, [KP23] show that a pair of random codes (C_1, C_2) of any fixed rates < 1 are product-expanding with high probability for sufficiently large block lengths; we want to show that (C_1, C_2) are still product-expanding with high probability when conditioning on the event that $\mathbf{1} \in C_2$. To avoid redundancy with [KP23], we will simply describe the necessary modifications to their proof of Proposition 10 (Theorem 1 in their paper [KP23]) for this case where we condition on $\mathbf{1} \in C_2$.

To begin, we recall the following definitions from [KP23]. Below we denote the q-ary entropy function by $H_q : [0,1] \to [0,1]$, so that

$$H_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x).$$

We also say that a subspace $V \subseteq \mathbb{F}_q^n$ is *α-sparse* if it is spanned by vectors of Hamming weight $\leq \alpha n$.

Definition 35 (Property $(*)$ [KP23])**.** A code $C \subseteq \mathbb{F}_q^n$ of dimension $n - r$ has **property** $(*)$ if the following holds for $\alpha = H_q^{-1}(r/8n)$: for every $m \in \{1, \ldots, r\}$ and every α-sparse m-dimensional subspace $V \subseteq \mathbb{F}_q^n$, then $\dim(C \cap V) < m/2$.

[KP23] prove Proposition 10 (their Theorem 1) by consider the following two events E_1, E_2 for a pair of random codes $C_1, C_2 \subseteq \mathbb{F}_q^n$ of respective dimensions $k_1, k_2 \in (\epsilon n, (1-\epsilon)n)$. Recall here that $\epsilon > 0$ is any fixed constant, and $\rho = \rho(\epsilon) > 0$ is a sufficiently small constant depending only on ϵ.

1. E_1 is the event that there exists $x \in C_1 \otimes \mathbb{F}_q^\perp + \mathbb{F}_q^\perp \otimes C_2$ of weight $|x| \leq 2\rho n^2$ and of rank $\operatorname{rank}(x) \geq \epsilon$.
2. E_2 is the event that C_1 or C_2 does not have property $(*)$.

To prove Proposition 10, [KP23] show the following:

Lemma 36 ([KP23])**.** *If $C_1, C_2 \subseteq \mathbb{F}_q^n$ are random codes of respective dimensions $k_1, k_2 \in (\epsilon n, (1-\epsilon)n)$, then for all sufficiently large n,*

1. $\Pr[E_1] \leq 5q^{-\epsilon^2 n^2/8}$.
2. $\Pr[E_2] \leq 16q^{-\epsilon n/8}$.
3. *If E_1, E_2 do not occur then (C_1, C_2) is ρ-product-expanding.*

Lemma 36 directly implies Proposition 10. Similarly, the following lemma directly implies Proposition 34:

Lemma 37 ([KP23])**.** *If $C_1, C_2 \subseteq \mathbb{F}_q^n$ are random codes of respective dimensions $k_1, k_2 \in (\epsilon n, (1-\epsilon)n)$, and F denotes the event that $\mathbf{1} \in C_2$, then for all sufficiently large n,*

1. $\Pr[E_1 | F] \leq 5q^{-\epsilon^2 n^2/8+n}$.
2. $\Pr[E_2 | F] \leq 16q^{-\epsilon n/64}$.
3. *If E_1, E_2 do not occur then (C_1, C_2) is ρ-product-expanding.*

Item 3 in Lemma 36 and item 3 in Lemma 37 are identical. Furthermore, as $\Pr[F] \geq q^{-n}$, then $\Pr[E_1|F] \leq \Pr[E_1]/\Pr[F] \leq \Pr[E_1] \cdot q^n$, so item 1 in Lemma 36 directly implies item 1 in Lemma 37.

Therefore to prove Lemma 37 and therefore Proposition 34, it remains for us to prove item 2 of Lemma 37. That is, it suffices to show that conditioned on F, each of C_1 and C_2 has property $(*)$ with probability $\geq 1 - 8q^{-\epsilon n/64}$. As C_1 is a random code, Lemma 5 in [KP23] (which they use to show item 2 in Lemma 36) implies that C_1 has property $(*)$ with probability $\geq 1 - 8q^{-\epsilon n/8}$. Thus the following lemma completes the proof of Lemma 37, and therefore of Proposition 34, and in turn of Proposition 31.

Lemma 38. *If $C \subseteq \mathbb{F}_q^n$ is drawn uniformly at random from the set of $k = (n-r)$-dimensional codes that contain $\mathbf{1} \in \mathbb{F}_q^n$, then C has property $(*)$ with probability at least*

$$1 - \frac{4q^{-r/64}}{1 - q^{-r/64}}.$$

To prove Lemma 38, we will use the following technical result, which appears as Lemma 3 in [KP23].

