Quantum Science and Technology

PAPER • OPEN ACCESS

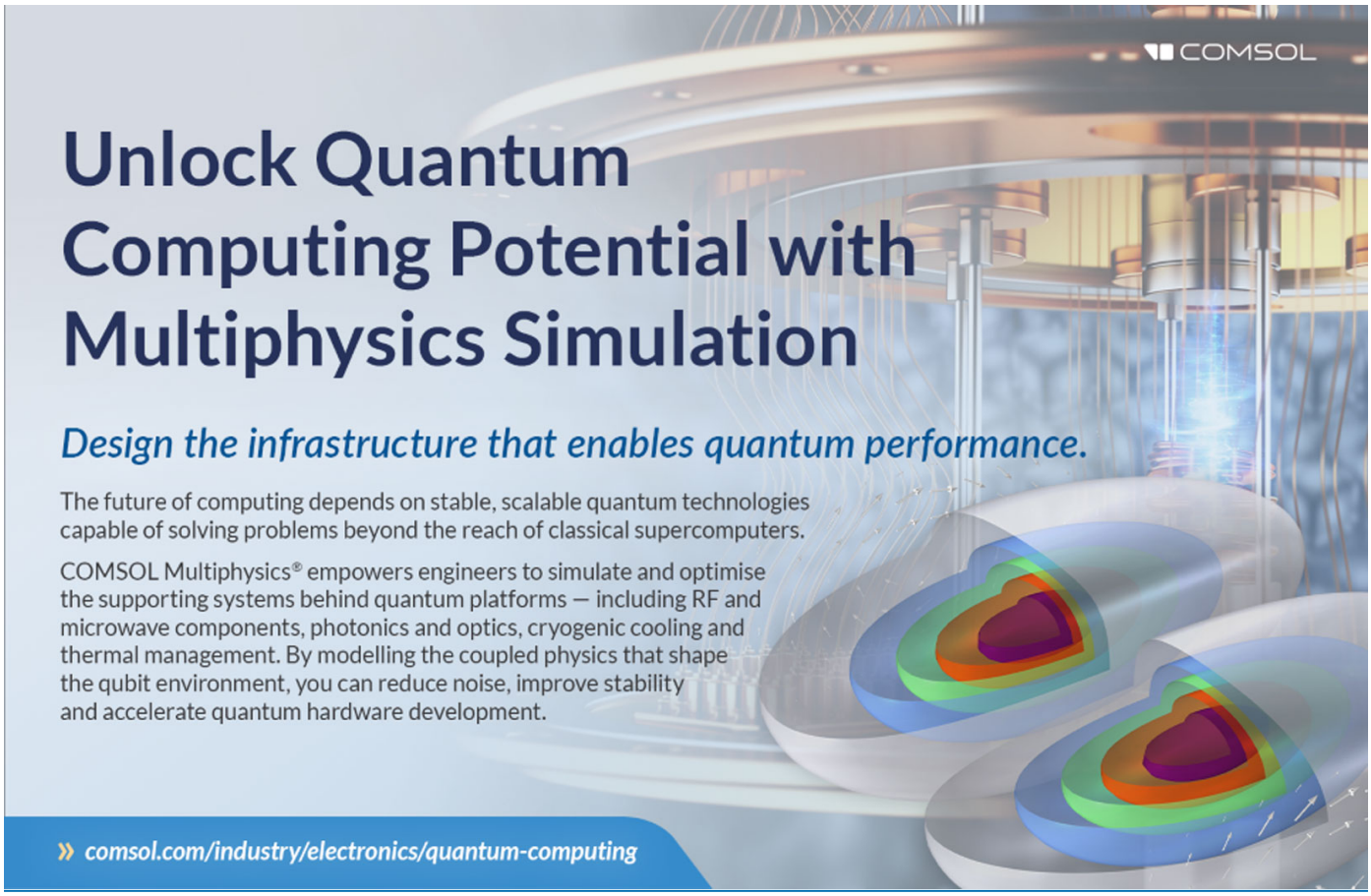
High-threshold magic state distillation with quantum quadratic residue codes

To cite this article: Michael Zurel *et al* 2026 *Quantum Sci. Technol.* **11** 045052

View the article online for updates and enhancements.

You may also like

- Resource-efficient quantum simulation of transport phenomena via Hamiltonian embedding
Joseph Li, Gengzhi Yang, Jiaqi Leng et al.
- Magic steady state production: non-Hermitian, dissipative, and stochastic pathways
Pablo Martinez-Azcona, Matthieu Sarkis, Alexandre Tkatchenko et al.
- Exact and approximate conditions of tabletop reversibility: when is Petz recovery cost-free?
Minjeong Song, Hyukjoon Kwon and Valerio Scarani



This content was downloaded from IP address 38.49.173.7 on 11/11/2026 at 15:26

Quantum Science and Technology

Check for updates

OPEN ACCESS

RECEIVED
31 March 2026

REVISED
11 July 2026

ACCEPTED FOR PUBLICATION
7 September 2026

PUBLISHED
25 September 2026

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

High-threshold magic state distillation with quantum quadratic residue codes

Michael Zurel*, Santanil Jana and Nadish de Silva

Department of Mathematics, Simon Fraser University, Burnaby, Canada
* Author to whom any correspondence should be addressed.

E-mail: mzurel@sfu.ca

Keywords: universal quantum computation, fault-tolerant quantum computation, magic state distillation, quantum codes, quadratic residue codes

Abstract

We present applications of quantum quadratic residue (QR) codes in magic state distillation. This includes showing that existing codes which are known to distill magic states, like the 5-qubit perfect code, the 7-qubit Steane code, and the 11-qutrit and 23-qubit Golay codes, are equivalent to certain quantum QR codes. We also present new examples of quantum QR codes that distill qubit T states and qutrit Strange states with high thresholds, and we show that there are infinitely many quantum QR codes that distill T states with a non-trivial threshold. All of these codes, including the codes with the highest currently known thresholds for T state and Strange state distillation, are unified under the umbrella of quantum QR codes.

Contents

1. Introduction 2
2. Background 3
   2.1. Classical coding theory 3
   2.2. Quantum coding theory 4
   2.3. Correspondence between classical and quantum codes 6
   2.4. Magic state distillation 7
3. QR codes 8
   3.1. Quantum QR codes 8
   3.2. Parameters of quantum QR codes 9
   3.3. CSS quantum QR codes 11
   3.4. Transversal Clifford gates of quantum QR codes 13
4. Magic state distillation with quantum QR codes 14
   4.1. Special cases 14
   4.2. Qubit T state distillation 14
      4.2.1. T state distillation and weight enumerators 14
      4.2.2. Distilling T states with quantum QR codes 18
      4.2.3. Infinitely many quantum QR codes distill T states 20
   4.3. Qutrit Strange state distillation 22
      4.3.1. Wigner functions and weight enumerators 22
      4.3.2. Distilling Strange states with quantum QR codes 24
5. Discussion 25
Acknowledgments 27
Data availability statement 27
Appendix A. The structure of magic state distillation protocols for qudits 27
Appendix B. Stabilizer operations in phase space 29

© 2026 The Author(s). Published by IOP Publishing Ltd

Appendix C. Weight enumerator tables 31
Appendix D. Post-distillation noise plots 33
Appendix E. Resource analysis for distillation with concatenated codes 34
References 37

1. Introduction

Magic state distillation is one of the leading candidates for scalable fault-tolerant quantum computation [1, 2]. In this approach, the stabilizer formalism describes the operations that can be performed fault-tolerantly, namely, preparation of stabilizer states, Clifford unitary gates, and Pauli measurements. But these operations alone are not universal for quantum computation [3–6]. Furthermore, the Eastin–Knill theorem shows that no universal set of unitary gates can be implemented transversally in any fixed stabilizer code [7].

Quantum computational universality is recovered by injecting 'magic' (non-stabilizer) states into otherwise stabilizer quantum circuits. The problem of fault-tolerantly implementing a universal set of computational primitives is thus shifted from that of fault-tolerantly implementing a universal set of unitary gates, to that of preparing sufficiently high-fidelity magic states. For this task, Bravyi and Kitaev showed that a supply of noisy magic states can be consumed by a so-called *magic state distillation protocol* in order to return fewer purified copies of the magic state [1]. Repeated application of these protocols allow for the production of nearly pure magic states, enabling universal quantum computation. Since then, many new magic states and distillation protocols have been discovered [8–19].

A foundational question for this quantum computational model, with implications for quantum computation and quantum information more broadly, is: *which quantum states can be distilled using only stabilizer operations to useful magic states?* We call these states *universal* for quantum computation. For qubits, a long-standing conjecture is that any state that is not a convex mixture of stabilizer states is universal [9, 14, 20–22]. The corresponding conjecture for odd-prime-dimensional qudits is that any state whose Wigner function takes negative values is universal [13, 16, 23–25].

Mixtures of stabilizer states for qubits and Wigner nonnegative states for odd-prime-dimensional qudits are known to not be distillable to useful magic states [3, 21, 23, 24]. Thus, nonstabilizerness and Wigner negativity are necessary preconditions for universality. Both conjectures align with traditional notions of non-classicality. Stabilizer states with stabilizer operations are efficiently simulatable classically [3, 4, 6], and therefore are classical from a computational point of view, whereas Wigner negativity is a traditional notion of non-classicality in physics [26–31]. Furthermore, negativity in the Wigner function implies the failure of a classical (i.e. noncontextual) hidden variable description of the computation [32–36]. The conjectures assert that the respective notions of non-classicality (non-stabilizerness for qubits, Wigner negativity for odd-prime-dimensional qudits) are not only necessary but also sufficient for quantum computational universality. Both conjectures remain open.

An obstruction to proving these conjectures is that no protocol of finite size can distill useful magic states from resource states arbitrarily close to certain regions of the boundary of the known non-universal regions. Campbell and Browne showed that, for qubits, no fixed finite-size stabilizer code can distill useful magic states from states arbitrarily close to a face of the stabilizer polytope [37]. For any given protocol size there is always a finite gap between the best achievable noise threshold of a protocol of that size and the face of the stabilizer polytope. Therefore, a positive resolution of the qubit conjecture would require a sequence of distillation protocols of increasing size whose distillation thresholds converge to the stabilizer polytope boundary. Prakash and Gupta established an analogous limitation for odd-dimensional qudits [25], implying that any proof of the Wigner negativity conjecture would similarly need an infinite family of protocols with thresholds approaching the relevant boundary.

Simultaneously, there is a common expectation that increasing code size tends to make distillation thresholds worse rather than better. This expectation is supported empirically by the observation that the codes with the best known noise thresholds for particular target magic states are often among the smallest codes that distill those states. For example, for qubit T-type magic states, the best known threshold is achieved by the 5-qubit code [1]. For qutrit 'Strange' states, the best known threshold is achieved by the 11-qutrit Golay code [16, 19]. Together these observations suggest that, while large codes are necessary in principle to get a sequence of codes with distillation thresholds approaching the boundary of the non-universal region, identifying code families that maintain high thresholds with increasing code size is non-trivial.

This motivates the present focus on quantum quadratic residue (QR) codes and their role in defining such infinite families of magic state distillation protocols. The first contribution of this paper is to show

that several known distillation protocols, such as those based on the 5-qubit perfect code (for T state distillation), the 7-qubit Steane code (for H state distillation), and the 11-qutrit Golay code (for Strange state distillation), are equivalent to certain quantum QR codes. The second contribution is to identify new members of the quantum QR code family that distill qubit T states and qutrit Strange states.

Although none of the new protocols we find surpass the noise thresholds of the 5-qubit code or the 11-qutrit Golay code, we do obtain fairly large codes that distill these states with high thresholds. This fact, together with the unification of the best known codes for T state and Strange state distillation under a single umbrella, make quantum QR codes a natural candidate family for constructing infinite sequences of magic state distillation protocols that could resolve the above conjectures. However, these codes are notoriously hard to analyze and so, for now, computing the noise thresholds of these codes is a computational problem that can only be carried out for some relatively small examples, and we cannot conclude much about the asymptotic thresholds of these codes or any subsequence thereof. That said, one asymptotic result we do prove is that there are infinitely many quantum QR codes which distill qubit T states with a non-trivial threshold.

