Metric Self-Dual Completion and Optimal Additive Hardness for Quantum and Graph-State Distance*

Rafail Ostrovsky
UCLA

Abstract

We prove that the quantum code distance is NP-hard to approximate within an additive error of cN, for some constant $c > 0$, where N is the number of qubits. Our reductions are deterministic. This improves the previous square-root additive gap to $\Omega(N)$ and resolves the explicitly stated linear-gap question of Kapshikar and Kundu. Our result holds for CSS codes with identical X- and Z-check spaces, and with a constant rate and constant relative distance. For every fixed $\lambda > 1$, there is a constant $c > 0$ such that hardness still holds even when every nonidentity stabilizer has weight greater than λ times the quantum distance. We also improve the hardness gap of graph state distance on N vertices of Grigorescu, Jha, and Samperton from cube-root to $\Omega(N)$, resolving their explicitly stated open question. Both hardness results are asymptotically optimal since both distances are at most N. Our graph state distance hardness result holds for balanced bipartite graphs with a binary adjacency matrix that is its own inverse (mod 2).

Our main technique for both hardness bounds above is classical: we show how to convert any code C of length m into a self-dual code $A(C)$ of length $N = \Theta(m)$ while exactly doubling the original coset metric. The conversion is deterministic and efficient. We call it the *metric self-dual completion* of C. It comes with a linear embedding $\tau : \mathbb{F}_2^m \hookrightarrow \mathbb{F}_2^N$. The embedding doubles all Hamming distances between vectors in $\mathbb{F}_2^m$ and all pairwise distances between corresponding cosets. The embedding also guarantees that all codewords of $A(C)$ of weight at most $2m$ are exactly $\tau(C)$.

*Posted to arXiV on September 19^{th}, 2026 [Ost26].

Contents

1 Introduction **1**
- 1.1 Metric Self-Dual Completion . . . 1
- 1.2 Optimal additive hardness for quantum and graph-state distance . . . 2
- 1.3 Classical consequences and related work . . . 3

2 Preliminaries **5**
- 2.1 Notation and conventions . . . 5
- 2.2 A linear algebra view of quantum codes . . . 5
- 2.3 A linear algebra view of graph distance . . . 6

3 Metric self-dual completion **6**
- 3.1 Main construction: building metric self-dual . . . 6
- 3.2 Preservation of all source-coset distances . . . 8

4 Hardness applications **8**
- 4.1 Application to Quantum code distance . . . 9
- 4.2 Application to Graph-state distance . . . 9

5 Linear-size self-dual hosts **10**
- 5.1 Putting the all-ones word into any dual-good code. . . . 10
- 5.2 The prefix-sum construction. . . . 11
- 5.3 Constructing linear-size self-dual codes . . . 12

6 AI methodologies **12**

References **13**

A Type I and Type II metric self-dual completions **15**
- A.1 The completion and its recovery map . . . 15
- A.2 Type I completion . . . 16
- A.3 Type II completion . . . 16

1 Introduction

Recent work established $N^{1/2}$ additive hardness gap for quantum-code distance [KK23] and $N^{1/3}$ hardness gap for graph-state distance [GJS25]. These works left open the tantalizing question whether either problem admits a linear additive hardness gap. We answer both questions in the affirmative.

We stress that our main construction is classical. We convert any nonzero binary linear code into a self-dual code with constant length expansion while exactly doubling the original minimum distance. While simple repetition makes any code self-orthogonal, observe that the self-dual extension $\{(x, x) : x \in \mathbb{F}_2^m\}$ has a distance of two and is therefore useless.

Our main idea is to use a self-dual binary "host" code whose distance exceeds twice the maximum embedded weight. We only keep host words that are orthogonal to an embedded copy of the input code, then add that copy. The key argument is to show that the dimensions removed and later added are exactly equal. The modified code is self-orthogonal and therefore, having dimension $N/2$, is self-dual. Since the host distances are large, this guarantees that all corresponding coset distances of the original input code are exactly doubled.

The retained host subspace supplies CSS checks. We use the completed self-dual code for the already known graph reduction. All classical codes considered in this paper are binary linear codes.

1.1 Metric Self-Dual Completion

As usual, $d(C)$ denotes the minimum distance of C. Let $\mathrm{dist}(y, C)$ be the Hamming distance from y to C. Our main result (clauses (i)-(iv)) applies to every code $C \leqslant \mathbb{F}_2^m$. For fixed m and H, the same output length and embedding τ work for every C. To simplify the notation, we refer to the self-dual completion of C as A, instead of $A(C)$.

Main Theorem 1 (Metric self-dual completion). *$\forall\ m \geqslant 1$ and $H \geqslant 2m$ $\exists\ N = \Theta(H)$ and linear embedding $\tau : \mathbb{F}_2^m \hookrightarrow \mathbb{F}_2^N$, such that $\forall\ C \leqslant \mathbb{F}_2^m$ $\exists$ self-dual code $A = A^\perp \leqslant \mathbb{F}_2^N$ so that:*

(i) $\forall x \in \mathbb{F}_2^m,\ 2 \cdot \mathrm{wt}(x) = \mathrm{wt}(\tau(x))$;

(ii) $\forall x \in \mathbb{F}_2^m,\ \mathrm{dist}(\tau(x), A) = 2 \cdot \mathrm{dist}(x, C)$;

(iii) $\tau(C) = \{a \in A : \mathrm{wt}(a) \leqslant H\}$;

(iv) A generator matrix for A and matrix for τ can be computed in deterministic $\mathrm{poly}(N)$ *time.*

Furthermore, clause (ii) implies that the induced map $y + C \mapsto \tau(y) + A$ is well-defined and injective. Specifically, this correspondence doubles every pairwise coset distance.

Together with the self-dual code A, our construction produces a self-orthogonal subspace $K \leqslant A$ where $\dim K = N/2 - \dim C$ and $d(K) > 2H$. For $C \neq 0$, K satisfies:

$$\min_{v \in K^\perp \setminus K} \mathrm{wt}(v) = 2d(C)$$

By setting $H = 2m$, we get a self-dual code A of linear length (in the length of the original codeword) and, by clause (iv), do it in deterministic poly-time. All codewords of A of weight at most $2m$ are exactly the embedded codewords of $\tau(C)$ by clause (iii). That means that $d(A) = 2d(C)$ whenever $C \neq 0$, by clauses (i) and (iii). Thus, we obtain self-duality (with only a constant-factor blow-up) in the code length, while preserving the original coset geometry. Our main (classical) construction and its proof appear in Section 3.

We remark that the full source-coset metric isomorphism is a stronger guarantee than these applications require. Generalizations of clauses (i)-(iv) of our main theorem appear in the appendix.

1.2 Optimal additive hardness for quantum and graph-state distance

Recall that a distance problem has a *linear additive gap* if, for some constant $c > 0$, it is NP-hard to distinguish the YES case $f \leqslant T$ from the NO case $f > T + cN$. Such a gap implies NP-hardness of approximation within additive error $c_0 N$ for some constant $c_0 > 0$. Here f is the distance, N is the number of qubits or vertices, and the integer threshold T is part of the input. Inputs come from a "promise problem" and are guaranteed to satisfy one of the two conditions. We use many-one deterministic Karp reductions to prove NP-hardness.

Distance of a quantum code. Quantum distance is the smallest number of qubits on which an undetectable error can change the encoded information. We prove hardness for Calderbank-Shor-Steane (CSS) codes with identical X- and Z-check spaces. For such a code $Q = CSS(B, B)$, write N for the number of qubits, k for the number of encoded (logical) qubits, and $d(Q)$ for the logical distance of the quantum code. We write $d(B)$ for the minimum Hamming weight of a nonzero vector in B; equivalently, this is the minimum weight of a nonzero stabilizer label. See Section 2.2 for a purely linear-algebra formulation of these notions.

For a fixed constant $c > 0$, the additive-gap quantum-distance promise problem takes as input an N-qubit code $Q = \mathrm{CSS}(B, B)$ and an integer threshold T, with the promise that either $d(Q) \leqslant T$ or $d(Q) > T + cN$. The challenge is to determine which of these two conditions holds.