Lemma 39 ([KP23])**.** *For every v-dimensional subspace $V \subseteq \mathbb{F}_q^n$, the probability that a uniformly random u-dimensional subspace $U \subseteq \mathbb{F}_q^n$ has $\dim(U \cap V) \geq w$ is $\leq 4q^{-w(n+w-v-u)}$.*

We will also need the following result on random planted codes.

Lemma 40. *Let $C \subseteq \mathbb{F}_q^n$ be drawn uniformly at random from the set of k-dimensional codes that contain $\mathbf{1} \in \mathbb{F}_q^n$. Then for every fixed $(n-1)$-dimensional subspace $W \subseteq \mathbb{F}_q^n$ such that $\mathbf{1} \notin W$, the intersection $C \cap W$ is a uniformly random $(k-1)$-dimensional subspace of W.*

Proof. Because $\mathbf{1} \notin W$ and $\mathbf{1} \in C$, we must have $\dim(C \cap W) = k - 1$, so $C = \operatorname{span}\{\mathbf{1}\} \oplus (C \cap W)$. That is, C is uniquely determined from $C \cap W$, and each $(k-1)$-dimensional subspace $W' \subseteq W$ gives a distinct $C = \operatorname{span}\{\mathbf{1}, W'\}$. Thus because C is drawn uniformly from its allowable set of vector spaces, $\Pr[C \cap W = W']$ is the same for all W'. □

We are now ready to prove Lemma 38.

Proof of Lemma 38. We simply modify the proof of Lemma 5 in [KP23] to account for the condition that $\mathbf{1} \in C$. Recall that $\alpha = H_q^{-1}(r/8n)$. Let $C \subseteq \mathbb{F}_q^n$ be a uniformly random code that contains $\mathbf{1} \in \mathbb{F}_q^n$. We want to show that with high probability over the choice of C, it holds for every $m \in \{1, \ldots, r\}$ and every α-sparse m-dimensional subspace $V \subseteq \mathbb{F}_q^n$ that $\dim(C \cap V) < m/2$.

Fix such an m and V. We will first show that

$$\Pr[\dim(C \cap V) \geq m/2] \leq 4q^{-\frac{9}{64}mr}. \tag{6}$$

Union bounding over all m and V with this inequality will then give the desired result.

To prove (6), we consider the the cases of small m and large m separately. In both cases, we will use Lemma 40 to relate $C \cap V$ to the intersection of a truly random code with V, so that we can apply Lemma 39 to bound $\Pr[\dim(C \cap V)] \geq m/2$.

1. Assume that $m < 1/\alpha$. As V has a basis consisting of m α-sparse vectors, all elements of V are supported within $\leq \alpha n \cdot m < n$ components, so there is some component $i \in [n]$ on which all $x \in V$ have $x_i = 0$. That is, letting $\Pi_i : \mathbb{F}_q^n \to \mathbb{F}_q$ denote projection onto component i, then $V \subseteq \Pi_i^{-1}(0)$.

   Now as $\mathbf{1} \notin \Pi_i^{-1}(0)$, Lemma 40 implies that $C \cap \Pi_i^{-1}(0)$ is a uniformly random $(k-1)$-dimensional subspace of $\Pi_i^{-1}(0)$.

   Letting p be the probability that a random k-dimensional subspace of $\mathbb{F}_q^n$ lies inside $\Pi_i^{-1}(0)$, we may sample

$$c \sim \begin{cases} \mathrm{Unif}(\Pi_i^{-1}(0) \setminus (C \cap \Pi_i^{-1}(0))), & \text{with probability } p \\ \mathrm{Unif}(\Pi_i^{-1}(\mathbb{F}_q \setminus \{0\})), & \text{with probability } 1-p, \end{cases}$$

   so that the vector space

$$U := \mathrm{span}\{c, (C \cap \Pi_i^{-1}(0))\}$$

   is a uniformly random k-dimensional subspace of $\mathbb{F}_q^n$. Then because $V \subseteq \Pi_i^{-1}(0)$ so that

$$C \cap V = C \cap \Pi_i^{-1}(0) \cap V \subseteq U \cap V,$$

   it follows by Lemma 39 that

$$\begin{aligned} \Pr[\dim(C \cap V) \geq m/2] &\leq \Pr[\dim(U \cap V) \geq m/2] \\ &\leq 4q^{-m/2(n+m/2-m-k)} \\ &\leq 4q^{-mr/4}, \end{aligned}$$

   where the final inequality above holds because $k = n - r$ and $m \in \{1, \ldots, r\}$. Thus (6) holds in this case where $m < 1/\alpha$.