Outline. The rest of this paper is structured as follows. First, in section 2, we provide some background on classical and quantum coding theory and on magic state distillation, including a correspondence between quantum stabilizer codes and certain classical codes that is useful in defining and describing the codes and magic state distillation protocols of present interest. In section 3, we introduce classical and quantum QR codes and prove some of their properties. In section 4, we show that quantum QR codes distill magic states. This section includes rederivations of relations between the performance of a magic state distillation protocol and certain weight enumerators of the corresponding code [16, 19, 25, 38, 39], as well as methods for computing the weight enumerators of quantum QR codes. Finally, we conclude with a discussion of these results in section 5.

2. Background

Here we briefly review some background material on classical coding theory, quantum coding theory, and magic state distillation that will be necessary for stating and proving our main results.

2.1. Classical coding theory

An $[n,k,\delta]_q$ linear code is a k-dimensional subspace of the vector space $\mathbb{F}_q^n$, where $\mathbb{F}_q$ is the finite field of order q. Here, n is the length of the code, k is the dimension, q is the size of the alphabet, and δ is the minimum distance of the code. When the minimum distance is not relevant or not known, the notation $[n,k]_q$ is also used.

The vector space $\mathbb{F}_q^n$ comes equipped with the standard Euclidean inner product

$$\langle u,v\rangle = \sum_{k=1}^{n} u_k v_k \in \mathbb{F}_q. \tag{1}$$

The dual code of an $[n,k]_q$ code C is the $[n,n-k]_q$ code

$$C^{\perp} := \left\{ v \in \mathbb{F}_q^n \mid \langle u,v\rangle = 0 \ \forall u \in C \right\}. \tag{2}$$

A code C is called self-orthogonal (or, more specifically, Euclidean self-orthogonal) if $C \subseteq C^{\perp}$, and it is (Euclidean) self-dual if $C = C^{\perp}$.

When the size of the alphabet is a square, $q = r^2$, we can define another inner product on $\mathbb{F}_q^n$. To start, the Frobenius automorphism is $a \mapsto \bar{a} = a^r$, this is the generator of the Galois group $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_r) \cong \mathbb{Z}_2$. Then the Hermitian inner product on $\mathbb{F}_q^n$ is defined as

$$\langle u,v\rangle_{\mathrm{H}} = \sum_{k=1}^{n} \bar{u}_k v_k. \tag{3}$$

The Hermitian dual code $C^{\perp_{\mathrm{H}}}$ of a code C, along with Hermitian self-orthogonality and Hermitian self-duality, can be defined similar to the Euclidean case.

The weight enumerator polynomial of a code C is a useful invariant defined as

$$W_C(x,y) = \sum_{v\in C} x^{n-wt(v)} y^{wt(v)} = \sum_{w=0}^{n} A_w x^{n-w} y^{w}, \tag{4}$$

where $wt(v) := |\{k \mid v_k \neq 0\}|$ is the Hamming weight of the vector v, and the coefficients $A_w := |\{v \in C \mid wt(v) = w\}|$ determine the weight distribution of the code. The weight enumerator of a code C is related to that of the its dual code by the MacWilliams identity [40, 41],

$$W_{C^\perp}(x,y) = \frac{1}{|C|} W_C(x+(q-1)y, x-y). \tag{5}$$

This identity holds for both Euclidean and Hermitian duals [42].

Cyclic codes. A cyclic code of length n over $\mathbb{F}_q$ is a linear code $C \subset \mathbb{F}_q^n$ such that for any vector $c = (c_0, c_1, \ldots, c_{n-1}) \in C$, the code also contains the vector $(c_{n-1}, c_0, \ldots, c_{n-2})$ obtained by cyclically permuting the coordinates. We can identify cyclic codes with certain polynomials. First, vectors in $\mathbb{F}_q^n$ are in bijective correspondence with polynomials in $\mathbb{F}_q[x]$ of degree at most $n-1$ through the map $c = (c_0, c_1, \ldots, c_{n-1}) \mapsto c(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$. With multiplication modulo $x^n - 1$, a cyclic shift of the vector's coordinates corresponds to multiplication by x, $xc(x) = c_{n-1} + c_0 x + \cdots + c_{n-2}x^{n-1}$. Under this identification, cyclic codes are ideals of the ring $\mathbb{F}_q[x]/(x^n-1)$. The ring $\mathbb{F}_q[x]/(x^n-1)$ is a principal ideal ring, and every cyclic code (ideal of $\mathbb{F}_q[x]/(x^n-1)$), C, is generated by a unique monic polynomial $g(x)$ of minimum degree dividing $x^n - 1$. The polynomial $g(x)$ is called the generator polynomial of the code C, and $\dim(C) = n - \deg(g)$ [43].

Assume that $\gcd(n,q) = 1$, let $\mathbb{F}_{q^m}$ be the splitting field of $x^n - 1$ over $\mathbb{F}_q$, and let $\zeta \in \mathbb{F}_{q^m}$ be a primitive nth root of unity. The defining set, or set of zeros, of the code C is the set

$$\mathcal{Z}(C) := \left\{k \in \mathbb{Z}_n \mid g\left(\zeta^k\right) = 0\right\}. \tag{6}$$

Since, with the assumption $\gcd(n,q) = 1$, $x^n - 1 = \prod_{k\in\mathbb{Z}_n}(x - \zeta^k)$ in $\mathbb{F}_{q^m}[x]$, the zeros uniquely determine the generator polynomial, and hence also the code. The Euclidean dual code $C^\perp$ of a cyclic code C is a cyclic code with zeros [42]

$$\mathcal{Z}\left(C^\perp\right) = \mathbb{Z}_n \setminus (-\mathcal{Z}(C)), \tag{7}$$

and when $q = r^2$, the Hermitian dual $C^{\perp_H}$ of a cyclic code C is also a cyclic code with zeros [44, proposition 3.4]

$$\mathcal{Z}\left(C^{\perp_H}\right) = \mathbb{Z}_n \setminus (-r\mathcal{Z}(C)). \tag{8}$$

An idempotent of a cyclic code C is a polynomial $e(x) \in C$ such that $e(x)^2 = e(x)$ in $\mathbb{F}_q[x]/(x^n-1)$. When $\gcd(n,q) = 1$, each cyclic code has a unique idempotent generator, called the idempotent of the code.

Additive codes. Additive codes are more general than the linear codes described above. An additive code over a finite field $\mathbb{F}_q$ is not necessarily $\mathbb{F}_q$-linear, but it is closed under addition. In other words, an additive code $C \subset \mathbb{F}_q^n$ is a subgroup of $(\mathbb{F}_q^n, +)$. Additive codes will appear briefly in section 2.3, but most of this paper will be about linear codes.

See, for example, [42, 43] for a more thorough introduction to coding theory.

2.2. Quantum coding theory

Our setting is quantum computation on systems of multiple qudits, where each qudit has prime Hilbert space dimension d. For our main results, we restrict to the cases $d = 2$ and $d = 3$.

The single-qudit generalized Pauli operators are defined as

$$Z = \sum_{j\in\mathbb{Z}_d} \omega^j |j\rangle\langle j| \quad \text{and} \quad X = \sum_{j\in\mathbb{Z}_d} |j+1\rangle\langle j| \tag{9}$$

where $\omega = \exp(2\pi \mathrm{i}/d)$. The n-qudit Pauli operators are constructed from tensor products of the single-qudit operators. We label the n-qudit Pauli operators up to overall phases by the elements of the vector space $\mathbb{F}_d^{2n}$ as follows. For each $v = (v_z, v_x) \in \mathbb{F}_d^n \times \mathbb{F}_d^n$,

$$T_v = \begin{cases} \mathrm{i}^{-\langle v_z, v_x\rangle} \bigotimes_{k=1}^n Z^{[v_z]_k} X^{[v_x]_k} & \text{if } d = 2, \\ \omega^{-\langle v_z, v_x\rangle \cdot 2^{-1}} \bigotimes_{k=1}^n Z^{[v_z]_k} X^{[v_x]_k} & \text{if } d \text{ is odd}, \end{cases} \tag{10}$$

where 2^{-1} is the multiplicative inverse of 2 in $\mathbb{F}_d$. The Pauli group $\mathcal{P}$ is the group generated by the n-qudit Pauli operators.

The symplectic form on $\mathbb{F}_d^{2n}$ defined by $[u,v] := \langle u_z, v_x\rangle - \langle u_x, v_z\rangle$ computes the commutator of the Pauli operators through the relation

$$[T_u, T_v] := T_u T_v T_u^{-1} T_v^{-1} = \omega^{[u,v]}\mathbb{1}. \tag{11}$$

We define a function β that tracks the phases that appear when commuting Pauli operators compose through the relation

$$T_u T_v = \omega^{-\beta(u,v)} T_{u+v}, \qquad \forall u, v \in \mathbb{F}_d^{2n} \text{such that } [u,v] = 0. \tag{12}$$

With the phase convention chosen in equation (10), when d is odd, we have the identities $T_u^{-1} = T_u^\dagger = T_{-u}$ and $T_u T_v = \omega^{[u,v]\cdot 2^{-1}} T_{u+v}$ for all $u, v \in \mathbb{F}_d^{2n}$, and so $\beta \equiv 0$. These are the finite-dimensional analogues of the Weyl relations. For $d = 2$, there is no corresponding choice of phase convention that reproduces these relations [36].

The Clifford group, $\mathcal{C}\ell$, is the normalizer of the Pauli group in the unitary group, i.e. the subgroup of unitary operators that map Pauli operators to Pauli operators under conjugation. They act on Pauli operators as

$$g T_u g^\dagger = \omega^{\Phi_g(u)} T_{S_g u}, \qquad \forall g \in \mathcal{C}\ell \tag{13}$$

where $S_g \in \mathrm{Sp}(\mathbb{F}_d^{2n})$ is a symplectic map on $\mathbb{F}_d^{2n}$. The Pauli group is a normal subgroup of the Clifford group such that $\mathcal{C}\ell/\mathcal{P} \cong \mathrm{Sp}(\mathbb{F}_d^{2n})$.

A projector onto the eigenspace of a set $I \subset \mathbb{F}_d^{2n}$ of compatible Pauli observables corresponding to eigenvalues $\{\omega^{r(v)} \mid v \in I\}$ is given by

$$\Pi_I^r = \frac{1}{|I|} \sum_{v \in I} \omega^{-r(v)} T_v. \tag{14}$$

In order for Π_I^r to be a valid projector, the set I and the function $r : I \to \mathbb{F}_d$ must satisfy some conditions. First, to have idempotence, $(\Pi_I^r)^2 = \Pi_I^r$, I must be closed under addition, i.e. I is a subspace of $\mathbb{F}_d^{2n}$. Second, the operators labeled by I must commute so I must be an isotropic subspace, i.e. a subspace on which the symplectic form vanishes. Finally, the eigenvalues must be compatible so r must be a consistent (noncontextual) value assignment for I. That is, $r : I \to \mathbb{F}_d$ is a function satisfying

$$\omega^{-r(u)-r(v)} T_u T_v = \omega^{-r(u+v)} T_{u+v} \quad \forall u, v \in I, \tag{15}$$

or equivalently, with equation (12),

$$r(u) + r(v) - r(u+v) = -\beta(u,v) \qquad \forall u, v \in I. \tag{16}$$

An $[\![n, k, \delta]\!]_d$ quantum stabilizer code is a rank $n-k$ Abelian subgroup of the n-qudit Pauli group that contains no element of the center of the Pauli group other than $\mathbb{1}$. Such a code can be identified with a pair (I, r) where I is an $n-k$-dimensional isotropic subspace of $\mathbb{F}_d^{2n}$, and $r : I \to \mathbb{F}_d$ is a consistent value assignment for I. The code space is the subspace of the n-qudit Hilbert space fixed by Π_I^r. For some purposes, we can ignore the phases r and identify the code with the subspace I. Note that since I is a subspace of a vector space over a finite field, we can regard it as a classical linear code which is self-orthogonal with respect to the symplectic inner product.