Theorem 2 (Optimal additive hardness for quantum distance). *For every fixed $\lambda > 1$, there are positive constants c, r, δ with the following property. The promise problem above is NP-hard even when Q has at least rN logical qubits, has distance at least δN, and every nontrivial stabilizer has weight greater than $\lambda \cdot d(Q)$.*

Thus, even for codes that encode a linear number of qubits and correct a linear number of errors, estimating the smallest fraction of qubits supporting a nontrivial logical Pauli is NP-hard to some fixed constant additive accuracy. We stress that the construction does not guarantee sparse check matrices. Thus, the above theorem does not establish hardness for quantum LDPC codes.

Optimal additive hardness for quantum distance has received some attention in the recent literature. Specifically, Kapshikar and Kundu [KK23, Corollary 2 and Section V] established a randomized reduction square-root additive hardness gap and asked whether a linear gap could be obtained. Grigorescu, Jha, and Samperton [GJS25, Theorem 1 and Sections 1.1-1.2] provided an alternative proof of the square-root hardness under deterministic Karp reductions and posed the linear-gap question as a main motivation.

Grigorescu, Jha, and Samperton considered hypergraph-product reduction with full-rank factor checks for this problem, but showed that their approach cannot yield a linear additive gap under deterministic Karp reductions unless P=NP [GJS25, Proposition 22]. Our construction uses a self-dual host and a subspace exchange to obtain linear output length. Our Theorem 2 resolves the linear-gap question posed in both papers above, going all the way from $N^{1/2}$ to $\Omega(N)$ with a deterministic Karp reduction.

Graph-state Distance. We use the graph-state distance definition of Kovalev, Dumer, and Pryadko [KDP11]. For a simple graph G, with adjacency matrix Γ, its graph-state distance $d(\Gamma)$ is the minimum weight of a nontrivial element of the stabilizer of the graph state associated with G. We provide a linear-algebraic definition of this graph-state distance in Section 2.3. Grigorescu, Jha, and Samperton [GJS25, Theorem 2 and Section 5.3] established an $N^{1/3}$ additive gap for this problem and asked whether it could be made linear. Our Theorem 3 resolves their question. Furthermore, it

answers their question even when G is bipartite and balanced, and when its adjacency matrix is an involution (over $\mathbb{F}_2$):

Theorem 3 (Optimal additive hardness for graph-state distance). *$\exists\, c, \delta > 0$ s.t. Graph-state minimum distance is NP-hard to approximate within an additive error of $c \cdot N$ even when G is bipartite and balanced and its adjacency matrix Γ satisfies $\Gamma^2 = I$ and $d(\Gamma) \geqslant \delta N$.*

We start with Grigorescu, Jha, and Samperton's code-to-graph mapping [GJS25, proof of Theorem 2]. We then use our metric self-dual completion as the missing piece: our construction doubles the source distance and produces a self-dual input. Self-duality forces the graph to be balanced and gives $\Gamma^2 = I$ as needed.

problem	previous additive gap	this paper
CSS quantum distance	$N^{1/2}$	$\Omega(N)$
graph-state distance	$N^{1/3}$	$\Omega(N)$

Since distances are at most N, our additive gaps are asymptotically optimal for both problems.

1.3 Classical consequences and related work

We first explore the implications of clauses (i), (iii), and (iv) of our main theorem (of metric self-dual completion). We start with the result of Dumer, Micciancio, and Sudan [DMS03] who showed a linear additive (hardness) gap for the minimum-distance code, using a randomized reduction. For multiplicative approximation, Cheng and Wan [CW12] proved hardness within every fixed constant factor under deterministic reductions. The same result was later shown using a PCP-free proof by Bhattiprolu, Guruswami, Lee, and Ren [BGLR25]. For codes of constant rate and relative distance, Austrin and Khot [AK14] obtained a deterministic constant-factor hardness.

We stress that none of these reductions guarantee self-duality of the output codes, as explicitly discussed in Kapshikar and Kundu (in connection to quantum applications, see [KK23, Section I]). Our main result, specifically clauses (i), (iii), and (iv), allows us to adopt these classical hardness results to self-dual codes.

Austrin and Khot [AK14, Theorem 1.1] showed that, for some constants $\varepsilon > 0$ and $\gamma > 1$, it is NP-hard under a deterministic reduction to distinguish $d(C) \leqslant t$ from $d(C) > \gamma t$, where $t > 0$ is an integer and $C \leqslant \mathbb{F}_2^m$ satisfies $\dim C \geqslant \varepsilon m$ and $d(C) \geqslant \varepsilon m$. Combining this with clauses (i), (iii), and (iv) of our main theorem gives:

Corollary 4 (Optimal additive hardness for self-dual distance). *$\exists\, c, \delta > 0$ such that, given a self-dual $A \leqslant \mathbb{F}_2^N$ and threshold T, it is NP-hard (under Karp reductions) to distinguish $d(A) \leqslant T$ vs. $d(A) > T + cN$, even when $d(A) \geqslant \delta N$.*

Proof. Use Austrin-Khot hard instance (C, t) as input to Theorem 1, using clauses (i), (iii), and (iv), with $H = 2m$. For some fixed constant $\alpha > 0$ observe that $d(A) = 2d(C)$ and $N \leqslant \alpha m$. Assign $T = 2t$, $\delta = 2\varepsilon/\alpha$, and $c = \delta(1 - 1/\gamma)$. We therefore have $d(A) \geqslant \delta N$. The YES case gives $d(A) \leqslant T$. The NO case gives $d(A) > \gamma T$, thus:

$$d(A) - T > \left(1 - \frac{1}{\gamma}\right) d(A) \geqslant cN \qquad \square$$

The hard instances of Corollary 4 have a rate of $1/2$ and a relative distance of at least δ. Next, observe that clauses (i) and (iii) of Theorem 1 give $d(A) = 2d(C)$, which preserves every multiplicative approximation gap when the threshold is also doubled. Hence, when we apply the PCP-free hardness result [BGLR25, Theorem 1.3], this gives the following:

Corollary 5 (Hardness within every constant factor for self-dual distance). *For every fixed constant $\gamma > 1$, given a self-dual $A \leqslant \mathbb{F}_2^N$ and a threshold T, it is hard (under deterministic Karp reductions) to distinguish $d(A) \leqslant T$ from $d(A) > \gamma T$. This hardness proof is PCP-free.*

Proof. Choose $\gamma' > \gamma$. We reduce a nonzero instance (C, t) of Bhattiprolu, Guruswami, Lee, and Ren, with $C \leqslant \mathbb{F}_2^m$, $t > 0$, and hardness factor γ', to an instance (A, T) of the problem in the corollary. We apply our completion Theorem 1, using clauses (i), (iii), and (iv), with $H = 2m$ to efficiently construct a self-dual A where $d(A) = 2d(C)$ and $T = 2t$. In the YES case, $d(A) \leqslant T$. In the NO case, $d(A) \geqslant \gamma' T > \gamma T$. Finally, our reduction, together with Bhattiprolu, Guruswami, Lee, and Ren's proof, remains efficient and PCP-free. □

Remark: The above corollaries rely on preserving distance, whereas a different direction minimizes the number of coordinates added to achieve self-orthogonality. An, Kaplan, Kim, Luo, and Wang [AKK+25] study minimizing the number of columns that must be added to a generator matrix to make its code self-orthogonal while retaining its input dimension. Their embeddings (of Hamming codes of length $2^r - 1$ with $r \geqslant 3$) are self-dual. However, their result does not guarantee self-duality for arbitrary input codes and does not preserve the source distance gap. In contrast, our metric completion allows the dimension to grow and guarantees self-duality.

Type I and Type II generalizations. Appendix A extends clauses (i)-(iv) of our main theorem to Type I and Type II completions with distance scaling factors of two and four, respectively. These extensions also provide linear recovery maps. We stress that these generalizations are not used in the main body of the paper.

Concurrent work. In a preprint posted after ours, Kapshikar [Kap26] also obtains a linear additive gap for quantum code distance under randomized reductions and deterministic truth-table hardness under a circuit-hardness hypothesis. Our NP-hardness is unconditional and holds under deterministic Karp reductions.