2. Assume that $m \geq 1/\alpha$. Again sampling U as in the $m < 1/\alpha$ case above, then U, C are k-dimensional spaces with $\dim(U \cap C) \geq k - 1$. Thus $\dim(C \cap V) \leq \dim(U \cap V) + 1$, so by Lemma 39,

$$\begin{aligned} \Pr[\dim(C \cap V) \geq m/2] &\leq \Pr[\dim(U \cap V) \geq m/2 - 1] \\ &\leq 4q^{-(m/2-1)(n+m/2-1-m-k)} \\ &\leq 4q^{-(1/2-\alpha)m(n+(1/2-\alpha)m-m-(n-r))} \\ &\leq 4q^{-\frac{3}{8}m\left(r-\frac{5}{8}m\right)} \\ &\leq 4q^{-\frac{9}{64}mr}. \end{aligned}$$

   where the third inequality above holds because $m \geq 1/\alpha$ and $k = n - r$, and the fourth inequality holds because $\alpha = H_q^{-1}(r/8n) \leq H_q^{-1}(1/8) \leq 1/8$, and the fifth inequality holds because $m \in \{1, \ldots, r\}$. Thus (6) holds in this case where $m \geq 1/\alpha$.

Now it is well-known that there are $|\{x \in \mathbb{F}_q^n : |x| \leq \alpha n\}| \leq q^{H_q(\alpha)n}$ distinct α-sparse vectors (see for instance Proposition 3.3.3 of [GRS22]), so there are at most $q^{mH_q(\alpha)} = q^{mr/8n}$ distinct α-sparse m-dimensional subspace $V \subseteq \mathbb{F}_q^n$, where we are using the definition of $\alpha = H_q^{-1}(r/8n)$.

Union bounding over all such V, it follows from (6) that the probability that there exists an α-sparse m-dimensional subspace V with $\dim(C \cap V) \geq m/2$ is at most

$$q^{mr/8n} \cdot 4q^{-\frac{9}{64}mr} = 4q^{-mr/64}.$$

Finally union bounding over $m \in \{1, \ldots, r\}$, we conclude that the probability that that there exists some $m \in \{1, \ldots, r\}$ with some α-sparse m-dimensional subspace V with $\dim(C \cap V) \geq m/2$ is at most

$$\sum_{m=1}^{r} 4q^{-mr/64} \leq \frac{4q^{-r/64}}{1 - q^{-r/64}}.$$

Thus C has property $(*)$ with probability $\geq 1 - 4q^{-r/64}/(1 - q^{-r/64})$, as desired. □

5.4 Construction of Strongly Explicit Expanders

In this section, we present the proof of Theorem 14, which follows from the expander construction of [LW93] along with the expansion amplification technique of [JMRW22]. However, there are some details we need to verify to prove Theorem 14, specifically that the construction of [LW93] is strongly explicit.

We begin by describing the Cayley expanders given in Example 3.4 of [LW93], which are notable because the number of vertices equals a power of any desired prime p. In particular, by choosing an odd prime, we obtain Cayley expanders on an odd number of vertices, which is important for our planted quantum Tanner codes over $\mathbb{F}_2$ due to Condition 2 in Proposition 30. In contrast, the Ramanujan Cayley graphs of [LPS88] and [Mor94] have an even number of vertices.

5.4.1 Base Expander Construction [LW93]

This section presents the Cayley expanders given in Example 3.4 of [LW93]. For simplicity we restrict attention to expanders constructed over SL_2, but [LW93] shows that a similar construction over SL_k for any $k \geq 2$ also yields Cayley expanders.

Fix a prime p. For every $t \in \mathbb{N}$, define

$$G(t) = \ker(SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/t\mathbb{Z})),$$

where the homomorphism on the right hand side above simply maps all matrix entries to their value (mod t). [LW93] shows that there exists a finite generating set $S = S(p)$ for $G(p)$. They then define the Cayley graphs

$$\Gamma_m^0 = \mathrm{Cay}(G(p)/G(p^{m+1}), S), \tag{7}$$

for which they show the following:

Theorem 41 ([LW93])**.** *There exists a constant $\lambda < \Delta$ such that each Cayley graph Γ_m^0 given by (7) has spectral expansion $\leq \lambda$. Furthermore, Γ_m^0 has p^{3m} vertices and is $\Delta = |S|$-regular.*

5.4.2 Strong Explicitness of Base Expander Construction

In this section, we show that the expanders in Section 5.4.1 are strongly explicit.