The symplectic weight, or Pauli weight, of a vector $v = (v_z, v_x) \in \mathbb{F}_d^n \times \mathbb{F}_d^n$ is

$$swt(v) = \left|\left\{k \in \mathbb{Z}_n \mid [v_z]_k \neq 0 \vee [v_x]_k \neq 0\right\}\right|. \tag{17}$$

Equivalently, it is the number of qudits on which the Pauli operator T_v acts non-trivially. The distance δ of a stabilizer code is the minimum symplectic weight of the nonzero vectors in $I^\perp \setminus I$. Symplectic weight distributions and symplectic weight enumerators can be defined analogous to equation (4).

For more background on quantum error correction and fault-tolerant quantum computation see, for example, [45, 46].

2.3. Correspondence between classical and quantum codes

There is a correspondence between stabilizer codes on d-dimensional qudits where d is a prime, and certain additive codes over $\mathbb{F}_{d^2}$ [47–49]. To describe this correspondence, we first need a map $\iota : \mathbb{F}_d^{2n} \to \mathbb{F}_{d^2}^n$ defined as follows. Fix a quadratic extension $\mathbb{F}_{d^2}/\mathbb{F}_d$ and choose an $\alpha \in \mathbb{F}_{d^2} \setminus \mathbb{F}_d$ so that $\{1, \alpha\}$ is an $\mathbb{F}_d$-basis for $\mathbb{F}_{d^2}$. Then for each $v = (v_z, v_x) \in \mathbb{F}_d^n \times \mathbb{F}_d^n$,

$$\iota(v) = v_z + \alpha v_x \in \mathbb{F}_{d^2}^n. \tag{18}$$

Let $x \mapsto \bar{x} := x^d$ denote the Frobenius automorphism, this is the generator of the Galois group $\mathrm{Gal}(\mathbb{F}_{d^2}/\mathbb{F}_d) \cong \mathbb{Z}_2$. Then $\mathbb{F}_{d^2}^n$ comes equipped with a Hermitian inner product,

$$\langle x, y\rangle_{\mathrm{H}} = \sum_{k=1}^{n} \bar{x}_k y_k \in \mathbb{F}_{d^2}. \tag{19}$$

Another inner product on $\mathbb{F}_{d^2}^n$ will also be useful. First, we define the field trace map $Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d} : \mathbb{F}_{d^2} \to \mathbb{F}_d$ as[1]

$$Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(x) = x + \bar{x}. \tag{20}$$

Then, letting $\lambda = (\alpha - \overline{\alpha})^{-1} \in \mathbb{F}_{d^2}$ (this is well-defined since $\alpha \neq \overline{\alpha}$ for $\alpha \in \mathbb{F}_{d^2} \setminus \mathbb{F}_d$), we define an $\mathbb{F}_d$-bilinear alternating form on $\mathbb{F}_{d^2}^n$ by

$$\langle x, y\rangle_{\mathrm{MTH}} := Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda \langle x, y\rangle_{\mathrm{H}}) \in \mathbb{F}_d. \tag{21}$$

We call this the modified Trace-Hermitian inner product. If $\lambda = 1$, then this reduces to the usual Trace-Hermitian inner product $\langle x, y\rangle_{\mathrm{TH}} = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\langle x, y\rangle_{\mathrm{H}})$.

Through the map ι, the standard symplectic form on $\mathbb{F}_d^{2n}$ is exactly reproduced by the modified Trace-Hermitian form on $\mathbb{F}_{d^2}^n$. More precisely, for all $u, v \in \mathbb{F}_d^{2n}$,

$$[u, v] := \langle u_z, v_x\rangle - \langle u_x, v_z\rangle = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda \langle \iota(u), \iota(v)\rangle_{\mathrm{H}}) =: \langle \iota(u), \iota(v)\rangle_{\mathrm{MTH}}. \tag{22}$$

To see this, write $\iota(u) = u_z + \alpha u_x$ and $\iota(v) = v_z + \alpha v_x$. Then, expanding $\langle \iota(u), \iota(v)\rangle_{\mathrm{H}}$ gives

$$\langle \iota(u), \iota(v)\rangle_{\mathrm{H}} = \langle u_z, v_z\rangle + \alpha\langle u_z, v_x\rangle + \overline{\alpha}\langle u_x, v_z\rangle + \overline{\alpha}\alpha\langle u_x, v_x\rangle. \tag{23}$$

Since $Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda) = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\overline{\alpha}\alpha) = 0$, $Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\alpha) = 1$, and $Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\overline{\alpha}) = -1$, multiplying equation (23) by λ and taking a trace yields equation (22).

Therefore, an isotropic subspace $I \subseteq \mathbb{F}_d^{2n}$ (which specifies a stabilizer code up to phases) corresponds via ι to an $\mathbb{F}_d$-additive code $C = \iota(I) \subseteq \mathbb{F}_{d^2}^n$ which is self-orthogonal with respect to the inner product equation (21). Conversely, if $C \subseteq \mathbb{F}_{d^2}^n$ is an $\mathbb{F}_d$-additive code satisfying $C \subseteq C^{\perp_{\mathrm{MTH}}}$, then $I = \iota^{-1}(C)$ is an isotropic subspace of $\mathbb{F}_d^{2n}$, and hence defines a stabilizer code on n qudits with $k = n - \log_d |I|$ logical qudits.

The symplectic weight of $v = (v_z, v_x) \in \mathbb{F}_d^{2n}$ is equal to the Hamming weight of $\iota(v) \in \mathbb{F}_{d^2}^n$, since $\iota(v)_k = 0 \in \mathbb{F}_{d^2}$ if and only if $([v_z]_k, [v_x]_k) = (0, 0) \in \mathbb{F}_d^2$. Therefore, the distance of the corresponding stabilizer code is

$$\delta = \min\left\{wt(c) \mid c \in C^{\perp_{\mathrm{MTH}}} \setminus C\right\}. \tag{24}$$

Hermitian self-orthogonal codes (with respect to equation (19)) are necessarily also self-orthogonal with respect to the modified Trace-Hermitian inner product equation (21). Such codes are a special case of stabilizer codes. Similarly, stabilizer codes on d-dimensional qudits are only required to be $\mathbb{F}_d$-linear (i.e. the image under ι is an additive code over $\mathbb{F}_{d^2}$). Codes which under the map ι are $\mathbb{F}_{d^2}$-linear codes are also a special case of stabilizer codes. The codes of interest in this paper will happen to be $\mathbb{F}_{d^2}$-linear and Hermitian self-orthogonal.

The following lemma will be useful later on.

Lemma 1. *If a code $C \subseteq \mathbb{F}_{d^2}^n$ is $\mathbb{F}_{d^2}$-linear, then $C^{\perp_{\mathrm{MTH}}} = C^{\perp_{\mathrm{H}}}$.*

[1] The simple form of the trace map here comes from the fact that $\mathrm{Gal}(\mathbb{F}_{d^2}/\mathbb{F}_d)$ has order 2. In general, the field trace sums over every element of the Galois group.

Proof of lemma 1. If $y \in C^{\perp_{\mathrm{H}}}$, then $\langle x,y\rangle_{\mathrm{H}} = 0$ for all $x \in C$, and hence,

$$\langle x,y\rangle_{\mathrm{MTH}} = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\langle x,y\rangle_{\mathrm{H}}) = 0 \qquad \forall x \in C. \tag{25}$$

Therefore, $C^{\perp_{\mathrm{H}}} \subseteq C^{\perp_{\mathrm{MTH}}}$.

Conversely, let $y \in C^{\perp_{\mathrm{MTH}}}$, fix any $x \in C$, and let $b := \langle x,y\rangle_{\mathrm{H}}$. Since C is $\mathbb{F}_{d^2}$-linear, we have $ax \in C$ for every $a \in \mathbb{F}_{d^2}$, and so

$$0 = \langle ax,y\rangle_{\mathrm{MTH}} = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\langle ax,y\rangle_{\mathrm{H}}) = Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(\lambda\bar{a}b) \qquad \forall a \in \mathbb{F}_{d^2}. \tag{26}$$

Since a ranges over all of $\mathbb{F}_{d^2}$, so does $\lambda\bar{a}$. Hence, $Tr_{\mathbb{F}_{d^2}/\mathbb{F}_d}(zb) = 0$ for all $z \in \mathbb{F}_{d^2}$. This is only possible if $b = 0$, i.e. $\langle x,y\rangle_{\mathrm{H}} = 0$. Since $x \in C$ was arbitrary, $y \in C^{\perp_{\mathrm{H}}}$. Thus $C^{\perp_{\mathrm{MTH}}} = C^{\perp_{\mathrm{H}}}$.

□

In this paper, we will focus on two cases in particular, namely, qubits and qutrits.

Qubits. In this case, we take $\mathbb{F}_4 \cong \mathbb{F}_2[\alpha]/(\alpha^2+\alpha+1)$ with Frobenius conjugation $x \mapsto \bar{x} = x^2$ and trace map $Tr_{\mathbb{F}_4/\mathbb{F}_2}(x) = x + x^2$. For any $\alpha \in \mathbb{F}_4 \setminus \mathbb{F}_2$ we have $\alpha - \overline{\alpha} = 1$, hence $\lambda = 1$ and the modified Trace-Hermitian inner product reduces to the usual Trace-Hermitian inner product. We can choose an explicit identification of $\mathbb{F}_2^2 = \{0 = (0,0), z = (1,0), x = (0,1), y = (1,1)\}$ with $\mathbb{F}_4$, for example $(0,z,x,y) \longleftrightarrow (0,1,\alpha,\alpha^2)$. Then additive codes $C \subseteq \mathbb{F}_4^n$ which are Trace-Hermitian self-orthogonal correspond to n-qubit stabilizer codes.

Qutrits. Here, take $\mathbb{F}_9 \cong \mathbb{F}_3[\alpha]/(\alpha^2+1)$ with conjugation $x \mapsto \bar{x} = x^3$ and trace map $Tr_{\mathbb{F}_9/\mathbb{F}_3}(x) = x + x^3$[2]. In this case, additive codes $C \subseteq \mathbb{F}_9^n$ which are self-orthogonal with respect to equation (21) correspond to n-qutrit stabilizer codes.

2.4. Magic state distillation

A magic state distillation protocol is a procedure using stabilizer operations (preparations of stabilizer states, Clifford unitary gates, Pauli measurements, tracing out subsystems, post-selection and/or classical conditioning on measurement outcomes), that consumes a number of copies of a noisy magic state to produce fewer copies of a state with higher fidelity to the pure target magic state. In this paper, we focus on n-to-1 qudit distillation protocols.

Formally, we fix a target pure magic state $|M\rangle$ and an experimentally-preparable mixed resource state ρ. An n-to-1 qudit magic state distillation protocol is a sequence of stabilizer operations described by a completely positive, trace non-increasing map Φ. If the protocol distills $|M\rangle$, then there is a distillable region (a neighborhood of states) around $|M\rangle$ such that if the resource state ρ is in this distillable region, then with success probability $P_{\mathrm{s}} = Tr(\Phi(\rho^{\otimes n}))$, the post-distillation state $\rho' = \Phi(\rho^{\otimes n})/P_{\mathrm{s}}$ has greater fidelity with respect to $|M\rangle$ than the resource state, $Tr(|M\rangle\langle M|\,\rho') > Tr(|M\rangle\langle M|\,\rho)$. Iterating the protocol pushes the noise to zero allowing us to prepare essentially pure magic states.