Organization. Section 2 sets up the notation and presents self-contained linear-algebra descriptions of quantum codes and graph distance. We also present a code-to-graph reduction that we use in the graph-state application in a later section. Section 3 proves clauses (i)-(iv) of our main theorem: it explains the construction and proof of the metric self-dual completion theorem and the companion guarantee for K subspace, which will be helpful in applications. Section 4 gives the hardness results for both quantum distance and graph-distance applications of clauses (i), (iii), and (iv) of our main theorem. Section 5 builds the host construction, which is one of the input ingredients needed for clauses (ii)-(iv) of our main theorem. Appendix A extends clauses (i)-(iv) of our main theorem to Type I and Type II codes with a linear recovery map.

2 Preliminaries

2.1 Notation and conventions

All lengths, dimensions, indices, repetition parameters, weight thresholds, and Hamming-ball radii are integers. Vector spaces are over $\mathbb{F}_2$. $\mathrm{wt}(x)$ is the Hamming weight, and $\mathrm{supp}(x)$ is the support of x. $U \leqslant V$ means that U is a linear subspace of V. For $U, W \leqslant V$ when $U \cap W = \{0\}$, define *direct sum* as $U \oplus W \stackrel{\text{def}}{=} \{u + w : u \in U,\ w \in W\}$. $C^\perp$ is the orthogonal complement of C. A code C is *self-orthogonal* if $C \leqslant C^\perp$ and *self-dual* if $C = C^\perp$. Recall that the distance of a code is $d(C) = \min_{0 \neq c \in C} \mathrm{wt}(c)$, where $d(0) = +\infty$. As usual, $\mathrm{dist}(y, C) = \min_{c \in C} \mathrm{wt}(y + c)$. The quotient metric is $\mathrm{dist}(x + C, y + C) \stackrel{\text{def}}{=} \mathrm{dist}(x + y, C)$. Self-dual codes are *Type II* if all weights are divisible by four; they are *Type I* otherwise, and, of course, of even length. We will use

$$\mathrm{wt}(x + y) = \mathrm{wt}(x) + \mathrm{wt}(y) - 2|\,\mathrm{supp}(x) \cap \mathrm{supp}(y)| \tag{1}$$

Codes are specified by bases. Generator matrices have basis rows. Vectors are columns. Matrices also represent linear maps. All our algorithms are deterministic.

2.2 A linear algebra view of quantum codes

For $V = \mathbb{F}_2^N \times \mathbb{F}_2^N$, *symplectic weight* is defined as $\mathrm{swt}(x, z) \stackrel{\text{def}}{=} |\,\mathrm{supp}(x) \cup \mathrm{supp}(z)|$ where $x, z \in \mathbb{F}_2^N$. The *symplectic pairing* (also called *symplectic form*) is defined as $\omega((x, z), (x', z')) \stackrel{\text{def}}{=} x^\top z' + z^\top x'$. Observe that it is nondegenerate and bilinear. For $S \leqslant V$ *symplectic orthogonal complement* is defined as $S^{\perp_s} \stackrel{\text{def}}{=} \{v \in V : \forall s \in S,\ \omega(v, s) = 0\}$. S is *isotropic* if $S \leqslant S^{\perp_s}$. Nondegeneracy gives $\dim S^{\perp_s} = 2N - \dim S$. An isotropic subspace S is called a *stabilizer space* [KK23, Sections I and II-B1] and jointly with N describes, for distance calculations, a *stabilizer code* Q with parameter $k = N - \dim S$ (where k is the number of qubits that Q encodes). When $k > 0$ the *logical distance* of Q is defined as:

$$d(Q) \stackrel{\text{def}}{=} \min_{v \in S^{\perp_s} \setminus S} \mathrm{swt}(v) \tag{2}$$

When $K \leqslant \mathbb{F}_2^N$ is self-orthogonal, observe that $S = K \times K$ is isotropic. The quantum code is denoted as $Q = \mathrm{CSS}(K, K)$, named after Calderbank and Shor [CS96] and Steane [Ste96]. Observe that $S^{\perp_s} = K^\perp \times K^\perp$, $k = N - 2\dim K$, and, for $k > 0$,

$$d(Q) = \min_{u \in K^\perp \setminus K} \mathrm{wt}(u) \tag{3}$$

Since $S^{\perp_s} = K^\perp \times K^\perp$ and $S = K \times K$, every $(x, z) \in S^{\perp_s} \setminus S$ has at least one component in $K^\perp \setminus K$. Thus, for $(x, z) \in S^{\perp_s} \setminus S$, $\mathrm{swt}(x, z) \geqslant \max\{\mathrm{wt}(x), \mathrm{wt}(z)\} \geqslant \min_{u \in K^\perp \setminus K} \mathrm{wt}(u)$, with equality at $(u, 0)$ for a minimum-weight $u \in K^\perp \setminus K$.

The minimum symplectic weight of a nonzero vector in $S = K \times K$, called the *minimum nonzero stabilizer weight*, is defined as:

$$d(K) = \min_{0 \neq s \in S} \mathrm{swt}(s) \tag{4}$$

with the minimum equal to $+\infty$ when $K = 0$.

2.3 A linear algebra view of graph distance

Let G be a simple unweighted graph on N vertices. We denote by Γ its adjacency matrix. The notion of *graph-state stabilizer space* is defined as $S \stackrel{\text{def}}{=} \{(x, \Gamma x) : x \in \mathbb{F}_2^N\}$. Since G is a simple, $\Gamma = \Gamma^\top$, which makes S isotropic, with $\dim S = N$. Thus, $S = S^{\perp_s}$. That means that $k = 0$ (hence, no qubits are encoded). The distance of G is defined as the smallest symplectic weight of a nonzero stabilizer label [GJS25, Section 1.2]:

$$d(\Gamma) \stackrel{\text{def}}{=} \min_{x \in \mathbb{F}_2^N \setminus \{0\}} |\operatorname{supp}(x) \cup \operatorname{supp}(\Gamma x)| \tag{5}$$

If x is a selector function for a vertex set X, then $\operatorname{supp}(\Gamma x)$ consists of vertices with an odd number of neighbors in X, giving an equivalent combinatorial definition.

Graph distance minimizes over $S \setminus \{0\}$, whereas logical distance in equation (2) minimizes over $S^{\perp_s} \setminus S$ for a code Q with $k > 0$, i.e., encoding at least one qubit.

Lemma 6 (Grigorescu, Jha, and Samperton, [GJS25, proof of Theorem 2]). *Let $A \leqslant \mathbb{F}_2^N$ with dimension $0 < r < N$. After a coordinate permutation, take the parity-check matrix of A and put it into systematic form $[I_{N-r} \mid R]$. A simple bipartite graph with adjacency matrix*

$$\Gamma = \begin{pmatrix} 0 & R \\ R^\top & 0 \end{pmatrix}$$

has N vertices, and $d(\Gamma) = \min\{d(A), d(A^\perp)\}$.

Furthermore, GJS's graph-state distance construction shows that a graph stabilizer indexed by (u, v) has support equal to the union of the supports of $(u, R^\top u) \in A^\perp$ and $(Rv, v) \in A$. Therefore, the minimum nonzero stabilizer weight is $\min\{d(A), d(A^\perp)\}$, which we get by setting one index to zero.

If A is self-dual, this graph is balanced and has $d(\Gamma) = d(A)$. Moreover, R is square, and self-duality gives $RR^\top = R^\top R = I$, so $\Gamma^2 = I$.

3 Metric self-dual completion

In this section, we prove clauses (i)-(iv) of our main result and its companion guarantees. The only ingredient whose proof we delay to Section 5 is the following proposition:

Proposition 7 (Linear-size self-dual code). *For every $\Delta \geqslant 1$, there is a deterministically and efficiently constructible self-dual $P \leqslant \mathbb{F}_2^N$ such that $4 \mid N$, $d(P) \geqslant \Delta$, and $N = \Theta(\Delta)$.*

Section 5 proves this proposition starting from the dual-good construction in Shpilka's paper, whose Appendix A is by Guruswami [Shp09]. In that section, we show how to go from a dual-good code to a self-dual code without losing constant relative rate and constant relative distance. We now proceed with our main construction.