We begin with the following well-known lemma. For completeness we present a proof.[4]

Lemma 42 (Well known)**.** *The map $SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/t\mathbb{Z})$ is surjective.*

Proof. Given any matrix $A \in SL_2(\mathbb{Z}/t\mathbb{Z})$, we can perform row reduction to obtain a diagonal matrix $A' = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ for some $a \in \mathbb{Z}/t\mathbb{Z}$, which means that A equals A' times a product of elementary matrices (i.e. matrices with all 1s on the diagonal, and with a single nonzero off-diagonal entry). But it also holds that

$$A' = \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -a^{-1} & 1 \end{pmatrix} \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Thus $A \in SL_2(\mathbb{Z}/t\mathbb{Z})$ is a product of elementary matrices along with (possibly) the matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. But all of these matrices also belong to $SL_2(\mathbb{Z})$, so A is the image under the map $SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/t\mathbb{Z})$ of the product of these matrices in $SL_2(\mathbb{Z})$. Thus this map is surjective, as desired. □

We now show the following lemma, which provides a more tractable characterization of the group $G(p)/G(p^{m+1})$. The proof is fairly standard, but we provide it for completeness.

Lemma 43. *The natural map $G(p) \to SL_2(\mathbb{Z}/p^{m+1})$ induces an isomorphism*

$$G(p)/G(p^{m+1}) \cong \ker(SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}) \to SL_2(\mathbb{Z}/p\mathbb{Z})). \tag{8}$$

Proof. Consider the sequence of homormorphisms

$$SL_2(\mathbb{Z}) \xrightarrow{\phi_1} SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}) \xrightarrow{\phi_2} SL_2(\mathbb{Z}/p\mathbb{Z}).$$

By Lemma 42, ϕ_1 and $\phi_2 \circ \phi_1$ are surjective, so ϕ_2 is also surjective. By definition $G(p^{m+1}) = \ker \phi_1$ and $G(p) = \ker(\phi_2 \circ \phi_1)$. Now the surjectivity of ϕ_1 implies that ϕ_1 induces an isomorphism

$$\tilde{\phi}_1 : SL_2(\mathbb{Z})/\ker \phi_1 \xrightarrow{\sim} SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}).$$

As $G(p)/G(p^{m+1}) = \ker(\phi_2 \circ \phi_1)/\ker(\phi_1)$ is a subgroup of $SL_2(\mathbb{Z})/\ker \phi_1$, we obtain a restricted isomorphism

$$G(p)/G(p^{m+1}) = \ker(\phi_2 \circ \phi_1)/\ker(\phi_1) \xrightarrow{\sim} \tilde{\phi}_1(\ker(\phi_2 \circ \phi_1)/\ker(\phi_1)).$$

But the right hand side above by definition equals $\ker(\phi_2)$, as it holds that $x \in \tilde{\phi}_1(\ker(\phi_2 \circ \phi_1)/\ker(\phi_1))$ if and only if $\phi_2(x)$ is the identity. Thus we have shown that ϕ_1 induces a natural isomorphism

$$G(p)/G(p^{m+1}) \cong \ker(\phi_2),$$

as desired. □

[4] This proof is fairly standard, though our presentation follows ideas sugggested by [Jul13, Ebe13].

Let G_m denote the group in (8), so that the expanders of [LW93] described in Section 5.4.1 are Cayley graphs $\mathrm{Cay}(G_m, S)$ for $m \in \mathbb{N}$. The following lemma shows that we can enumerate elements of G_m using 3-tuples of elements of $\mathbb{Z}/p^m\mathbb{Z}$.

Lemma 44. *There is a bijection*

$$\phi : (\mathbb{Z}/p^m\mathbb{Z})^3 \to G_m = \ker(SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}) \to SL_2(\mathbb{Z}/p\mathbb{Z}))$$

given by

$$\phi(a,b,c) = I + p\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

for $d \in \mathbb{Z}/p^m\mathbb{Z}$ given by

$$d = (1+pa)^{-1}(pbc - a). \tag{9}$$

Proof. By definition $1+pa$ is invertible in $\mathbb{Z}/p^m\mathbb{Z}$, so d is well defined, and we have

$$\phi(a,b,c) = \det\left(I + p\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (1+pa)(1+pd) - p^2bc = 1,$$

so indeed ϕ maps (a,b,c) to an element of G_m. We must verify that ϕ is injective and surjective.

To see that ϕ is injective, observe that if $(a,b,c) \neq (a',b',c')$, then (pa,pb,pc) and (pa',pb',pc') are distinct tuples in $(\mathbb{Z}/p^{m+1})^3$, so $\phi(a,b,c)$ and $\phi(a',b',c')$ are distinct matrices in $(\mathbb{Z}/p^{m+1}\mathbb{Z})^{2\times 2}$.

To see that ϕ is surjective, consider that every matrix $M \in G_m = \ker(SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}) \to SL_2(\mathbb{Z}/p\mathbb{Z}))$ is by definition of the form $M = I + p\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a,b,c,d \in \mathbb{Z}/p^m\mathbb{Z}$. Now because this matrix has determinant 1, we have $(1+a)(1+d) - bc = 1$, which simplifies to (9), so $M = \phi(a,b,c)$. Thus ϕ is surjective, as desired. □

As a side note, Lemma 44 also recovers the fact that $|G_m| = |\mathbb{Z}/p^m\mathbb{Z}|^3 = p^{3m}$.