Campbell and Browne [8] showed that, in analyzing useful magic state distillation protocols, it suffices to focus on protocols with a certain structure. In particular, they showed that for any n-to-1 qudit protocol Φ and resource state ρ, there is another protocol with the same or better performance (measured by the fidelity of the output state), of the form

$$\rho \longmapsto \rho' = \frac{\mathcal{R}(\rho^{\otimes n})}{Tr(\mathcal{R}(\rho^{\otimes n}))}, \tag{27}$$

where the recovery map $\mathcal{R}(A)$ is defined as

$$\mathcal{R}(A) = Tr_{2,\ldots,n}\left(D\Pi_I^r A \Pi_I^r D^\dagger\right). \tag{28}$$

Here, $I \subset \mathbb{F}_d^{2n}$ is an $n-1$-dimensional isotropic subspace, $r : I \to \mathbb{F}_d$ is a consistent value assignment, and D is a decoding Clifford operation that maps the code space of (I,r) onto the first physical qudit. In words, we (i) prepare n copies of the resource state, (ii) measure commuting Pauli observables labeled by $I \subset \mathbb{F}_d^{2n}$ and post-select on measurement outcomes $r : I \to \mathbb{F}_d$ (equivalently, project with Π_I^r), (iii) use a Clifford group operation to decode the code space, $I^\perp/I \cong \mathbb{F}_d^2$, onto the first physical qudit, and (iv) trace out the last $n-1$ qudits.

[2] Different choices of the irreducible quadratic polynomial in $\mathbb{F}_9 \cong \mathbb{F}_3[\alpha]/f(\alpha)$ (and different choices of basis element α) lead to different but equivalent maps ι and forms equation (21).

Although [8] proves this structural result only for qubits, the same reduction extends to prime-dimensional qudits with essentially the same argument, since the argument depends only on the properties of the stabilizer formalism that have direct analogues for odd-prime dimensions [5]. We provide a sketch of the proof for prime-dimensional qudits in appendix A.

We emphasize that this result concerns only the fidelity of the post-selected output state, and hence also the achievable noise threshold of the protocols. That is, the reduction of a general magic state distillation protocol to one of the form equation (27) preserves or improves the conditional output state fidelity, but it is not guaranteed to preserve the success probability of the protocol, and in fact, typically decreases it. Since the primary figure of merit in this paper is the distillation threshold, the restriction to stabilizer reductions is without loss of generality for our purposes.

3. QR codes

QR codes are three related families of codes defined by the set of QRs modulo an odd prime p. To distinguish them, they are called the augmented, extended, and expurgated QR codes.

Let $\mathbb{F}_q$ be the finite field of order q with q a prime power, let p be an odd prime such that $\gcd(p,q)=1$ and q is a QR modulo p, and let ζ be a primitive pth root of unity in the splitting field $\mathbb{F}_{q^m}$ of x^p-1 over $\mathbb{F}_q$. Let $\mathcal{Q}$ be the set of nonzero QRs modulo p, and let $\mathcal{N}=\mathbb{F}_p^\times\setminus\mathcal{Q}$ be the set of non-residues modulo p. We have $|\mathcal{Q}|=|\mathcal{N}|=(p-1)/2$.

The length p augmented QR code over $\mathbb{F}_q$ is a $[p,(p+1)/2]_q$ cyclic code with generator polynomial $\prod_{s\in\mathcal{Q}}(x-\zeta^s)$. The length p expurgated QR code over $\mathbb{F}_q$ is a $[p,(p-1)/2]_q$ cyclic code with generator polynomial $(x-1)\prod_{s\in\mathcal{Q}}(x-\zeta^s)$.

Extending the augmented QR code by adding a parity check digit gives the extended QR code, a $[p+1,(p+1)/2]_q$ code. The augmented QR code can be obtained from the extended QR code by puncturing on any coordinate, and the expurgated QR code can be obtained from the extended QR code by shortening on any coordinate.

See, for example, [42, 43] for more background on QR codes.

3.1. Quantum QR codes

When a code over $\mathbb{F}_{d^2}$ is Hermitian self-orthogonal, it can be identified with a quantum stabilizer code on d-dimensional qudits via the correspondence described in section 2.3. This motivates the following definition.

Definition 1. A quantum QR code on d-dimensional qudits is the image under the map ι^{-1} defined in section 2.3 of a Hermitian self-orthogonal expurgated QR code over $\mathbb{F}_{d^2}$.

Quantum QR codes have previously been introduced [49, 50], but only as examples of the procedure of constructing quantum codes from classical codes described in section 2.3. In this paper, we explore these codes in depth, along with their properties and applications.

The following theorem identifies when an expurgated QR code is Hermitian self-orthogonal, and thus, when it can be used to define a quantum QR code.

Theorem 1. *A length p expurgated QR code over $\mathbb{F}_q$ where $q=r^2$ is Hermitian self-orthogonal if and only if $-r$ is a quadratic non-residue modulo p.*

Proof of theorem 1. We use two facts about cyclic codes: (i) for two cyclic codes C_1 and C_2 with zeros $\mathcal{Z}(C_1)$ and $\mathcal{Z}(C_2)$, we have $C_1\subseteq C_2$ if and only if $\mathcal{Z}(C_2)\subseteq\mathcal{Z}(C_1)$ [44, lemma 3.5], and (ii) for a cyclic code $C\subset\mathbb{F}_q^n$, $q=r^2$, with zeros $\mathcal{Z}(C)$, the hermitian dual $C^{\perp_H}$ of C is a cyclic code with zeros $\mathcal{Z}(C^{\perp_H})=\mathbb{Z}_n\setminus(-r\mathcal{Z}(C))$ [44, proposition 3.4].

Let C be the length p expurgated QR code over $\mathbb{F}_q$, with $q=r^2$. By definition, C is the cyclic code with zeros

$$\mathcal{Z}(C):=\{0\}\cup\mathcal{Q}\subset\mathbb{Z}_p. \tag{29}$$

Then the Hermitian dual $C^{\perp_H}$ of C is a cyclic code with zeros

$$\mathcal{Z}\left(C^{\perp_H}\right)=\mathbb{Z}_p\setminus(\{0\}\cup-r\mathcal{Q}). \tag{30}$$

Therefore, $C\subseteq C^{\perp_H}$ if and only if $\mathbb{Z}_p\setminus(\{0\}\cup-r\mathcal{Q})\subseteq\{0\}\cup\mathcal{Q}$, or equivalently, $C\subseteq C^{\perp_H}$ if and only if $\mathcal{N}\subseteq-r\mathcal{Q}$. Since $\mathcal{Q}\leqslant\mathbb{F}_p^\times$ is the index-2 subgroup of squares modulo p, multiplication by $a\in\mathbb{F}_p^\times$ preserves $\mathcal{Q}$ and

$\mathcal{N}$ if and only if $a \in \mathcal{Q}$, and it swaps $\mathcal{Q}$ and $\mathcal{N}$ if and only if $a \in \mathcal{N}$. Thus, $\mathcal{N} \subseteq -r\mathcal{Q}$ if and only if $-r \in \mathcal{N}$, i.e. if and only if $-r$ is a quadratic non-residue modulo p. □

For the cases $q = 4$ and $q = 9$, we have the following corollaries.

Corollary 1. *An expurgated QR code over $\mathbb{F}_4$ is Hermitian self-orthogonal if and only if the code length is* 5 *or* 7 *modulo* 8.

Proof of corollary 1. Let C be a length p expurgated QR code over $\mathbb{F}_4$. From theorem 1, the condition for C to be Hermitian self-orthogonal is that -2 is a quadratic non-residue modulo p. Using multiplicativity of the Legendre symbol and the supplements to the law of quadratic reciprocity [51],

$$\left(\frac{-2}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot (-1)^{\frac{p^2-1}{8}}. \tag{31}$$

Now we check the four possible cases for an odd prime p mod 8.

$$\begin{cases} \text{If} p \equiv 1 \mod 8, \text{then} \frac{p-1}{2} \text{is even and} \frac{p^2-1}{8} isevenso\left(\frac{-2}{p}\right) = 1. \\ \text{If} p \equiv 3 \mod 8, \text{then} \frac{p-1}{2} \text{is odd and} \frac{p^2-1}{8} isoddso\left(\frac{-2}{p}\right) = 1. \\ \text{If} p \equiv 5 \mod 8, \text{then} \frac{p-1}{2} \text{is even and} \frac{p^2-1}{8} isoddso\left(\frac{-2}{p}\right) = -1. \\ \text{If} p \equiv 7 \mod 8, \text{then} \frac{p-1}{2} \text{is odd and} \frac{p^2-1}{8} isevenso\left(\frac{-2}{p}\right) = -1. \end{cases} \tag{32}$$

Therefore, -2 is a quadratic non-residue modulo p if and only if $p \equiv 5 \mod 8$ or $p \equiv 7 \mod 8$. □

Corollary 2. *An expurgated QR code over $\mathbb{F}_9$ is Hermitian self-orthogonal if and only if the code length is* 5 *or* 11 *modulo* 12.

Proof of corollary 2. Let C be a length p expurgated QR code over $\mathbb{F}_9$. From theorem 1, the condition for C to be Hermitian self-orthogonal is that -3 is a quadratic non-residue modulo p. Using multiplicativity of the Legendre symbol and quadratic reciprocity [51],

$$\left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = (-1)^{\frac{p-1}{2}} \cdot \left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}} = \left(\frac{p}{3}\right). \tag{33}$$

We have $\left(\frac{p}{3}\right) = -1$ if and only if $p \equiv 2 \mod 3$. An odd prime p must be 1, 3, 5, 7, or 11 modulo 12, and of these cases only 5 and 11 are compatible with p being 2 modulo 3. Therefore, -3 is a quadratic non-residue mod p if and only if $p \equiv 5 \mod 12$ or $p \equiv 11 \mod 12$. □

In the definition of QR codes above, we have some conditions on the length p of the code that are required in order for the code to exist. Namely, we assume that p is an odd prime, that $\gcd(p, q) = 1$, and that q is a QR modulo p. In the case of quantum QR codes on d-dimensional qudits, the corresponding expurgated QR code is over $\mathbb{F}_{d^2}$. That is, $q = d^2$ is a square. Therefore, q is a QR modulo any prime p such that $\gcd(p, q) = 1$, and so the third condition is satisfied automatically. For the case $q = 4$, any odd prime p satisfies $\gcd(p, 4) = 1$, and so the conclusion of corollary 1 is that there exists a quantum QR code on qubits for any odd-prime length p such that p is 5 or 7 modulo 8. For the case $q = 9$, any odd prime $p > 3$ satisfies $\gcd(p, 9) = 1$, and so with corollary 2, there exists a quantum QR code on qutrits for any odd-prime length p such that p is 5 or 11 modulo 12.

3.2. Parameters of quantum QR codes

As shown in the previous section, a quantum QR code on d-dimensional qudits exists for any odd-prime length p such that $\gcd(p, d^2) = 1$, and $-d$ is a quadratic non-residue modulo p. In particular, quantum QR codes on qubits exist for any odd-prime length that is 5 or 7 modulo 8, and quantum QR codes on qutrits exist for any odd-prime length that is 5 or 11 modulo 12.

The cardinality of a length p expurgated QR code over $\mathbb{F}_q$, $q = r^2$, is $|C| = q^{(p-1)/2} = r^{p-1}$. Therefore, a length p quantum QR code on d-dimensional qudits has a stabilizer group of size d^{p-1}, and has code space of dimension $d^{p-(p-1)} = d$. That is, quantum QR codes always encode exactly one qudit.

We can also determine the distance of quantum QR codes in terms of the distance of the corresponding expurgated QR codes. This is achieved by the following theorem.