3.1 Main construction: building metric self-dual

We start with an arbitrary input length m and a threshold $H \in \mathbb{Z}$ with $H \geqslant 2m$. We use Proposition 7 with $\Delta = 2H + 1$ to choose a self-dual *host* $M \leqslant \mathbb{F}_2^N$ with $d(M) > 2H$ and $N = \Theta(H)$. The host and its length depend only on H. $N \geqslant d(M) > 2H \geqslant 2m$. We define $\tau : \mathbb{F}_2^m \hookrightarrow \mathbb{F}_2^N$ as follows:

$$\tau(x) \stackrel{\text{def}}{=} (x, x, 0^{N-2m})$$

It is immediate that the image $W = \tau(\mathbb{F}_2^m)$ is self-orthogonal (since $\tau(x)^\top \tau(y) = x^\top y + x^\top y = 0$ for all $x, y \in \mathbb{F}_2^m$). Moreover, for every $x \in \mathbb{F}_2^m$,

$$\mathrm{wt}(\tau(x)) = 2\,\mathrm{wt}(x) \tag{6}$$

Therefore, every word of W has weight at most $2m \leqslant H$. Since $d(M) > 2H$ and $\mathrm{wt}(\tau(x)) \leqslant H$, the triangle inequality gives, $\forall\ u \in M \setminus \{0\}$, and $\forall\ x \in \mathbb{F}_2^m$,

$$\mathrm{wt}(u + \tau(x)) > H \tag{7}$$

We call the host M *protected through* H when equation (7) holds. Furthermore, observe that $M \cap W = 0$.

As a second step, starting with any $C \leqslant \mathbb{F}_2^m$, we want to insert $\tau(C)$ into host M while preserving self-duality. Since we need to keep the subspace self-orthogonal, we will keep only host words that are orthogonal to $\tau(C)$. So define

$$K \stackrel{\text{def}}{=} M \cap \tau(C)^\perp$$

As a third step, we add $\dim(\tau(C))$ dimensions to K to obtain:

$$A \stackrel{\text{def}}{=} K + \tau(C)$$

K is self-orthogonal since M is, and $K \leqslant M$. Furthermore, $\tau(C)$ is self-orthogonal by our construction. K and $\tau(C)$ are orthogonal to each other since $K \leqslant \tau(C)^\perp$. Since $K \cap \tau(C) \subseteq M \cap W = \{0\}$, we have $A = K \oplus \tau(C)$. Therefore, A is self-orthogonal.

To prove self-duality, the only missing step is to show that when we go from M to K, the operation removes from M exactly $\dim C$ dimensions to get K. When we build A we add $\tau(C)$ dimensions back to K and therefore A has dimension $N/2$. The same calculation also identifies the short vectors in $K^\perp \setminus K$, which will later determine the quantum distance.

Unique decomposition. Using the fact that $K \leqslant M$, $M \cap W = 0$, and $\tau(C) \leqslant W$, we obtain the following equations:

$$A = K \oplus \tau(C) \tag{8}$$

and

$$A + W = K \oplus W \tag{9}$$

Therefore, every word of $A + W$ can be written as $u + \tau(x)$, with $u \in K$ and $x \in \mathbb{F}_2^m$.

Observe that M, length $N = \Theta(H)$, and embedding τ only depends on m, H, so the same choices work for any code of length m. Furthermore, the host M is computable in deterministic poly-time in N as shown in Proposition 7. The rest is just Gaussian elimination and direct-sum decomposition. Therefore, we can compute in $\mathrm{poly}(N)$ bases for $W, K, A, A + W$.

Lemma 8 (Self-dual completion). *For $K = M \cap \tau(C)^\perp$ and $A = K + \tau(C)$ constructed above, $\dim K = N/2 - \dim C$, and A is self-dual. Moreover, K is self-orthogonal, $d(K) > 2H$, and*

$$\{v \in K^\perp \setminus K : \mathrm{wt}(v) \leqslant H\} = \tau(C \setminus \{0\}) \tag{10}$$

Consequently, if $C \neq 0$, then

$$\min_{v \in K^\perp \setminus K} \mathrm{wt}(v) = 2d(C) \tag{11}$$

Proof. We have already established in Section 3.1 that A is self-orthogonal and $A = K \oplus \tau(C)$. Since τ is injective, we have $\dim C = \dim \tau(C)$ and therefore we only need to show that $\dim K = N/2 - \dim C$. Self-duality of M gives us

$$K = M^\perp \cap \tau(C)^\perp = (M + \tau(C))^\perp$$

But we also know that $M \cap W = 0$ and $\tau(C) \leqslant W$. Therefore $M + \tau(C)$ is direct. That is, $K = (M \oplus \tau(C))^\perp$. Therefore

$$K^\perp = M \oplus \tau(C) \tag{12}$$

Since $\dim K = N - \dim K^\perp$, it follows that $\dim K = N - \dim(M + \tau(C)) = N/2 - \dim C$. By equation (8), $\dim A = \dim K + \dim C = N/2$. Finally, since A is self-orthogonal, $A = A^\perp$.

Since $K \leqslant M$, is self-orthogonal and $d(K) \geqslant d(M) > 2H$, by equation (12), every word of $K^\perp$ is uniquely defined as $u + \tau(c)$ with $u \in M$ and $c \in C$. If $u \neq 0$, its weight exceeds H by equation (7). If $u = 0$, its weight is $2\,\mathrm{wt}(c) \leqslant H$, and it lies outside K exactly when $c \neq 0$, because $K \cap \tau(C) = 0$. This proves equation (10). For $C \neq 0$, these short words have minimum weight $2d(C)$, which gives equation (11). □

3.2 Preservation of all source-coset distances

For any $y \in \mathbb{F}_2^m$, let's consider the coset $\tau(y) + A$, and its relation to the original coset $y + C$. Since we have already shown that $A = K \oplus \tau(C)$, every word of $\tau(y) + A$ has a unique representation that can be written as $u + \tau(x)$, with $u \in K$ and $x \in y + C$. By the argument already presented in the proof of Lemma 8, every minimum-weight representative of $\tau(y) + A$ has zero host component $u = 0$ and weight exactly $2\,\mathrm{dist}(y, C)$. This proves

$$\mathrm{dist}(\tau(y), A) = 2\,\mathrm{dist}(y, C) \tag{13}$$

The same argument gives

$$\{p \in \tau(y) + A : \mathrm{wt}(p) \leqslant H\} = \tau(y + C) \tag{14}$$

The direct sum equations in equations (8) and (9) show that τ induce an isomorphism $\mathbb{F}_2^m/C \to (A + W)/A$. Applying this to $y + z$ shows that this isomorphism exactly doubles the distance between any two source cosets $y + C$ and $z + C$.

Proof of Theorem 1. The construction in Section 3.1 picks M, N, τ independently of C, with $d(M) > 2H$ and $N = \Theta(H)$. By Lemma 8, we know that $A = A^\perp$. Now, equations (6), (13) and (14) give clauses (i)-(iii) of the main theorem, where we set $y = 0$ in equation (14). The coset correspondence and our pairwise distance guarantees are already proved in Section 3.2. For $C \neq \{0\}$, equations (6) and (14) (with y set to 0) give

$$d(A) = 2d(C) \leqslant H \tag{15}$$

For $C = \{0\}$, our construction already sets $A = M$, so $d(A) > 2H$. The running time in clause (iv) is analyzed in Section 3.1. □

4 Hardness applications

Our two additive hardness gaps use different outputs of our main construction. For quantum minimum distance, we use K for CSS check spaces. We use a self-dual code A as the input to the code-to-graph reduction of Grigorescu, Jha, and Samperton [GJS25, proof of Theorem 2]. Both hardness constructions exactly double the classical source distance and retain the linear output length. The Austrin-Khot hardness theorem supplies us with the hardness gap to get a linear additive gap for both settings.