We are now ready to show that the expanders described in Section 5.4.1 are strongly explicit.

Proposition 45. *For every fixed p, the family of Cayley graphs $(\Gamma_m^0 = \mathrm{Cay}(G_m, S))_{m\in\mathbb{N}}$ presented in Section 5.4.1 is strongly explicit.*

Proof. By Lemma 43, we have $G_m \cong \ker(SL_2(\mathbb{Z}/p^{m+1}\mathbb{Z}) \to SL_2(\mathbb{Z}/p\mathbb{Z}))$, and the generating set S consists of a fixed finite set of matrices in $\mathbb{Z}^{n\times n}$, which can be viewed as matrices in G_m by replacing each entry with its value $\pmod{p^{m+1}}$. Note that S depends on p, but here we assume p is fixed, and S does not depend on m (except for our interpretation of its entries as integers $\pmod{p^{m+1}}$).

Now Lemma 44 implies that the elements of G_m are index by tuples $(a,b,c) \in (\mathbb{Z}/p^m\mathbb{Z})^3$, or equivalently, by tuples $(a,b,c) \in \{0,1,\dots,p^m-1\}^3$. Furthermore, we can go between the tuple representation (a,b,c) and the matrix representation $\phi(a,b,c)$ defined in Lemma 44 in $\mathrm{poly}(\log p^m) = \mathrm{poly}(m)$ time, as the conversion simply requires computing $d = (1+pa)^{-1}(pbc-a)$ and then multiplying $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ by p and adding I; the reverse conversion (from matrix to tuple) is similarly efficient. Thus we can perform group multiplication and inversion in time $\mathrm{poly}(m)$, as these operations are simply given by matrix multiplication and inversion.

Thus we have shown that the group operations of G_m run in $\mathrm{poly}(\log|G_m|)$ time, and the generating set S can be computed in constant time, so the family of Cayley graphs $\Gamma_m^0 = \mathrm{Cay}(G_m, S)$ for $m \in \mathbb{N}$ is strongly explicit, as desired. □

5.4.3 Amplifying the Expansion to Almost-Ramanujan

In this section, we prove Theorem 14 by applying the expansion amplification technique of [JMRW22]. In particular, [JMRW22] show the following (see for instance their Theorem 1.2).

Theorem 46 ([JMRW22])**.** *Let $(\Gamma_m^0 = \mathrm{Cay}(G_m, S_m))_{m\in\mathbb{N}}$ be a strongly explicit family of Δ_0-regular Cayley graphs with spectral expansion bounded by some constant $\lambda < \Delta_0$. Then there exists an infinite set $\boldsymbol{\Delta} \subseteq \mathbb{N}$ for which there is a strongly explicit family $(\Gamma_{m,\Delta} = \mathrm{Cay}(G_m, S_{m,\Delta}))_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}}$ of almost-Ramanujan Cayley graphs, where $\Gamma_{m,\Delta}$ has degree $|S_{m,\Delta}| = \Delta$.*

Theorem 14 now follows almost immediately from Proposition 45 and Theorem 46:

Proof of Theorem 14. We apply Theorem 46 to the Cayley graphs $(\Gamma_m^0 = \mathrm{Cay}(G_m, S))_{m\in\mathbb{N}}$ presented in Section 5.4.1. These graphs have constant degree by definition, have constant spectral expansion $\lambda < \Delta$ by Theorem 41, and are strongly explicit by Proposition 45. Therefore Theorem 46 gives a strongly explicit family $(\Gamma_{m,\Delta} = \mathrm{Cay}(G_m, S_{m,\Delta}))_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}}$ of almost-Ramanujan Cayley graphs over the groups G_m given by (8), which have order p^{3m}.

It only remains to ensure that $\boldsymbol{\Delta}$ has the property that if $\Delta \in \boldsymbol{\Delta}$, then either $\Delta - 1 \in \boldsymbol{\Delta}$ or $\Delta+1 \in \boldsymbol{\Delta}$. For this purpose, for each graph $\Gamma_{m,\Delta} = \mathrm{Cay}(G_m, S_{m,\Delta})$ in our family such that $\Delta \in \boldsymbol{\Delta}$ but $\Delta + 1 \notin \boldsymbol{\Delta}$, we may also add the graph $\Gamma_{m,\Delta+1} := \mathrm{Cay}(G_m, S_{m,\Delta} \cup \{\mathrm{id}\})$ obtained by adding the identity element as a Cayley generator. The resulting family of graphs $\Gamma_{m,\Delta}$ ranges over the possible degrees $\Delta \in \boldsymbol{\Delta}' := \boldsymbol{\Delta} \cup \{\Delta + 1 : \Delta \in \boldsymbol{\Delta}\}$, which has the desired property that if $\Delta \in \boldsymbol{\Delta}'$, then either $\Delta - 1 \in \boldsymbol{\Delta}'$ or $\Delta + 1 \in \boldsymbol{\Delta}'$.