Theorem 2. *Let C be a Hermitian self-orthogonal expurgated QR code with parameters $[p, (p-1)/2, \delta_C]_{d^2}$. The distance of the corresponding $[\![p, 1, \delta_Q]\!]_d$ quantum QR code $\iota^{-1}(C)$ is $\delta_Q = \delta_C - 1$.*

The proof of this theorem relies on Prange's Theorem.

Theorem (Prange's Theorem). *Let C be a homogeneous $[n,k,\delta]$ code over $\mathbb{F}_q$ with $\delta > 1$. Let C^* be the code obtained from C by puncturing on some coordinate, and let C_* be the code obtained from C by shortening on some coordinate. Then for $0 \leqslant i \leqslant n-1$, we have*

$$\begin{cases} A_i(C^*) = \frac{n-i}{n} A_i(C) + \frac{i+1}{n} A_{i+1}(C), \\ A_i(C_*) = \frac{n-i}{n} A_i(C). \end{cases} \tag{34}$$

The version of Prange's theorem quoted here is Theorem 10.12 in [43]. An immediate consequence of this theorem is that the weight enumerators of the codes C^* and C_* can be obtained by taking partial derivatives of the weight enumerator polynomial of C. In particular,

$$W_{C^*}(x,y) = \frac{1}{n}\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) W_C(x,y) \tag{35}$$

and

$$W_{C_*}(x,y) = \frac{1}{n}\frac{\partial}{\partial x} W_C(x,y). \tag{36}$$

The precondition for Prange's theorem to apply is that the code in question must be homogeneous. A sufficient condition for homogeneity is that the automorphism group of the code acts transitively on the coordinate positions [43]. The permutation action of the automorphism group on the coordinate positions of the extended QR code is transitive (in fact, it is 2-transitive). This follows from the Gleason–Prange theorem [52, 53], which states that, labeling the coordinates of the extended QR code by the elements of the projective line $\mathrm{P}^1(\mathbb{F}_p) = \{0,1,\ldots,p-1,\infty\}$, the permutation action of the automorphism group contains a copy of the projective special linear group $\mathrm{PSL}(2,p)$. The action of $\mathrm{PSL}(2,p)$ on the projective line is 2-transitive, and so Prange's theorem applies to extended QR codes. The augmented and expurgated QR codes are obtained from the extended QR code by puncturing and shortening respectively, and so we can obtain the weight distributions of these codes from the weight distribution of the extended code via Prange's theorem.

We can now proceed to the proof of theorem 2.

Proof of theorem 2. Let C be a $[p,(p-1)/2,\delta_C]_{d^2}$ Hermitian self-orthogonal expurgated QR code, and let $I = \iota^{-1}(C)$ be the corresponding $[\![p,1,\delta_Q]\!]_d$ quantum QR code. Since C is $\mathbb{F}_{d^2}$-linear, by lemma 1, $C^{\perp_{\mathrm{MTH}}} = C^{\perp_{\mathrm{H}}}$, so

$$\delta_Q := \min\left\{swt(v) \mid v \in I^\perp \setminus I\right\} = \min\left\{wt(c) \mid c \in C^{\perp_{\mathrm{H}}} \setminus C\right\}. \tag{37}$$

The zeros of the expurgated QR code are $\mathcal{Z}(C) = \{0\} \cup \mathcal{Q}$. Therefore, with equation (8), we have $\mathcal{Z}\left(C^{\perp_H}\right) = \mathbb{Z}_p \setminus (\{0\} \cup (-d\mathcal{Q}))$. Since C is Hermitian self-orthogonal, $-d \in \mathcal{N}$ by theorem 1, so multiplication by $-d$ swaps $\mathcal{Q}$ and $\mathcal{N}$, and therefore $-d\mathcal{Q} = \mathcal{N}$. Thus $\mathcal{Z}(C^{\perp_{\mathrm{H}}}) = \mathbb{Z}_p \setminus (\{0\} \cup \mathcal{N}) = \mathcal{Q}$, i.e. $C^{\perp_{\mathrm{H}}}$ is the length-p augmented QR code.

Now let $\tilde{C}$ be the length $p+1$ extended QR code over $\mathbb{F}_{d^2}$. Puncturing $\tilde{C}$ on any coordinate gives the augmented code, and shortening on any coordinate gives the expurgated code (see, e.g. [43]). Prange's theorem, equation (34), implies that shortening preserves the minimum distance while puncturing decreases it by exactly 1. Hence $\delta(C^{\perp_{\mathrm{H}}}) = \delta(\tilde{C}) - 1$ and $\delta(C) = \delta(\tilde{C})$, so $\delta(C^{\perp_{\mathrm{H}}}) = \delta_C - 1$.

Because, by definition, C has no nonzero vector of weight less than δ_C, any vector in $C^{\perp_{\mathrm{H}}}$ of weight $\delta_C - 1$ lies in $C^{\perp_{\mathrm{H}}} \setminus C$, and therefore $\delta_Q = \min\{wt(c) \mid c \in C^{\perp_{\mathrm{H}}} \setminus C\} = \delta_C - 1$. □

Corollary 3. *The minimum distance of a $[\![p,1,\delta]\!]_d$ quantum QR code satisfies*

$$\delta \geqslant \lceil\sqrt{p}\rceil - 1. \tag{38}$$

Proof of corollary 3. This follows from theorem 2 together with the square root bound for classical QR codes [42, chapter 16]. □

The parameters of some of the Hermitian self-orthogonal expurgated QR codes over $\mathbb{F}_4$ and $\mathbb{F}_9$ are shown in table 1, together with the parameters of the corresponding qubit and qutrit quantum QR codes.

Table 1. Parameters of classical expurgated quadratic residue codes over $\mathbb{F}_4$ and $\mathbb{F}_9$ which are Hermitian self-orthogonal, and the corresponding parameters of the qubit and qutrit quantum quadratic residue codes.

Classical code	Quantum code
$[5,2,4]_4$	$[\![5,1,3]\!]_2$
$[7,3,4]_4$	$[\![7,1,3]\!]_2$
$[13,6,6]_4$	$[\![13,1,5]\!]_2$
$[23,11,8]_4$	$[\![23,1,7]\!]_2$
$[29,14,12]_4$	$[\![29,1,11]\!]_2$
$[31,15,8]_4$	$[\![31,1,7]\!]_2$
$[37,18,12]_4$	$[\![37,1,11]\!]_2$
$[47,23,12]_4$	$[\![47,1,11]\!]_2$
$[53,26,16]_4$	$[\![53,1,15]\!]_2$
$[61,30,18]_4$	$[\![61,1,17]\!]_2$
$[71,35,12]_4$	$[\![71,1,11]\!]_2$
Classical code	Quantum code
$[5,2,4]_9$	$[\![5,1,3]\!]_3$
$[11,5,6]_9$	$[\![11,1,5]\!]_3$
$[17,8,8]_9$	$[\![17,1,7]\!]_3$
$[23,11,9]_9$	$[\![23,1,8]\!]_3$
$[29,14,12]_9$	$[\![29,1,11]\!]_3$
$[41,20,14]_9$	$[\![41,1,13]\!]_3$
$[47,23,15]_9$	$[\![47,1,14]\!]_3$

3.3. CSS quantum QR codes

In some cases, quantum QR codes happen to be CSS codes. This is the result of the following theorem.

Theorem 3. *Let C be a Hermitian self-orthogonal expurgated QR code of length p over $\mathbb{F}_{d^2}$. The corresponding quantum QR code, $\iota^{-1}(C)$, is a CSS code if and only if d is a QR modulo p. Furthermore, if d is a QR modulo p, then an expurgated QR code D of length p over $\mathbb{F}_d$ exists, and the Z and X stabilizers of $\iota^{-1}(C)$ are each given by D. That is, $\iota^{-1}(C)$ has a generator matrix of the form*

$$G = \left[\begin{array}{c|c} H & 0 \\ 0 & H \end{array}\right], \tag{39}$$

where H is a generator matrix of D.

Proof of theorem 3. First we consider the 'only if' direction. Suppose $I = \iota^{-1}(C)$ is a CSS code, that is, $I = (D_Z \times \{0\}) \oplus (\{0\} \times D_X)$ for classical codes $D_Z, D_X \subseteq \mathbb{F}_d^p$, so that $C = D_Z + \alpha D_X$. Since C is a linear code, it is preserved by multiplication by α. Therefore, for any $v \in D_Z$, the image of αv under ι^{-1} is $(0, v)$, so $D_Z \subseteq D_X$. Similarly, for $v \in D_X$ the image of $\alpha^2 v$ under ι^{-1} has Z component a nonzero multiple of v, so $D_X \subseteq D_Z$. Hence $D_Z = D_X =: D$ and $C = D \otimes_{\mathbb{F}_d} \mathbb{F}_{d^2}$, with $\dim_{\mathbb{F}_d}(D) = (p-1)/2$. The code D inherits cyclicity from C, and its generator polynomial, being the minimal-degree monic polynomial of C, is $g(x) = (x-1)\prod_{s\in\mathcal{Q}}(x-\zeta^s)$. Applying the Frobenius automorphism $y \mapsto y^d$ to the coefficients fixes $g \in \mathbb{F}_d[x]$ while permuting its roots by $\zeta^s \mapsto \zeta^{ds}$. Therefore $d\mathcal{Q} = \mathcal{Q}$, and since $\mathcal{Q}$ is the subgroup of squares in $\mathbb{F}_p^\times$, $d \in \mathcal{Q}$.

Now consider the 'if' direction. Assume d is a QR modulo p. Then an expurgated QR code D of length p over $\mathbb{F}_d$ exists [42]. Let $R = \mathbb{F}_d[x]/(x^p - 1)$, and let $e(x) \in R$ be the idempotent of D. By Theorem 12.3 of [43], $e(x)$ is also the idempotent of the expurgated QR code of length p over any extension field of $\mathbb{F}_d$, and in particular, $e(x)$ is the idempotent of C. That is, $D = \langle e(x)\rangle_{\mathbb{F}_d}$ and $C = \langle e(x)\rangle_{\mathbb{F}_{d^2}}$.

Now use the same choice of $\alpha \in \mathbb{F}_{d^2} \setminus \mathbb{F}_d$ as in section 2.3, so that $\mathbb{F}_{d^2} = \mathbb{F}_d \oplus \alpha\mathbb{F}_d$ as an $\mathbb{F}_d$-vector space, and hence $\mathbb{F}_{d^2}[x]/(x^p - 1) = R \oplus \alpha R$. Therefore,

$$C = (R \oplus \alpha R)\, e = Re + \alpha Re = D + \alpha D. \tag{40}$$

Then, by definition of the bijection ι, we have $\iota^{-1}(C) = D \times D$. Therefore, the corresponding quantum code is a CSS stabilizer code with both Z-checks and X-checks given by D.