4.1 Application to Quantum code distance

Lemma 9 (Identical-check Minimum Distance). *For all nonzero $C \leqslant \mathbb{F}_2^m$ and $H \geqslant 2m$, in poly$(m+H)$ time we can construct $K \leqslant \mathbb{F}_2^N$ such that K is self-orthogonal and $Q = \mathrm{CSS}(K,K)$ where $N = \Theta(H)$, $k = 2\dim C$, $d(Q) = 2d(C)$, and $d(K) > 2H$.*

Proof. Start with K from Lemma 8 and let $Q = \mathrm{CSS}(K,K)$. Our construction guarantees that K is self-orthogonal, $d(K) > 2H$, and $\dim K = N/2 - \dim C$. Section 2.2 shows that $k = N - 2\dim K$ and $k = 2\dim C$. Therefore, using equations (3) and (11), gives us $d(Q) = 2d(C)$. The reduction is deterministic and efficient Section 3.1. □

Proof of Theorem 2. Pick a constant $\lambda > 1$ and start with an Austrin-Khot hard instance (C,t), where $C \leqslant \mathbb{F}_2^m$, from Section 1.3. They show that, for some constant $\gamma > 1$, it is NP-hard to distinguish $d(C) \leqslant t$ from $d(C) > \gamma t$, even when the dimension and distance of C are linear in m (i.e., $\dim C = \Omega(m)$ and $d(C) = \Omega(m)$).

We now apply Lemma 9 to C with $H = 2\lceil\lambda\rceil m$. Observe that $Q = \mathrm{CSS}(K,K)$ acts on $N = \Theta(m)$ qubits. Moreover, $k = 2 \cdot \dim C = \Omega(N)$ and $d(Q) = 2 \cdot d(C) = \Omega(N)$.

Set $T = 2t$. By the argument already presented in the proof of Corollary 4, the same linear additive gap follows. So we can find $c, r, \delta > 0$ so that $k \geqslant rN$ and $d(Q) \geqslant \delta N$, with an additive hardness gap of cN.

Moreover, every nontrivial stabilizer has weight $\geqslant d(K)$. Since $C \leqslant \mathbb{F}_2^m$, its distance is at most m. Since we set $H = 2\lceil\lambda\rceil m$ and have $d(Q) = 2d(C)$, this guarantees that $d(K) > H \geqslant \lambda \cdot d(Q)$. Construction of Q is poly(N)-time and deterministic. □

4.2 Application to Graph-state distance

By Lemma 6, a self-dual code and its graph have the same distance. Combining this fact with our completion gives the following:

Theorem 10 (Exact reduction to graph-state distance). *For nonzero $C \leqslant \mathbb{F}_2^m$, one can construct in poly(m) time a simple balanced bipartite graph on $N = \Theta(m)$ vertices with adjacency matrix Γ that satisfies $\Gamma^2 = I$ and $d(\Gamma) = 2d(C)$.*

Proof. As in the previous case, apply Theorem 1, using clauses (i), (iii), and (iv), with $H = 2m$. We get a self-dual A of length $N = \Theta(m)$, and by equation (15) $d(A) = 2d(C)$. Lemma 6 shows how to build a simple balanced bipartite graph G on N vertices with adjacency matrix Γ where $\Gamma^2 = I$ and $d(\Gamma) = d(A) = 2d(C)$. Both constructions are deterministic and efficient. □

Proof of Theorem 3. We start with a hard self-dual instance (A,T) of length N from Corollary 4 and apply Lemma 6. The reduction $(A,T) \mapsto (\Gamma,T)$ preserves both distance and length, is efficient and deterministic. The resulting simple balanced bipartite graph satisfies $\Gamma^2 = I$. Thus, the reduction preserves the linear additive gap and a constant relative distance. Finally, observe that if $c_0 N$ is the source gap, an estimate within additive error $c_0 N/2$ distinguishes the YES and NO instances by comparison with $T + c_0 N/2$. Now, set $c = c_0/2$. □

5 Linear-size self-dual hosts

We now prove Proposition 7, which provides the host M that we use in Section 3. We stress that this proof is independent of the completion and hardness arguments presented in the previous sections. We start with a theorem of Shpilka and Guruswami which gives an explicit "dual-good" code with constant rate and constant relative distance, whose dual also has a constant rate and constant relative distance. Let us first restate Theorem 4 in Shpilka's paper in our notation. The construction and proof are given in Appendix A of Shpilka's paper, which states that the appendix was written by Guruswami:

Theorem 11 (Shpilka-Guruswami, [Shp09, Theorem 4]). *$\exists$ constant $\zeta \geqslant 1/30$ such that, $\forall$ $j \geqslant 1$, $\exists$ code $D_j \leqslant \mathbb{F}_2^{n_j}$ with $n_j = 42 \cdot 8^{j+1}$, with $\dim D_j = n_j/2$, and with $\min\{d(D_j), d(D_j^\perp)\} \geqslant \zeta n_j$. Construction, encoding, and decoding from a fraction $\zeta/2$ of errors all take poly(n_j) time and are deterministic.*

The statement in Shpilka's paper denotes the two relative distances by $\eta, \eta' \geqslant \zeta$. For our purposes, set $\delta = 1/30$ and observe that distances are at least δn_j, and $n_{j+1} = 8n_j$. We compute a basis of D_j by encoding the standard basis vectors of $\mathbb{F}_2^{n_j/2}$. We do not need the decoding algorithm, and we do not make an assumption that D_j contains the all-one word.

Instead, we take the Shpilka-Guruswami theorem as a black box construction and modify any such dual-good code in two modular steps using elementary linear algebra:

1. We first show how to put an all-ones word into any dual-good code without losing its dual-good property. The idea is to append the code and its dual, then apply an invertible linear map to each coordinate pair.

2. Next, we show how, starting with any dual-good code containing an all-ones vector, we can construct a self-dual code with a constant rate and constant relative distance. The idea is to use prefix sums to obtain self-duality.

Each step of this construction preserves a lower bound on distances proportional to the original length and doubles the length of the code. We denote $\mathbf{1}_n$ for the all-ones vector in $\mathbb{F}_2^n$.

5.1 Putting the all-ones word into any dual-good code.

The following normalization preserves the smaller of the code and dual distances. Swapping the two coordinate blocks of the constructed code gives its dual.

Lemma 12 (All-ones normalization). *For every code $D \leqslant \mathbb{F}_2^n$, where $n \geqslant 1$, one can efficiently construct a code $E \leqslant \mathbb{F}_2^{2n}$ such that $\dim E = n$, $\mathbf{1}_{2n} \in E$, and $\min\{d(E), d(E^\perp)\} \geqslant \min\{d(D), d(D^\perp)\}$.*

Proof. For every $h \in D \cap D^\perp$, we have $\mathbf{1}_n^\top h = h^\top h = 0$ (in $\mathbb{F}_2$). Thus, $\mathbf{1}_n \in (D \cap D^\perp)^\perp = (D + D^\perp)$. We solve for $a \in D$ and $b \in D^\perp$ with $a + b = \mathbf{1}_n$, and define E where multiplication is coordinatewise as:

$$E = \{(ax + by,\ x + y) : x \in D,\ y \in D^\perp\}$$

Observe that from any output $(r, s) = (ax + by, x + y)$, we can recover the input $x = r + bs$ and $y = r + as$. Thus, the map is linear and injective. Therefore $\dim E = n = \dim D + \dim D^\perp$. If we set $(x, y) = (a, b)$ this gives $(r, s) = (\mathbf{1}_n, \mathbf{1}_n)$. Therefore $\mathbf{1}_{2n} \in E$.

Now, we can show that swapping the two blocks of E gives its dual: consider any two output pairs (r, s) and (r', s') that came from two input pairs (x, y) and (x', y'). Expanding and using $a + b = \mathbf{1}_n$ gives: $r^\top s' + s^\top r' = x^\top y' + y^\top x' = 0 + 0$. Therefore, the code obtained when we swap the two blocks of E is contained in $E^\perp$. Since both have dimension n, $E^\perp = \{(s, r) : (r, s) \in E\}$, and $d(E^\perp) = d(E)$.