We must verify that this larger family $(\Gamma_{m,\Delta})_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}'}$ is still almost-Ramanujan. Adding the identity element as a Cayley generator increases the graph degree by 1, and increases the spectral expansion by at most 1. This latter claim holds because adding the identity to the Cayley generating set has the effect of adding the identity matrix to the adjacency matrix, which increases all eigenvalues by 1, and thus increases the spectral expansion by at most 1. Therefore if the original degree-Δ graphs $\Gamma_{m,\Delta}$ have spectral expansion at most $\lambda(\Delta) = \Delta^{1/2+o(1)}$, then the added degree-$(\Delta + 1)$ graphs have spectral expansion at most $\lambda(\Delta + 1) := \lambda(\Delta) + 1$, which still grows as $(\Delta + 1)^{1/2+o(1)}$. Thus our final augmented family $(\Gamma_{m,\Delta})_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}'}$ is almost-Ramanujan, as desired. Note that the strong explicitness of this augmented family is immediate from the strong explicitness of the original family $(\Gamma_{m,\Delta})_{m\in\mathbb{N},\Delta\in\boldsymbol{\Delta}}$. □

Remark 47. The expansion amplification of [JMRW22] in Theorem 46, along with our technique of adding the identity to to the Cayley generating set in the proof above, assume that the Cayley generating sets of our graphs are actually *multisets*. That is, we allow repeated Cayley generators, so our Cayley graphs are actually multigraphs, meaning there can be multiple distinct edges between the same two vertices.

Fortunately, all of our applications of these graphs apply equally well to multigraphs and simple (non-multi) graphs. Indeed, just as there are no complications in defining a classical Tanner code on a multigraph, there are no complications in defining a quantum Tanner code using multigraphs; the edge and face sets simply become multisets.

5.5 Application to Strongly Explicit Sum-of-Squares Lower Bounds

In this section, we describe how we use our planted quantum Tanner codes to obtain *strongly* explicit lower bounds against a linear number of levels of the SoS hierarchy, thereby improving

upon the weakly explicit SoS lower bounds of Hopkins and Lin [HL22].

Hopkins and Lin [HL22] show that quantum LDPC codes with small-set boundary and coboundary expansion yield CSPs that are hard for a linear number of levels of the Sum-of-Squares SDP hierarchy. Specifically, the CSPs they use are instances of ℓ-LIN over $\mathbb{F}_2$ (or equivalently, ℓ-XOR) given in Definition 48 below.

Recall that in general, an instance of ℓ-LIN consists of a vector of m variables $y = (y_1, \dots, y_m)$ along with a set of n (affine) linear constraints over $\mathbb{F}_q$, each of which involves $\leq \ell$ variables. Formally, such a system of equations can be expressed in matrix notation as $Ay = \beta$ for some $A \in \mathbb{F}_q^{n \times m}$ and some $\beta \in \mathbb{F}_q^n$, where each row of A has $\leq \ell$ nonzero entries.

Below, we let the *locality* of a CSS code $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ refer to the maximum Hamming weight of any row or column of H_X or H_Z. The qLDPC codes we consider by definition have locality $\ell = O(1)$ as $n \to \infty$.

Definition 48 (ℓ-LIN instances from qLDPC codes [HL22])**.** Let $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ be a CSS code of locality ℓ. Also fix any $\beta \in C_X \setminus C_Z^\perp$. Then define the associated ℓ-LIN instance $\mathcal{I}_{\mathcal{C},\beta}$ to have $m = m_Z$ variables $y_1, \dots, y_m \in \mathbb{F}_q$ and n linear constraints over $\mathbb{F}_q$ given by the system of equations $H_Z^\top y = \beta$, where $y = (y_1, \dots, y_m)$.

[HL22] instantiates this definition with quantum Tanner codes. Although quantum Tanner codes are strongly explicit, meaning that the matrices H_X, H_Z are strongly explicit, any ℓ-LIN instance $\mathcal{I}_{\mathcal{C},\beta}$ from these codes requires a description of some $\beta \in C_X \setminus C_Z^\perp$. Previously, the only known method for finding such a codword was via Gaussian elimination, which runs in $\mathrm{poly}(n)$ time, and thus only yields a (weakly) explicit construction of β and of $\mathcal{I}_{\mathcal{C},\beta}$.

In contrast, our planted quantum Tanner codes in Theorem 32 are guaranteed to have the all-1s vector $\mathbf{1} \in C_X \setminus C_Z^\perp$, which is by definition strongly explicit. As such, we immediately obtain the following.