Finally, if H is a generator matrix for D, then the rows of $[H|0]$ and $[0|H]$ generate $\iota^{-1}(C) \cong D \times D \subset \mathbb{F}_d^n \times \mathbb{F}_d^n$. Therefore $\iota^{-1}(C)$ has generator matrix

$$G = \left[\begin{array}{c|c} H & 0 \\ 0 & H \end{array}\right], \tag{41}$$

as claimed. □

Corollary 4. *A quantum QR code on qubits is a CSS code if and only if the length of the code is a prime $p \equiv 7$ mod 8.*

Proof of corollary 4. An expurgated QR code of odd-prime length p over $\mathbb{F}_2$ exists if and only if 2 is a QR modulo p. By the second supplement to the law of quadratic reciprocity [51], this occurs if and only if $p \equiv \pm 1 \mod 8$. Incorporating the Hermitian self-orthogonality condition corollary 1, a length p quantum QR code on qubits is a CSS code if and only if $p \equiv 7 \mod 8$. □

Corollary 5. *A quantum QR code on qutrits is a CSS code if and only if the length of the code is a prime $p \equiv 11$ mod 12.*

Proof of corollary 5. An expurgated QR code of odd-prime length p over $\mathbb{F}_3$ exists if and only if 3 is a QR modulo p. Using the third supplement to the law of quadratic reciprocity [51], this occurs if and only if $p \equiv \pm 1 \mod 12$. With the Hermitian self-orthogonality condition corollary 2, a length p quantum QR code on qutrits is a CSS code if and only if $p \equiv 11 \mod 12$. □

For qubits, the smallest example of a non-CSS quantum QR code is the $[\![5,1,3]\!]_2$ code. This code is equivalent to the well-known 5-qubit perfect code [54, 55] (also called the Laflamme code) with stabilizer group

$$\langle XZZXI, IXZZX, XIXZZ, ZXIXZ \rangle, \tag{42}$$

where by equivalent we mean up to permutations of qudits and local Cliffords. This can be verified directly, for example, using the Magma computer algebra system [56]. The smallest example of a CSS code among qubit quantum QR codes is the $[\![7,1,3]\!]_2$ code. This is the 7-qubit Steane code [57], a CSS code with stabilizer group

$$\langle IIIXXXX, IXXIIXX, XIXIXIX, IIIZZZZ, IZZIIZZ, ZIZIZIZ \rangle. \tag{43}$$

Each of the Z and X stabilizers are formed from a copy of the dual of the $[7,4,3]_2$ binary Hamming code, which itself is equivalent to the length 7 binary QR code.

For qutrits, the smallest example of a CSS code among quantum QR codes is the 11-qutrit Golay code [16]. This is a CSS code with Z and X stabilizers each given by a copy of the dual of the length 11 ternary Golay code. The length 11 ternary Golay code is equivalent to the length 11 ternary QR code [42], and so the 11-qutrit Golay code is equivalent to the 11-qutrit quantum QR code. The smallest non-CSS code is the $[\![5,1,3]\!]_3$ code with stabilizer generator matrix[3]

$$\left[\begin{array}{cc|cc} \mathbb{1}_\ell & A_Z & \mathbb{0}_\ell & A_X \\ \mathbb{0}_\ell & 2A_X & \mathbb{1}_\ell & A_Z \end{array}\right] \tag{44}$$

where $\ell = 2$,

$$A_Z = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{and} \quad A_X = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 2 & 0 \end{bmatrix}. \tag{45}$$

The next non-CSS code is the $[\![17,1,7]\!]_3$ code with stabilizer generator matrix equation (44) where $\ell = 8$,

$$A_Z = \begin{bmatrix} 2 & 2 & 1 & 2 & 0 & 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 & 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 & 1 & 2 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 2 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 2 & 1 & 1 & 2 & 1 \\ 1 & 1 & 2 & 2 & 1 & 0 & 0 & 2 & 2 \\ 2 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 0 & 0 & 2 & 1 & 2 & 2 \end{bmatrix}, \quad \text{and} \quad A_X = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 2 & 0 & 0 & 2 \\ 2 & 1 & 2 & 2 & 2 & 2 & 0 & 2 & 2 \\ 2 & 2 & 2 & 1 & 0 & 1 & 0 & 2 & 2 \\ 2 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 & 0 & 1 & 0 & 0 & 2 \\ 2 & 2 & 0 & 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 & 2 & 2 & 1 & 2 \\ 2 & 0 & 0 & 2 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{46}$$

[3] Here we use the conventions established in section 2.3 for the identification of $\mathbb{F}_9^n$ with $\mathbb{F}_3^{2n}$.

3.4. Transversal Clifford gates of quantum QR codes

In this section, we describe the transversal Clifford gates of quantum QR codes. Transversal gates act independently on the physical qudits, and therefore do not spread errors within a code block, making them a standard way of fault-tolerantly implementing logical operations [45, 46]. For the case of qubits, we can use the recent classification of [58], Theorem 6.1 therein provides a complete characterization of transversal Clifford gates of qubit stabilizer codes. This classification is a qubit-specific refinement of the approach based on endomorphism algebras of [50] which applies to stabilizer codes on qudits of any prime dimension.

Following [58], we consider ℓ code blocks of an n-qudit stabilizer code arranged as an $\ell \times n$ array of qudits with one code block per row, and we consider gates of the form $g^{\otimes n}$ where g is a fixed ℓ-qudit Clifford gate applied to each column. In [58], these are called *diagonal transversal gates* (note that the term 'diagonal' refers to their action on code blocks, the gates themselves need not be diagonal). Conjugating by g induces a symplectic transformation $S_g \in \mathrm{Sp}(\mathbb{F}_d^{2\ell})$ of the phase space $\mathbb{F}_d^{2\ell}$ (i.e. the labels of the Pauli operators equation (10)) of a single column, and conjugating by $g^{\otimes n}$ preserves the stabilizer group of ℓ code blocks if and only if the column-wise action of S_g preserves $I^{\oplus \ell}$. Therefore, we identify the group of transversal Clifford gates on ℓ code blocks with

$$G_I^\ell = \left\{ S \in \mathrm{Sp}\left(\mathbb{F}_d^{2\ell}\right) \mid S \text{applied to each column preserves } I^{\oplus \ell} \right\}. \tag{47}$$

Phases can be ignored for this identification since a transversal Clifford gate preserving the stabilizer group up to phases can be corrected by composing with a Pauli operator [30, 58].

Let I be a quantum QR code and C the corresponding linear code over $\mathbb{F}_{d^2}$, and $C^{(\ell)}$ denote ℓ code blocks of C. Identify the phase space of a column with $\mathbb{F}_{d^2}^\ell$ via the map ι defined in section 2.3 applied to each of the ℓ qudits. Any matrix $A \in M_\ell(\mathbb{F}_{d^2})$ acting column-wise preserves $C^{(\ell)}$, since the rows of the image are $\mathbb{F}_{d^2}$-linear combinations of codewords of C, and by equation (22) its action is symplectic if and only if $\bar{A}^T A = \mathbb{1}$, where $\bar{A}$ denotes the entry-wise Frobenius conjugate. Writing $\mathrm{U}(\mathbb{F}_{d^2}^\ell) := \{A \in M_\ell(\mathbb{F}_{d^2}) \mid \bar{A}^T A = \mathbb{1}\}$ for the unitary group of the Hermitian form equation (19), we obtain $\mathrm{U}(\mathbb{F}_{d^2}^\ell) \subseteq G_I^\ell$ for all $\mathbb{F}_{d^2}$-linear codes [50, 58], and in particular, for quantum QR codes. When the code is not CSS, these are all of the transversal Clifford gates, $G_I^\ell = \mathrm{U}(\mathbb{F}_{d^2}^\ell)$. This follows from Theorem 15 and Example 2 of [50] together with Theorem 5.5 of [58]. On the other hand, when the codes are $\mathbb{F}_{d^2}$-linear, and are also CSS codes with symmetric Z and X stabilizers, we have additional transversal gates, namely, $G_I^\ell = \mathrm{Sp}(\mathbb{F}_d^{2\ell})$ [19, 58].

We now specialize to qubit and qutrit quantum QR codes.

Qubits. For qubit QR codes we have two cases. First, if the length of the code is $p \equiv 7 \mod 8$, then the code is $\mathbb{F}_4$-linear, and corollary 4 shows that it is a CSS code with symmetric Z and X stabilizers. For such codes, the group of transversal Clifford gates is $\mathrm{Sp}(\mathbb{F}_2^{2\ell})$ [58]. That is, the codes have a complete transversal Clifford group. On the other hand, corollary 4 shows that codes of length $p \equiv 5 \mod 8$ are still $\mathbb{F}_4$-linear, but they are not CSS codes. The group of transversal Clifford gates for these codes is $\mathrm{U}(\mathbb{F}_4^\ell)$ [58]. Specializing to a single code block, $\ell = 1$, $\mathrm{U}(\mathbb{F}_4^1) = \{\mu \in \mathbb{F}_4^\times \mid \overline{\mu}\mu = \mu^3 = 1\} = \mathbb{F}_4^\times \cong \mathbb{Z}_3$. This group is generated by the transversal M_3 gate (also called the facet gate), which cyclically permutes the Pauli X, Y, and Z operators. Furthermore, these codes have no non-trivial transversal two-qubit gates, the only transversal two-qubit gates are swap gates followed by transversal single-qubit gates. This is spelled out explicitly for the case of the $[\![5,1,3]\!]_2$ code in corollaries 7.2 and 7.3 of [58]. The result is the same for all qubit QR codes of length $p \equiv 5 \mod 8$.

Qutrits. For qutrits, we also have two cases depending on whether or not the code is a CSS code. First, for codes of length $p \equiv 11 \mod 12$, the codes are $\mathbb{F}_9$-linear, and according to corollary 5 they are CSS codes with symmetric Z and X stabilizers. Such codes have $\mathrm{Sp}(\mathbb{F}_3^{2\ell})$ as their group of transversal Clifford gates, i.e. they have a full transversal Clifford group [19, lemma 1]. On the other hand, corollary 5 says that qutrit QR codes of length $p \equiv 5 \mod 12$ are $\mathbb{F}_9$-linear, but they are not CSS codes. Such codes have $\mathrm{U}(\mathbb{F}_9^\ell)$ as their group of transversal gates [50]. On a single code block, $\mathrm{U}(\mathbb{F}_9^1) = \{\mu \in \mathbb{F}_9^\times \mid \overline{\mu}\mu = \mu^4 = 1\} \cong \mathbb{Z}_4$, namely, the cyclic group $\langle \alpha \rangle = \{\pm 1, \pm\alpha\}$. Multiplication by α acts on $(\nu_z|\nu_x)$ as $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, the symplectic action of the inverse of the qutrit Fourier gate $F = \frac{1}{\sqrt{3}} \sum_{j,k \in \mathbb{Z}_3} \omega^{jk} |j\rangle\langle k|$. That is, non-CSS qutrit quantum QR codes have a transversal Fourier gate, with F^2 the parity gate $|j\rangle \mapsto |-j\rangle$ and $F^4 = \mathbb{1}$. In contrast with the qubit case, across two blocks the group $\mathrm{U}(\mathbb{F}_9^2)$, of order 96, strictly contains the subgroup of monomial transformations of order 32, and so there are non-trivial transversal two-qutrit gates.

4. Magic state distillation with quantum QR codes

Quantum QR codes are well-suited for defining n-to-1 qudit magic state distillation protocols. In this section, we first show that some quantum QR codes are equivalent to codes that are known to distill magic states, including the qubit H and T states and the qutrit Strange state. Then we analyze the performance of new quantum QR codes for distilling qubit T states and qutrit Strange states.

4.1. Special cases

Some quantum QR codes are equivalent to codes that have previously been shown to distill magic states, where by equivalent we mean up to permutations of qudits and local Cliffords. For example, as mentioned above, the qubit quantum QR code of length 5 is equivalent to the 5-qubit perfect code [54, 55] (also called the Laflamme code) with parameters $[\![5,1,3]\!]_2$. This code was the first code shown to distill T states [1], and it still has the best threshold among all codes known to distill T states.

As another example, the 7-qubit quantum QR code is equivalent to the Steane code [57], with parameters $[\![7,1,3]\!]_2$. This code distills qubit H states [20]. Since the length of the code is 7 mod 8, corollary 4 predicts that this code is a CSS code, as expected.

Finally, the 11-qutrit Golay code is known to distill qutrit Strange states [16]. As mentioned above, this code is equivalent to the 11-qutrit quantum QR code. Since the code length is 11 mod 12, this is a CSS code consistent with corollary 5. Similarly, the 23-qubit Golay code, a CSS code constructed from two copies of the dual of the length 23 binary Golay code, is equivalent to the 23-qubit quantum QR code. This code distills H states [20].