Now, observe that $\mathrm{wt}(r, s) = \mathrm{wt}(r) + \mathrm{wt}(s) \geqslant \max\{\mathrm{wt}(x), \mathrm{wt}(y)\}$ This follows from $x = r + bs$ and $y = r + as$. When $(r, s) \neq 0$, at least one of x, y is nonzero; therefore, $\mathrm{wt}(r, s) \geqslant \min\{d(D), d(D^\perp)\}$. This means that both E and $E^\perp$ have a distance of at least $\min\{d(D), d(D^\perp)\}$. □

5.2 The prefix-sum construction.

We next turn a code E containing the all-ones word into a self-dual code. We repeat each coordinate of a word in E and add consecutive prefix sums of a word in $E^\perp$. Summing each output pair recovers the latter word. Let $E \leqslant \mathbb{F}_2^n$ and assume that $\mathbf{1}_n \in E$. For $w \in E^\perp$, we define a prefix sum of all bits of w: $s_0 \stackrel{\text{def}}{=} s_0(w) = 0$ and for $1 \leqslant i \leqslant n$: $s_i \stackrel{\text{def}}{=} s_i(w) = \sum_{j=1}^{i} w_j$. Observe that $s_0 = 0$ by definition. Since $w \in E^\perp$, it is orthogonal to $\mathbf{1}_n$, and thus $s_n(w) = 0$. Define $\Phi : E \times E^\perp \to \mathbb{F}_2^{2n}$ by

$$\Phi(c, w) \stackrel{\text{def}}{=} (\Phi_1(c, w), \ldots, \Phi_n(c, w))$$

where $\Phi_i(c, w) = \big(c_i + s_{i-1}(w),\ c_i + s_i(w)\big)$. The output consists of n pairs, iterating over coordinates. Let $P = \operatorname{im} \Phi$. Our new code is

$$P = \Phi(E, E^\perp)$$

Lemma 13 (Prefix-sum self-dual construction). *$\forall$ codes $E \leqslant \mathbb{F}_2^n$ s.t. $\mathbf{1}_n \in E$, $\forall\ n \geqslant 1$, the map $\Phi : E \times E^\perp \to P$ is an invertible linear map. $\forall\ c, c' \in E$ and $w, w' \in E^\perp$,*

$$\Phi(c, w)^\top \Phi(c', w') = c^\top w' + w^\top c'$$

Moreover, code P is self-dual and has distance $d(P) \geqslant \min\{d(E^\perp),\ 2d(E)\}$.

Proof. **Linearity and dimension:** Φ is linear because prefix sums and coordinate addition are linear. To prove that it is one-to-one, consider an output $z = \Phi(c, w)$. From z, we can use the prefix-sum structure to recover w and then recover c: $w_i = z_{2i-1} + z_{2i}$ and $c_i = z_{2i-1} + s_{i-1}(w)$. Thus, $\dim P = \dim E + \dim E^\perp$ which is n.

Self-orthogonality: Let $c, c' \in E$ and $w, w' \in E^\perp$. Now, we want to show that it is always the case that $\Phi(c, w)^\top \Phi(c', w') = 0$. Linearity of Φ allows us to split arguments as: $\Phi(c, w) = \Phi(c, 0) + \Phi(0, w)$ where type one is: $\Phi(c, 0) = (c_1, c_1, \ldots, c_n, c_n)$ and type two is: $\Phi(0, w) = (0, s_1(w), s_1(w), \ldots, s_{n-1}(w), s_{n-1}(w), 0)$.

For $\Phi(c, 0)^\top \Phi(c', 0)$ each coordinatewise product $c_i c'_i$ occurs twice, and therefore cancels over $\mathbb{F}_2$. For $\Phi(0, w)^\top \Phi(0, w')$, each coordinatewise product $s_i(w) s_i(w')$, for $1 \leqslant i < n$, also occurs twice, in positions $(2i, 2i + 1)$, and the first and last points contribute zero, which results in cancellation as well.

Now, let's consider cross-products. For $\Phi(c, 0)^\top \Phi(0, w')$, output pair i contributes $c_i(s_{i-1}(w') + s_i(w')) = c_i w'_i$. Hence $\Phi(c, 0)^\top \Phi(0, w') = c^\top w' = 0$. Observe that the last equality holds since $c \in E$ and $w' \in E^\perp$. The same argument applies to the other cross-product pair $\Phi(0, w)^\top \Phi(c', 0) = w^\top c'$. Finally, expanding the dot product of two complete outputs gives four terms, where

$$\Phi(c, w)^\top \Phi(c', w') = c^\top w' + w^\top c' = 0$$

Thus $P \subseteq P^\perp$. Finally, since we already showed that $\dim P = n$ and the length of P is $2n$, we have $\dim P^\perp = 2n - n = n$, so $P = P^\perp$.

Distance. Take a nonzero output $z = \Phi(c, w)$. If $w \neq 0$, for coordinates where $w_i = 1$, pair i has the form:

$$(c_i + s_{i-1}(w), c_i + s_{i-1}(w) + 1)$$

which forces this pair to be either $(1, 0)$ or $(0, 1)$ and therefore

$$\mathrm{wt}(z) \geqslant \mathrm{wt}(w) \geqslant d(E^\perp)$$

On the other hand, if $w = 0$, every prefix sum is zero, so z repeats each coordinate of c. We know that $z \neq 0$, we have $c \neq 0$, and $c \in E$ contributes to the weight:

$$\mathrm{wt}(z) = 2\,\mathrm{wt}(c) \geqslant 2d(E)$$

These two cases prove the distance bound.

Construction time. Gaussian elimination computes a basis of $E^\perp$ from a basis of E. Applying Φ to the pairs $(u, 0)$ and $(0, v)$, where u and v range over these two bases, produces a basis of P by injectivity. All algorithms are deterministic and polynomial-time. □

5.3 Constructing linear-size self-dual codes

We can now go directly from dual-good to self-dual codes:

Proposition 14 (From dual-good to self-dual codes). *Let $\delta > 0$. Given $D \leqslant \mathbb{F}_2^n$, with both $d(D) \geqslant \delta n$, and $d(D^\perp) \geqslant \delta n$. We can deterministically and in poly(n)-time construct a self-dual code $P \leqslant \mathbb{F}_2^{4n}$ with $\dim P = 2n$ and $d(P) \geqslant \delta n$. Thus P has rate $1/2$ and relative distance $\geqslant \delta/4$.*

Proof. Apply all-ones Lemma 12 to D to get $E \leqslant \mathbb{F}_2^{2n}$, where $\mathbf{1}_{2n} \in E$, and $\min\{d(E), d(E^\perp)\} \geqslant \min\{d(D), d(D^\perp)\}$. Then apply prefix-sum Lemma 13 to E to get a self-dual $P \leqslant \mathbb{F}_2^{4n}$ with $d(P) \geqslant \min\{d(E^\perp), 2d(E)\} \geqslant \min\{d(D), d(D^\perp)\}$. Since P is self-dual, $\dim P = 2n$, which gives us $1/2$ rate. The distance bound above gives $d(P) \geqslant \delta n$. Therefore, the relative distance $\geqslant \delta/4$. □

We now apply Proposition 14 to Shpilka-Guruswami [Shp09] to construct:

Proof of Proposition 7. Apply Theorem 11 to construct dual-good code $D_j \leqslant \mathbb{F}_2^{n_j}$ with $\min\{d(D_j), d(D_j^\perp)\} \geqslant \delta n_j$. Choose the smallest $j \geqslant 1$ such that $\delta n_j \geqslant \Delta$. Shpilka-Gurswami guarantee that $n_j = \Theta(\Delta)$. (to see this, observe that δ and the first available length are constants, $n_{j+1} = 8n_j$, and for smallest j, we have $n_j = \Theta(\Delta)$). Now, apply (to D_j) the transformation from dual-good to self-dual codes from Proposition 14. This yields a self-dual P of length $N = 4n_j$ with $d(P) \geqslant \min\{d(D_j), d(D_j^\perp)\} \geqslant \delta n_j \geqslant \Delta$. Thus, $4 \mid N$ and $N = \Theta(\Delta)$ are as stated. Both Guruswami's construction of D_j and our conversions are deterministic and poly(N) time. □

6 AI methodologies

The mathematical ideas, results, proofs, and exposition in this paper are entirely the author's. After completing the manuscript, the author used OpenAI's ChatGPT-6 Astra to check for possible mathematical inconsistencies and typographical errors. It identified a few typographical errors but no errors in the mathematical arguments. The author takes full responsibility for the contents of this paper.