Lemma 49. *If $\mathcal{C}$ is chosen from a family of planted quantum Tanner codes from Theorem 32 and $\beta = \mathbf{1}$, then $\mathcal{I}_{\mathcal{C},\beta}$ gives a family of strongly explicit ℓ-LIN instances for a constant $\ell = O(1)$.*

Formally, [HL22] obtain their SoS lower bounds by showing the following result, which they applied to quantum Tanner codes. Below, recall that an ℓ-LIN instance is *μ-satisfiable* if there exists an assignment of the variables satisfying $\geq \mu$-fraction of the linear constraints. We refer to [HL22] and the references within for background on the SoS SDP hierarchy.

Theorem 50 ([HL22])**.** *Let $\mathcal{C} = \mathrm{CSS}(C_X = \ker H_X, C_Z = \ker H_Z)$ be a quantum LDPC code of locality ℓ with (c_1, c_2)-small-set boundary and coboundary expansion over a prime-sized alphabet $\mathbb{F}_p$. Then for every $\beta \in C_X \setminus C_Z^\perp$, the ℓ-LIN instance $\mathcal{I}_{\mathcal{C},\beta}$ with $m = m_Z$ variables and n constraints satisfies the following:*

1. *Soundness: $\mathcal{I}_{\mathcal{C},\beta}$ is at most $(1 - c_1)$-satisfiable.*
2. *Completeness: $\mathcal{I}_{\mathcal{C},\beta}$ cannot be refuted by $c_1 c_2 m / 4\ell$ levels of the SoS hierarchy.*

Although Hopkins and Lin [HL22] only showed Theorem 50 for the binary alphabet $\mathbb{F}_2$, their same proof extends to arbitrary fields $\mathbb{F}_p$ for prime p. Specifically, their proof uses small-set (co)boundary expansion to establish a bound on *refutation complexity*, which was then shown to imply an SoS bound for the binary alphabet $\mathbb{F}_2$ by Schoenebeck [Sch08], and for prime-sized alphabets $\mathbb{F}_p$ by Tulsiani [Tul09].

Thus as described above, [HL22] obtained (weakly) explicit, but not strongly explicit, lower bounds against $\Omega(n)$ levels of SoS by taking $\mathcal{C}$ to be a quantum Tanner code in Theorem 50. Meanwhile, applying our planted quantum Tanner codes in Theorem 32 with Lemma 49, we immediately obtain the following corollary to Theorem 50.

Corollary 51 (Strongly explicit SoS lower bounds for ℓ-LIN)**.** *The ℓ-LIN instances $\mathcal{I}_{\mathcal{C},\mathbf{1}}$ for planted quantum Tanner codes $\mathcal{C}$ over any fixed prime-sized alphabet $\mathbb{F}_p$ provide a family of strongly explicit instances with satisfiability $\leq (1-\Omega(1))$, such that no instance can be refuted by cn levels of the SoS hierarchy for a sufficiently small constant $c > 0$.*

[HL22] also showed a reduction that used their ℓ-XOR (i.e. ℓ-LIN over $\mathbb{F}_2$) SoS lower bounds to obtain 3-XOR SoS lower bounds, as stated below. Intuitively, the reduction works by introducing dummy variables to reduce the sizes of constraints.

Proposition 52 (Follows from Claim 6.5 in [HL22])**.** *Let $(\mathcal{I}_n)_{n\to\infty}$ be a strongly explicit family of ℓ-XOR instances such that each $\mathcal{I}_n$:*

1. *has $\Theta(n)$ variables and constraints,*
2. *has satisfiability $\leq (1-\Omega(1))$,*
3. *cannot be refuted by cn levels of the SoS hierarchy for a sufficiently small constant $c > 0$.*

Then there exists a strongly explicit family $(\mathcal{I}'_n)_{n\to\infty}$ of 3-XOR instances that also satisfies the three properties above.

While [HL22] only showed that the reduction behind Proposition 52 preserves weak explicitness, it by definition also preserves strong explicitness, so Proposition 52 holds. As a corollary of Proposition 52 and Corollary 51, we immediately obtain the following.

Corollary 53 (Strongly explicit SoS lower bounds for 3-XOR)**.** *There exists a strongly explicit family $(\mathcal{I}_n)_{n\to\infty}$ of 3-XOR instances such that each $\mathcal{I}_n$ has $\Theta(n)$ variables and constraints, has satisfiability $\leq (1-\Omega(1))$, and cannot be refuted by cn levels of the SoS hierarchy for a sufficiently small constant $c > 0$.*

We suspect a similar reduction should work for ℓ-LIN over arbitrary fields $\mathbb{F}_p$, but for conciseness we will not pursue this direction.

6 Acknowledgments

We thank Max Hopkins for numerous helpful discussions, and for bringing the problem of strongly explicit SoS lower bounds to our attention. We also thank Venkat Guruswami for helping to improve the exposition.