4.2. Qubit T state distillation

The state $|T\rangle = \cos\beta\,|0\rangle + e^{i\pi/4}\sin\beta\,|1\rangle$ with $\cos 2\beta = 1/\sqrt{3}$ is one of the first proposed magic states [1]. It is the single-qubit state with density matrix

$$|T\rangle\langle T| = \frac{1}{2}\left(I + \frac{1}{\sqrt{3}}(X+Y+Z)\right). \tag{48}$$

Rall showed that the performance of a magic state distillation protocol distilling T states is determined by the 'signed' weight enumerators of the corresponding code [38], and more recently, Kalra and Prakash showed that if the code has a transversal M_3 gate, then the performance is determined by the usual (unsigned) weight enumerators [39]. For completeness, we partially reproduce these results connecting distillation performance to weight enumerators here. Then we show how to compute the weight enumerators of quantum QR codes, and we analyze their performance for T state distillation.

4.2.1. T state distillation and weight enumerators

Let M_3 denote the single-qubit Clifford gate which cyclically permutes the Pauli operators, $M_3 X M_3^\dagger = Y$, $M_3 Y M_3^\dagger = Z$, and $M_3 Z M_3^\dagger = X$. Averaging over the group generated by M_3 gives a stabilizer channel

$$\mathcal{T}_T(\rho) := \frac{1}{3}\sum_{k\in\mathbb{Z}_3} M_3^k \rho M_3^{k\dagger}, \tag{49}$$

which maps arbitrary single-qubit states onto the line segment

$$\frac{1}{2}\mathbb{1} + \frac{t}{2\sqrt{3}}(X+Y+Z), \quad t\in[-1,1]. \tag{50}$$

For T state distillation, we assume we have access to a supply of noisy T states of the form

$$\rho_T(\epsilon) = (1-\epsilon)\,|T\rangle\langle T| + \frac{\epsilon}{2}\mathbb{1} = \frac{1}{2}\mathbb{1} + \frac{1-\epsilon}{2\sqrt{3}}(X+Y+Z). \tag{51}$$

No generality is lost in making this assumption since given any resource state we can always apply the twirl $\mathcal{T}_T$ to get a state of this form. Similarly, after a distillation protocol is performed, we can apply the twirl $\mathcal{T}_T$ again to get a new state $\rho_T(\epsilon')$ with a new noise parameter, ideally satisfying $\epsilon' < \epsilon$. A protocol is said to distill T states with threshold ϵ_{th} if there exists an $\epsilon_{\text{th}} > 0$ such that $\epsilon'(\epsilon) < \epsilon$ for all $0 < \epsilon < \epsilon_{\text{th}}$, where $\epsilon'(\epsilon)$ is the post-distillation noise as a function of the pre-distillation noise.

The starting point of an n-to-1 qudit distillation protocol is an n-fold tensor product of noisy T states. These states can be expanded in the Pauli basis as

$$\rho_T(\epsilon)^{\otimes n} = \frac{1}{2^n}\sum_{v\in\mathbb{F}_2^{2n}} t(\epsilon)^{swt(v)}\,T_v, \tag{52}$$

where $t(\epsilon) = (1-\epsilon)/\sqrt{3}$, and $swt(v)$ is the symplectic weight of the Pauli label $v \in \mathbb{F}_2^{2n}$.

Fix an n-to-1 qudit distillation protocol defined by a $[\![n,1]\!]_2$ stabilizer code $I \subset \mathbb{F}_2^{2n}$, post-selection on measurement outcomes $r : I \to \mathbb{F}_2$, and decoding Clifford D. The recovery map of the protocol is

$$\mathcal{R}\left(\rho^{\otimes n}\right) = Tr_{2,\ldots,n}\left(D\Pi_I^r \rho^{\otimes n} \Pi_I^r D^\dagger\right). \tag{53}$$

Let $P_s(\epsilon) = Tr(\mathcal{R}(\rho_T(\epsilon)^{\otimes n}))$ be the success probability of the protocol starting with resource state $\rho_T(\epsilon)$. This gives us the first connection to the signed weight enumerator of the code (I, r), defined as

$$W_I^r(x,y) := \sum_{v\in I} (-1)^{r(v)} x^{n-swt(v)} y^{swt(v)}, \tag{54}$$

analogous to equation (4). Namely, the success probability can be written as

$$\begin{aligned} P_s(\epsilon) &= Tr\left(\Pi_I^r \rho_T(\epsilon)^{\otimes n}\right) && (55)\\ &= \frac{1}{|I|}\sum_{v\in I}(-1)^{r(v)} Tr\left(T_v \rho_T(\epsilon)^{\otimes n}\right) && (56)\\ &= \frac{1}{2^{n-1}}\sum_{v\in I}(-1)^{r(v)} t(\epsilon)^{swt(v)} && (57)\\ &= \frac{1}{2^{n-1}} W_I^r(1, t(\epsilon)). && (58) \end{aligned}$$

Here we get the second line by expanding the projector as in equation (14) and using linearity of the trace, and the third line uses the Pauli basis expansion of noisy T states, equation (52).

Next we compute the post-distillation state,

$$\rho' := \frac{\mathcal{R}\left(\rho_T(\epsilon)^{\otimes n}\right)}{P_s(\epsilon)}, \tag{59}$$

which, after twirling by $\mathcal{T}_T$ gives a state $\mathcal{T}_T(\rho') = \rho_T(\epsilon')$ with noise parameter ϵ'. First, define the signed coset weight enumerators of the code (I, r) as

$$W_{u+I}^r(x,y) = \sum_{v\in I}(-1)^{r(v)+\beta(u,v)} x^{n-swt(u+v)} y^{swt(u+v)}, \quad \forall u \in I^\perp \setminus I, \tag{60}$$

and the signed dual weight enumerator as

$$W_{I^\perp}^r(x,y) = \sum_{u\in I^\perp/I} W_{u+I}^r(x,y) \tag{61}$$

where the outer sum is over representatives of each coset. For any non-identity single-qubit Pauli operator T_u, $u \in \{x,y,z\}$, the corresponding logical Pauli operator for the code is $T_{\bar{u}} = D^\dagger(T_u \otimes \mathbb{1}_{2^{n-1}})D$. Since $\bar{u} \in I^\perp$, $T_{\bar{u}}$ commutes with the projector Π_I^r, and so the Pauli expectation value of the post-distillation state is

$$\begin{aligned} Tr(T_u \rho') &= \frac{1}{P_s(\epsilon)} Tr\left(T_u Tr_{2,\ldots,n}\left(D\Pi_I^r \rho_T(\epsilon)^{\otimes n} \Pi_I^r D^\dagger\right)\right) && (62)\\ &= \frac{1}{P_s(\epsilon)} Tr\left(T_{\bar{u}} \Pi_I^r \rho_T(\epsilon)^{\otimes n}\right) && (63)\\ &= \frac{1}{2^{n-1}\cdot P_s(\epsilon)}\sum_{v\in I}(-1)^{r(v)+\beta(\bar{u},v)} t(\epsilon)^{swt(\bar{u}+v)} && (64)\\ &= \frac{W_{\bar{u}+I}^r(1,t(\epsilon))}{W_I^r(1,t(\epsilon))}. && (65) \end{aligned}$$

After twirling the post-distillation state by $\mathcal{T}_T$, we get a depolarized T state $\rho_T(\epsilon')$ with

$$\begin{aligned} \frac{1-\epsilon'}{\sqrt{3}} &= \frac{1}{3}\frac{W_{\bar{x}+I}^r(1,t(\epsilon)) + W_{\bar{y}+I}^r(1,t(\epsilon)) + W_{\bar{z}+I}^r(1,t(\epsilon))}{W_I^r(1,t(\epsilon))} && (66)\\ &= \frac{1}{3}\frac{W_{I^\perp}^r(1,t(\epsilon)) - W_I^r(1,t(\epsilon))}{W_I^r(1,t(\epsilon))}, && (67) \end{aligned}$$

where the last equality follows from $I^\perp = I \sqcup (\bar{x}+I) \sqcup (\bar{y}+I) \sqcup (\bar{z}+I)$. Solving this for ϵ', we get the post-distillation noise ϵ' as a function of the pre-distillation noise ϵ,

$$\epsilon'(\epsilon) = 1 - \frac{1}{\sqrt{3}} \frac{W^r_{I^\perp}(1, t(\epsilon)) - W^r_I(1, t(\epsilon))}{W^r_I(1, t(\epsilon))}. \tag{68}$$

At this stage, we will reduce generality by assuming the code (I, r) is an M_3 code, i.e. it has a transversal M_3 gate. This means a logical M_3 gate is simply a physical M_3 gate applied to each qubit. A transversal M_3 gate heavily constrains the signs $r : I \to \mathbb{F}_2$. Concretely, a transversal M_3 gate means that conjugation by $M_3^{\otimes n}$ must preserve the stabilizer group $S_I^r = \{(-1)^{r(a)} T_a \mid a \in I\}$, so that for every stabilizer group element $P \in S_I^r$, we also have $M_3^{\otimes n} P (M_3^{\otimes n})^\dagger \in S_I^r$. This gives a constraint on the signs known as Rall's rule, the proof of which we reproduce here.

Lemma (Rall's rule [38, 39]). *Let (I, r) be an $[\![n, 1]\!]_2$ stabilizer code with a transversal M_3 gate. Then for every $u \in I$, the stabilizer group element $P_u := (-1)^{r(u)} T_u$ has even symplectic weight, and its sign $(-1)^{r(u)}$ is determined by*

$$(-1)^{r(u)} = \begin{cases} +1, & swt(u) \equiv 0 \pmod 4, \\ -1, & swt(u) \equiv 2 \pmod 4. \end{cases} \tag{69}$$

Equivalently, since $swt(u)$ is even, $(-1)^{r(u)} = (-1)^{swt(u)/2}$.

Proof of Rall's rule. Fix $u \in I$ and let $P_u = (-1)^{r(u)} T_u \in S_I^r$. Define its M_3 conjugates

$$P'_u := M_3^{\otimes n} P_u \left(M_3^{\otimes n}\right)^\dagger, \qquad \text{and} \qquad P''_u := \left(M_3^{\otimes n}\right)^\dagger P_u M_3^{\otimes n}. \tag{70}$$

Because I is an M_3 code, $P'_u, P''_u \in S_I^r$. Since S_I^r is a stabilizer group, it is abelian and closed under multiplication, so $P_u P'_u P''_u \in S_I^r$.

On the other hand, M_3 cyclically permutes single-qubit Paulis without introducing additional signs. Therefore, on any qubit where T_u acts nontrivially (as X, Y, or Z), the product of its three M_3-orbit elements is $X \cdot Y \cdot Z = i\mathbb{1}$, and on any qubit where T_u acts trivially the product is $\mathbb{1}$. Thus,

$$P_u P'_u P''_u = (-1)^{r(u)} i^{swt(u)} \mathbb{1}. \tag{71}$$

But a stabilizer group cannot contain $-\mathbb{1}$, only $\mathbb{1}$, and so $(-1)^{r(u)} i^{swt(u)} = 1$. This forces $i^{swt(u)} \in \{\pm 1\}$, so $swt(u)$ is even, and furthermore $(-1)^{r(u)} = +1$ if $swt(u) \equiv 0 \pmod 4$ and $(-1)^{r(u)} = -1$ if $swt(u) \equiv 2 \pmod 4$.