References

[AK14] P. Austrin and S. Khot. A simple deterministic reduction for the gap minimum distance of code problem. *IEEE Transactions on Information Theory*, 60(10):6636–6645, 2014. https://doi.org/10.1109/TIT.2014.2340869. Full version: https://arxiv.org/abs/1010.1481. 3

[AKK+25] J. An, N. Kaplan, J.-L. Kim, J. Luo, and G. Wang. Shortest self-orthogonal embeddings of binary linear codes. arXiv:2511.05440, 2025. Revised 2026. https://doi.org/10.48550/arXiv.2511.05440. Full version: https://arxiv.org/abs/2511.05440. 4

[BGLR25] V. Bhattiprolu, V. Guruswami, E. Lee, and X. Ren. Inapproximability of finding sparse vectors in codes, subspaces, and lattices. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science*, pages 1295–1303, 2025. https://doi.org/10.1109/FOCS63196.2025.00068. Full version: https://arxiv.org/abs/2410.02636v3. 3, 4

[CS96] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, 1996. https://doi.org/10.1103/PhysRevA.54.1098. Full version: https://arxiv.org/abs/quant-ph/9512032. 5

[CW12] Q. Cheng and D. Wan. A deterministic reduction for the gap minimum distance problem. *IEEE Transactions on Information Theory*, 58(11):6935–6941, 2012. https://doi.org/10.1109/TIT.2012.2209198. 3

[DMS03] I. Dumer, D. Micciancio, and M. Sudan. Hardness of approximating the minimum distance of a linear code. *IEEE Transactions on Information Theory*, 49(1):22–37, 2003. https://doi.org/10.1109/TIT.2002.806118. Full version: https://cseweb.ucsd.edu/~daniele/papers/DMS.pdf. 3

[GJS25] E. Grigorescu, V. Jha, and E. Samperton. On the hardness of approximating distances of quantum codes. In *Proc. FSTTCS*, 2025. https://doi.org/10.4230/LIPIcs.FSTTCS.2025.34; full version: https://arxiv.org/abs/2509.21469. 1, 2, 3, 6, 8

[Kap26] U. Kapshikar. Hardness of approximating quantum code distance beyond $\sqrt{N}$, 2026. https://doi.org/10.48550/arXiv.2609.33956; full version: https://arxiv.org/abs/2609.33956. 4

[KDP11] A. A. Kovalev, I. Dumer, and L. P. Pryadko. Low-complexity quantum codes designed via codeword-stabilized framework. *Physical Review A*, 84(6):062319, 2011. https://doi.org/10.1103/PhysRevA.84.062319. Full version: https://arxiv.org/abs/1108.5490. 2

[KK23] U. Kapshikar and S. Kundu. On the hardness of the minimum distance problem of quantum codes. *IEEE Transactions on Information Theory*, 69(10):6293–6302, 2023. https://doi.org/10.1109/TIT.2023.3286870. Full version: https://arxiv.org/abs/2203.04262. 1, 2, 3, 5

[Ost26] Rafail Ostrovsky. Metric Self-Dual Completion and Optimal Additive Hardness for Quantum and Graph-State Distance. arXiv:2609.22669, 2026. First submitted to arXiv on September 19, 2026. https://arxiv.org/abs/2609.22669. 0

[Shp09] A. Shpilka. Constructions of low-degree and error-correcting ϵ-biased generators. *Computational Complexity*, 18(4):495–525, 2009. Appendix A by V. Guruswami. https://doi.org/10.1007/s00037-009-0281-5. Full version: https://www.cs.tau.ac.il/~shpilka/publications/Shpilka_e_bias.pdf. 6, 10, 12

[Ste96] A. M. Steane. Error correcting codes in quantum theory. *Physical Review Letters*, 77(5):793–797, 1996. https://doi.org/10.1103/PhysRevLett.77.793. 5

Appendix A Type I and Type II metric self-dual completions

We give two generalizations of clauses (i)-(iv) of our main theorem: to Type I and Type II codes. These generalizations are not used in the main body of the paper, but we present them here for completeness. Recall that a code is doubly even if every word weight is divisible by four.

A.1 The completion and its recovery map

Let m, s, H be positive integers, where $H \geqslant s \cdot m$. Similar to our Section 3.1, let $M \leqslant \mathbb{F}_2^N$ be self-dual and $\tau : \mathbb{F}_2^m \to \mathbb{F}_2^N$ be an injective linear map. Let $W \stackrel{\text{def}}{=} \tau(\mathbb{F}_2^m)$ be self-orthogonal, with $\mathrm{wt}(\tau(x)) = s \cdot \mathrm{wt}(x)$. Assume that M is protected through H, namely,

$$\mathrm{wt}(u + \tau(x)) > H \qquad \text{for every } u \in M \setminus \{0\} \text{ and } x \in \mathbb{F}_2^m \tag{16}$$

The same triangle inequality as in equation (7) shows that $d(M) > 2H$ is sufficient for equation (16). For $C \leqslant \mathbb{F}_2^m$, we use the same construction as in our main theorem clauses (i)-(iv): $K \stackrel{\text{def}}{=} M \cap \tau(C)^\perp$ and $A \stackrel{\text{def}}{=} K + \tau(C)$.

Lemma 15 (Completion and recovery). *As in Theorem 1, A is self-dual, and τ induces isomorphism $\mathbb{F}_2^m/C \to (A+W)/A$ that increases all coset distances by s factor. The following properties hold simultaneously:*

(i) For every $y \in \mathbb{F}_2^m$ and $0 \leqslant r \leqslant H$,

$$\{p \in \tau(y) + A : \mathrm{wt}(p) \leqslant r\} = \tau\big(\{x \in y + C : \mathrm{wt}(x) \leqslant \lfloor r/s \rfloor\}\big) \tag{17}$$

Thus, if $C \neq 0$, then $d(A) = s \cdot d(C)$, and if $C = 0$, then $d(A) > H$.

(ii) Every $p \in A + W = K \oplus W$ has a unique expression $p = u + \tau(x)$, with $u \in K$ and $x \in \mathbb{F}_2^m$.

(iii) The linear "recovery" map $\rho : A + W \to \mathbb{F}_2^m$

$$\rho(u + \tau(x)) \stackrel{\text{def}}{=} x \tag{18}$$

guarantees $\rho \circ \tau = \mathrm{id}_{\mathbb{F}_2^m}$ and $A = \rho^{-1}(C)$. Note that ρ induces the inverse quotient isomorphism. For every $p \in A + W$, $p + \tau(\rho(p)) \in A$ and $s \cdot \mathrm{wt}(\rho(p)) \leqslant \mathrm{wt}(p)$.

(iv) If M and W are doubly even, then A is Type II.

(v) If $M \cap W^\perp$ contains a word of weight $2 \bmod 4$, then A is Type I.

(vi) Given M, τ, C, the completion A and recovery map ρ are computable by Gaussian elimination in poly(N).

Proof. By the argument already presented in Section 3.1 and Lemma 8, using equation (16), the same proof gives $M \cap W = 0$, self-duality of A, and equations (8) and (9).

Clause (ii). The direct sum $A + W = K \oplus W$ and τ injectivity gives the unique decomposition $p = u + \tau(x)$.

Clause (i). Every word of $\tau(y) + A$ must have this expression with $x \in y + C$. If $\mathrm{wt}(p) \leqslant H$, then equation (16) forces $u = 0$, and hence $\mathrm{wt}(p) = s \cdot \mathrm{wt}(x)$. The converse is also immediate, therefore proving equation (17). By the argument already presented in Section 3.2, with distance

scaling factor s in place of 2, the same proof gives the coset isomorphism and its metric guarantee. It also gives $d(A) = s \cdot d(C)$ when $C \neq 0$. If $C = 0$, then $A = M$, and equation (16), (with $x = 0$) yields $d(A) > H$.