L. Golowich is supported by a National Science Foundation Graduate Research Fellowship under Grant No. DGE 2146752, and in part by V. Guruswami's Simons Investigator award and UC Noyce Initiative Award award. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing.

References

[ABN23] Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. NLTS Hamiltonians from Good Quantum Codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, pages 1090–1096, New York, NY, USA, June 2023. Association for Computing Machinery.

[DDHRZ20] Yotam Dikstein, Irit Dinur, Prahladh Harsha, and Noga Ron-Zewi. Locally testable codes via high-dimensional expanders. *arXiv:2005.01045 [cs]*, May 2020. arXiv: 2005.01045.

[DFHT21] Irit Dinur, Yuval Filmus, Prahladh Harsha, and Madhur Tulsiani. Explicit SoS Lower Bounds from High-Dimensional Expanders. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 38:1–38:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISSN: 1868-8969.

[DHLV23] Irit Dinur, Min-Hsiu Hsieh, Ting-Chun Lin, and Thomas Vidick. Good Quantum LDPC Codes with Linear Time Decoders. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, pages 905–918, New York, NY, USA, June 2023. Association for Computing Machinery.

[DLZ23] Irit Dinur, Siqi Liu, and Rachel Zhang. New Codes on High Dimensional Expanders, August 2023. ISSN: 1433-8092.

[Ebe13] Sean Eberhard. Answer to "Idea behind the factorization of the matrix $\operatorname{diag}(a,a^{-1})$ in algebraic K-Theory", April 2013.

[EH17] Lior Eldar and Aram W. Harrow. Local Hamiltonians Whose Ground States are Hard to Approximate. *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 427–438, October 2017. arXiv: 1510.02082.

[FKP19] Noah Fleming, Pravesh Kothari, and Toniann Pitassi. Semialgebraic Proofs and Efficient Algorithm Design. *Foundations and Trends® in Theoretical Computer Science*, 14(1-2):1–221, December 2019. Publisher: Now Publishers, Inc.

[GRS22] Venkatesan Guruswami, Atri Rudra, and Madhu Sudan. Essential coding theory. *Draft available at http://www. cse. buffalo. edu/ atri/courses/coding-theory/book*, 2022.

[HL22] Max Hopkins and Ting-Chun Lin. Explicit Lower Bounds Against Omega(n)-Rounds of Sum-of-Squares. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 662–673. IEEE Computer Society, October 2022.

[Hop23] Max Hopkins. Personal Communication, 2023.

[HT03] Holger F. Hofmann and Shigeki Takeuchi. Violation of local uncertainty relations as a signature of entanglement. *Physical Review A*, 68(3):032103, September 2003. Publisher: American Physical Society.

[JMRW22] Fernando Granha Jeronimo, Tushant Mittal, Sourya Roy, and Avi Wigderson. Almost Ramanujan Expanders from Arbitrary Expanders via Operator Amplification. pages 378–388. IEEE Computer Society, October 2022.

[Jul13] Julien. Answer to "Why is the quotient map $SL_n(\mathbb{Z})$ to $SL_n(\mathbb{Z}/p\mathbb Z)$ is surjective?", March 2013.

[KP23] Gleb Kalachev and Pavel Panteleev. Two-sided Robustly Testable Codes, August 2023. arXiv:2206.09973 [cs, math].

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, September 1988.

[LW93] A. Lubotzky and B. Weiss. Groups and Expanders. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 10:95–109, 1993.

[LZ22] Anthony Leverrier and Gilles Zémor. Quantum Tanner codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 872–883. IEEE Computer Society, October 2022.

[LZ23a] Anthony Leverrier and Gilles Zémor. Decoding Quantum Tanner Codes. *IEEE Transactions on Information Theory*, 69(8):5100–5115, August 2023. Conference Name: IEEE Transactions on Information Theory.

[LZ23b] Anthony Leverrier and Gilles Zémor. Efficient decoding up to a constant fraction of the code length for asymptotically good quantum codes. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, Proceedings, pages 1216–1244. Society for Industrial and Applied Mathematics, January 2023.

[Mor94] M. Morgenstern. Existence and Explicit Constructions of q + 1 Regular Ramanujan Graphs for Every Prime Power q. *Journal of combinatorial theory. Series B*, 62(1):44–62, 1994. Place: SAN DIEGO Publisher: Elsevier Inc.

[Nir23] Chinmay Nirkhe. Making the Leap to Quantum PCPs, July 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, pages 375–388, New York, NY, USA, June 2022. Association for Computing Machinery.

[Sch08] Grant Schoenebeck. Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602, October 2008. ISSN: 0272-5428.

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the lasserre hierarchy. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, STOC '09, pages 303–312, New York, NY, USA, May 2009. Association for Computing Machinery.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il