Clause (iii). By clause (ii), equation (18) makes ρ linear. For $p = u + \tau(x)$, $\rho \circ \tau = \mathrm{id}$, the condition that $p \in A$ holds exactly when $x \in C$, and $p + \tau(\rho(p)) = u \in K \leqslant A$. Thus, ρ enjoys all the stated algebraic and quotient properties. The weight bound holds with equality when $u = 0$. If $u \neq 0$, then equation (16) yields $s \cdot \mathrm{wt}(\rho(p)) \leqslant s \cdot m \leqslant H < \mathrm{wt}(p)$.

Clause (iv). Both $K \leqslant M$ and $\tau(C) \leqslant W$ are doubly even, and they are orthogonal. Hence, equation (1) shows that their sum A is doubly even.

Clause (v). A word in $M \cap W^{\perp}$ lies in $K \leqslant A$. If its weight is 2 mod 4, the self-dual code A is therefore Type I.

Clause (vi). Observe that Gaussian elimination computes K, A, and ρ in $\mathrm{poly}(N)$ time. □

A.2 Type I completion

Recall that puncturing a code in a coordinate means deleting that coordinate from every codeword. On the other hand, shortening in that coordinate means restricting only to codewords that are zero in that coordinate and then deleting this coordinate.

Theorem 16 (Type I metric self-dual completion). *$\forall\ m, H \in \mathbb{Z}_{>0}$, $H \geqslant 2m$, $\exists$ deterministic and efficient way to construct Type I code $M \leqslant \mathbb{F}_2^N$, $N = \Theta(H)$, protected through H, with one-to-one embedding $\tau : \mathbb{F}_2^m \to \mathbb{F}_2^N$: $\tau(x) = (x, x, 0^{N-2m})$. Our construction depends only on m and H and guarantees that for every code $C \leqslant \mathbb{F}_2^m$, the completion $A = K + \tau(C)$ is Type I and has all the properties of Lemma 15 with $s = 2$.*

Proof. We apply Proposition 7 with $\Delta = 2H + 2$ to obtain a self-dual code $P \leqslant \mathbb{F}_2^n$ with $4 \mid n$, $n = \Theta(H)$, and $d(P) \geqslant 2H + 2$. Puncture its last coordinate to form $E \leqslant \mathbb{F}_2^{n-1}$. Since every binary self-dual code contains the all-ones word, $\mathbf{1}_{n-1} \in E$ and $d(E) \geqslant d(P) - 1 > 2H$. Recall that puncturing and shortening are dual operations. Since $P = P^{\perp}$, we obtain

$$E^{\perp} = \{w \in \mathbb{F}_2^{n-1} : (w, 0) \in P\}$$

and $d(E^{\perp}) \geqslant d(P) > 2H$. By the prefix-sum construction in Lemma 13, $M = \Phi(E, E^{\perp})$ is self-dual, has length $N = 2(n-1) = \Theta(H)$, and has distance greater than $2H$. Observe that since $n \geqslant d(P) \geqslant 2H + 2$, we have $N = 2(n-1) \geqslant 4H + 2 \geqslant 2m$. By the argument already presented in equation (7), the same proof gives equation (16), with $\mathrm{wt}(\tau(x)) = 2\,\mathrm{wt}(x) \leqslant 2m \leqslant H$. Since $4 \mid n$, we have $N = 2(n-1) \equiv 2 \pmod 4$. Every binary self-dual code of length N contains $\mathbf{1}_N$, whose weight is 2 mod 4, and hence is Type I. Thus, M is Type I. The embedding doubles weights, and its image W is self-orthogonal. Every word of W has even weight, so $\mathbf{1}_N \in W^{\perp}$. Since $\mathbf{1}_N \in M$ and has weight 2 mod 4, clause (v) of Lemma 15 applies and therefore makes every completion A Type I. We now apply clauses (i)- (iii) of Lemma 15. These give all the recovery and metric properties, while clause (vi) gives efficient maps. The construction depends only on m and H. □

A.3 Type II completion

Theorem 17 (Type II metric self-dual completion). *$\forall\ m, H \in \mathbb{Z}_{>0}$ s.t. $4m \leqslant H$, there is a deterministic poly-time construction of Type II self-dual $M \leqslant \mathbb{F}_2^N$, $N = \Theta(H)$, with 1-to-1 embedding $\tau : \mathbb{F}_2^m \to \mathbb{F}_2^N$ defined as $\tau(x) = (x, x, x, x, 0^{N-4m})$, so that M is protected through H. M and the embedding τ depend only on m and H. That is, for every code $C \leqslant \mathbb{F}_2^m$, define (as in our*

main construction) $K \stackrel{\text{def}}{=} M \cap \tau(C)^{\perp}$ *and* $A \stackrel{\text{def}}{=} K + \tau(C)$*. Then* A *is a Type II self-dual code and has all the metric and recovery properties of Lemma 15 with* $s = 4$*.*

Proof. We first apply Proposition 7 with $\Delta = 2H + 1$ to obtain a self-dual code $P \leqslant \mathbb{F}_2^n$ with $n = \Theta(H)$ and $d(P) > 2H$. We then interleave two copies of P:

$$E \stackrel{\text{def}}{=} \{(a_1, b_1, \ldots, a_n, b_n) : a, b \in P\}$$

Interleaving only permutes the coordinates of $P \oplus P$. Thus, $E = E^{\perp}$, $\mathbf{1}_{2n} \in E$, and $d(E) = d(P)$. By Lemma 13, the code $M = \Phi(E, E)$ is self-dual, has length $N = 4n = \Theta(H)$, and has distance greater than $2H$. To show that M is Type II, we still need to check that it is doubly even. Take any $\Phi(c, w) \in M$ and re-write $\Phi(c, w) = \Phi(c, 0) + \Phi(0, w)$. Observe that $\Phi(c, 0)$ has weight $2\,\mathrm{wt}(c)$. As E is self-dual, all its codewords have even weight, so $2\,\mathrm{wt}(c)$ is divisible by 4. Now consider the second summand. For $w = (a_1, b_1, \ldots, a_n, b_n)$, define the prefix sums $s_0 \stackrel{\text{def}}{=} 0$ and $s_j \stackrel{\text{def}}{=} \sum_{\ell=1}^{j} w_\ell$ for $1 \leqslant j \leqslant 2n$. Since $w \in E$ has even weight, $s_{2n} = 0$. Each prefix sum s_j with $1 \leqslant j < 2n$ appears twice in $\Phi(0, w)$, and the first and last coordinates are zero. It follows that

$$\mathrm{wt}(\Phi(0, w)) = 2\big|\{j : 1 \leqslant j < 2n,\ s_j = 1\}\big|$$

We claim that the number of nonzero prefix sums on the right is even. Working over $\mathbb{F}_2$, we have

$$\sum_{j=1}^{2n-1} s_j = \sum_{i=1}^{n} (s_{2i-1} + s_{2i}) = \sum_{i=1}^{n} b_i = 0$$

The first (equality) step in the equation above uses the fact that $s_{2n} = 0$. For the second step, observe that consecutive prefix sums satisfy $s_{2i-1} + s_{2i} = b_i$ (this is exactly the same trick as in our Lemma 13 proof.) The last step holds because $b \in P$ has even weight. Thus, the number of nonzero prefix sums is even, and hence $\mathrm{wt}(\Phi(0, w)) \equiv 0 \pmod 4$.

Both summands belong to the self-dual code M, so they are orthogonal. By equation (1), their sum has weight divisible by four as well. This proves that M is Type II.

We next check the embedding. Since $n \geqslant d(P) > 2H \geqslant m$, we have $N = 4n \geqslant 4m$, so τ is well-defined. Its image is doubly even and hence self-orthogonal. Also, $\mathrm{wt}(\tau(x)) = 4\,\mathrm{wt}(x) \leqslant 4m \leqslant H$. Together with $d(M) > 2H$, this bound gives equation (16) by the triangle inequality.

All the hypotheses of Lemma 15 now hold. Clauses (i) - (iii) give the metric and recovery properties, and Clause (iv) shows that A is Type II. Efficient completion and recovery follow from Clause (vi). The construction of M and τ depends only on m, H and is deterministic poly(N)-time. □