□

Rall's rule implies that the signed weight enumerator W_I^r is determined completely by the (unsigned) weight distribution of I, and in particular, we can see that if (I, r) is an M_3 code, then $A_w = 0$ for all odd w, and

$$W_I^r(x, y) = \sum_{a \in I} (-1)^{r(a)} x^{n - swt(a)} y^{swt(a)} = \sum_{w=0}^{n} A_w (-1)^{w/2} x^{n-w} y^w. \tag{72}$$

Equivalently, $W_I^r(x, y) = W_I(x, iy)$ [39], where the right-hand side is real-valued because only even weights occur. In particular, we can rewrite the success probability as

$$P_s(\epsilon) = \frac{1}{2^{n-1}} W_I^r(1, t(\epsilon)) = \frac{1}{2^{n-1}} \sum_{w=0}^{n} A_w (-1)^{w/2} t(\epsilon)^w = \frac{1}{2^{n-1}} W_I(1, it(\epsilon)). \tag{73}$$

With the assumption of a transversal M_3 gate, a similar simplification is possible for the post-distillation state ρ'. We can choose the logical Pauli operators $\bar{x}, \bar{y}, \bar{z} \in I^\perp \setminus I$ such that the corresponding logical Pauli operators are cyclically permuted by the transversal M_3 gate, i.e. $M_3^{\otimes n} T_{\bar{x}} (M_3^{\otimes n})^\dagger = T_{\bar{y}}$, $M_3^{\otimes n} T_{\bar{y}} (M_3^{\otimes n})^\dagger = T_{\bar{z}}$, and $M_3^{\otimes n} T_{\bar{z}} (M_3^{\otimes n})^\dagger = T_{\bar{x}}$. Then the Pauli expectation values of the post-distillation state are related by

$$Tr\left(T_y\rho'\right) = Tr\left(M_3 T_x M_3^\dagger \rho'\right) \tag{74}$$

$$= \frac{1}{P_s(\epsilon)} Tr\left(M_3^{\otimes n} T_{\bar{x}} M_3^{\dagger\otimes n} \Pi_I^r \rho_T(\epsilon)^{\otimes n}\right) \tag{75}$$

$$= \frac{1}{P_s(\epsilon)} Tr\left(T_{\bar{x}} \Pi_I^r \left(M_3^\dagger \rho_T(\epsilon) M_3\right)^{\otimes n}\right) \tag{76}$$

$$= \frac{1}{P_s(\epsilon)} Tr\left(T_{\bar{x}} \Pi_I^r \rho_T(\epsilon)^{\otimes n}\right) = Tr\left(T_x\rho'\right). \tag{77}$$

Here in the third line we use the cyclic property of the trace and the fact that since $M_3^{\otimes n}$ is a logical gate, it must preserve the code space, and so it commutes with the projector Π_I^r. Similarly, $Tr(T_z\rho') = Tr(T_y\rho')$. Therefore, no twirling is necessary as the post-distillation state ρ' is already on the line segment equation (50). Every element of $I^\perp \setminus I$ lies in exactly one of the three cosets $\bar{x}+I$, $\bar{y}+I$, or $\bar{z}+I$, and we get a rule similar to Rall's rule for the signs of the elements of these cosets.

Lemma ([39]). *Let (I,r) be an $[\![n,1]\!]_2$ M_3 code, and choose $\bar{x},\bar{y},\bar{z} \in I^\perp \setminus I$ as above. For any $u \in I^\perp \setminus I$, define the signed Pauli operator P_u by writing $u = \bar{v} + w$ with $\bar{v} \in \{\bar{x},\bar{y},\bar{z}\}$ and $w \in I$, and setting*

$$P_u := (-1)^{r(w)+\beta(\bar{v},w)} T_u. \tag{78}$$

Then every $u \in I^\perp \setminus I$ has odd symplectic weight, and moreover, there exists a sign $\lambda \in \{\pm 1\}$, independent of u, such that

$$(-1)^{r(w)+\beta(\bar{v},w)} = \lambda (-1)^{(swt(u)-1)/2}. \tag{79}$$

Proof of lemma 2. Fix a $u \in I^\perp \setminus I$. We write $u = \bar{v} + w$ with $\bar{v} \in \{\bar{x},\bar{y},\bar{z}\}$ and $w \in I$ and define the corresponding signed Pauli operators $P_u = (-1)^{r(w)+\beta(\bar{v},w)} T_u$. Define its M_3 conjugates as

$$P_u' = M_3^{\otimes n} P_u \left(M_3^{\otimes n}\right)^\dagger \quad \text{and} \quad P_u'' = \left(M_3^{\otimes n}\right)^\dagger P_u M_3^{\otimes n}. \tag{80}$$

These three operators act on the code space as the three nontrivial logical Paulis in the order $\bar{X}\bar{Y}\bar{Z}$. Therefore, on the code space their product is

$$P_u P_u' P_u'' \Pi_I^r = \lambda \mathrm{i} \Pi_I^r \tag{81}$$

where $\lambda \in \{\pm 1\}$ is a fixed constant depending on the sign convention equation (10) for the logical Pauli operators.

On the other hand, M_3 cyclically permutes the single-qubit Paulis without introducing extra signs, so on each qubit where T_u acts nontrivially, the product $P_u P_u' P_u''$ contributes a factor of $\mathrm{i}\mathbb{1}$, and on each qubit where T_u acts trivially, the contribution is $\mathbb{1}$. Therefore,

$$P_u P_u' P_u'' = (-1)^{r(w)+\beta(\bar{v},w)} i^{swt(u)} \mathbb{1}. \tag{82}$$

Multiplying this expression by Π_I^r and comparing against the expression above, we see that this forces $swt(u)$ to be odd, and

$$(-1)^{r(w)+\beta(\bar{v},w)} = \lambda (-1)^{(swt(u)-1)/2}, \tag{83}$$

which proves the claim. □

Rall's rule shows that every word in I has even symplectic weight, while Lemma 2 shows that every word in $I^\perp \setminus I$ has odd symplectic weight. Therefore, if $\{B_w \mid w = 0,\ldots,n\}$ is the symplectic weight distribution of $I^\perp$, i.e.

$$W_{I^\perp}(x,y) = \sum_{w=0}^{n} B_w x^{n-w} y^w, \tag{84}$$

then $B_w = A_w$ for even w, and the odd-weight terms come entirely from $I^\perp \setminus I$. Using Lemma 2 for the numerator of equation (68), we get

$$W_{I^\perp}^r(1,t(\epsilon)) - W_I^r(1,t(\epsilon)) = \lambda \sum_{\substack{w=0 \\ w\text{odd}}}^{n} B_w (-1)^{(w-1)/2} t(\epsilon)^w = -i\lambda \left(W_{I^\perp}(1,it(\epsilon)) - W_I(1,it(\epsilon))\right). \tag{85}$$

This gives us three equivalent simplified expression for the post-distillation noise parameter,

$$\epsilon'(\epsilon) = 1 + \frac{\mathrm{i}\lambda}{\sqrt{3}} \frac{W_{I^\perp}(1, it(\epsilon)) - W_I(1, it(\epsilon))}{W_I(1, it(\epsilon))} \tag{86}$$

$$= 1 - \frac{\lambda}{\sqrt{3}} \frac{\sum_{w\mathrm{odd}} B_w (-1)^{(w-1)/2} t(\epsilon)^w}{\sum_{w\mathrm{even}} A_w (-1)^{w/2} t(\epsilon)^w} \tag{87}$$

$$= 1 - \frac{\lambda}{\sqrt{3}} \frac{\mathrm{Im}\, W_{I^\perp}(1, it(\epsilon))}{W_I(1, it(\epsilon))}. \tag{88}$$

Thus, up to the choice of $\lambda \in \{\pm 1\}$ (which essentially amounts to choosing $|T\rangle$ or $|T^\perp\rangle$ as your target magic state [39]), for M_3 codes the distillation map $\epsilon \mapsto \epsilon'(\epsilon)$ and the success probability $P_s(\epsilon)$ are determined entirely by the weight distributions A_w and B_w, or equivalently, by the simple (unsigned) weight enumerators. See [38, 39] for a more thorough derivation of these expressions for $\epsilon'(\epsilon)$ and $P_s(\epsilon)$.

4.2.2. Distilling T states with quantum QR codes

Throughout this section, let C be a $[p, (p-1)/2]_4$ Hermitian self-orthogonal expurgated QR code, and let $I = \iota^{-1}(C)$ be the corresponding $[\![p, 1]\!]_2$ quantum QR code.

We want to apply the results of the previous section to analyzing the performance of quantum QR codes for distilling T states. Any stabilizer code on n qubits corresponding to an isotropic subspace of dimension less than n that is $\mathbb{F}_4$-linear under the map ι is compatible with a transversal M_3 gate [58]. Since the expurgated QR code C is linear, this applies to the corresponding quantum QR code I. That is, there exists a choice of signs $r : I \to \mathbb{F}_2$ such that the code (I, r) has a transversal M_3 gate. This choice of signs allows us to use the simplified expressions equation (73) and equation (86) for the success probability $P_s(\epsilon)$ of the protocol and the post-distillation noise parameter $\epsilon'(\epsilon)$.

For a quantum QR code I, the correspondence of section 2.3 identifies $I \subset \mathbb{F}_2^n \times \mathbb{F}_2^n$ with an expurgated QR code $C \subset \mathbb{F}_4^n$, and with lemma 1, the symplectic complement $I^\perp$ is identified with the Hermitian dual $C^{\perp_\mathrm{H}}$. Under this identification, W_I and $W_{I^\perp}$ are exactly the Hamming weight enumerators W_C and $W_{C^{\perp_\mathrm{H}}}$, the latter of which can be obtained from W_C by the MacWilliams identity equation (5).

All that remains is to compute the weight enumerators of the expurgated QR codes. A direct enumeration of the weights of vectors in C is infeasible for the lengths of interest since the cardinality of a code, $|C| = 2^{p-1}$, is exponential in the code length. In order to compute the Hamming weight enumerators of expurgated QR codes over $\mathbb{F}_4$, we use Prange's theorem and classical methods [59].

The expurgated QR code C is obtained by shortening the extended QR code $\tilde{C}$ of length $n = p + 1$ on any coordinate. Therefore, W_C can be obtained from $W_{\tilde{C}}$ via Prange's theorem, equation (34). Since $C^{\perp_\mathrm{H}}$ is the augmented QR code of length p (see the proof of theorem 2), $C^{\perp_\mathrm{H}}$ is obtained by puncturing $\tilde{C}$ on any coordinate, the weight enumerator $W_{C^{\perp_\mathrm{H}}}$ can also be obtained from $W_{\tilde{C}}$ by Prange's theorem. Using the partial derivative form of Prange's theorem, equations (35) and (36), we get an even simpler expression for $\epsilon'(\epsilon)$, namely,

$$\epsilon'(\epsilon) = 1 + \frac{\mathrm{i}\lambda}{\sqrt{3}} \left. \frac{\frac{\partial}{\partial y} W_{\tilde{C}}(x, y)}{\frac{\partial}{\partial x} W_{\tilde{C}}(x, y)} \right|_{(x,y)=(1, it(\epsilon))}. \tag{89}$$

Then we only need to determine $W_{\tilde{C}}$. The code $\tilde{C}$ is Hermitian self-dual, and through invariant theory, we know that the Hamming weight enumerators of $\mathbb{F}_4$-linear Hermitian self-dual codes live in the ring $\mathbb{C}[i_2, h_6]$ where

$$\begin{cases} i_2(x, y) = x^2 + 3y^2 \\ h_6(x, y) = x^6 + 45x^2y^4 + 18y^6 \end{cases} \tag{90}$$

are the Hamming weight enumerators of the $[2, 1, 2]_4$ repetition code and the $[6, 3, 4]_4$ hexacode respectively [59, §7.3.1]. With $n = p + 1$, since $W_{\tilde{C}}(x, y)$ is homogeneous of degree n, we can write

$$W_{\tilde{C}}(x, y) = \sum_{j=0}^{\lfloor n/6 \rfloor} a_j i_2(x, y)^{\frac{n}{2} - 3j} h_6(x, y)^j. \tag{91}$$

Table 2. Noise thresholds of magic state distillation protocols for qubit T states based on quantum quadratic residue codes.

Code	Threshold
$[5,2,4]_4$	≈0.34535
$[23,11,8]_4$	≈0.32237
$[29,14,12]_4$	≈0.24190
$[47,23,12]_4$	≈0.05050
$[53,26,16]_4$	≈0.27343
$[71,35,12]_4$	≈0.00664

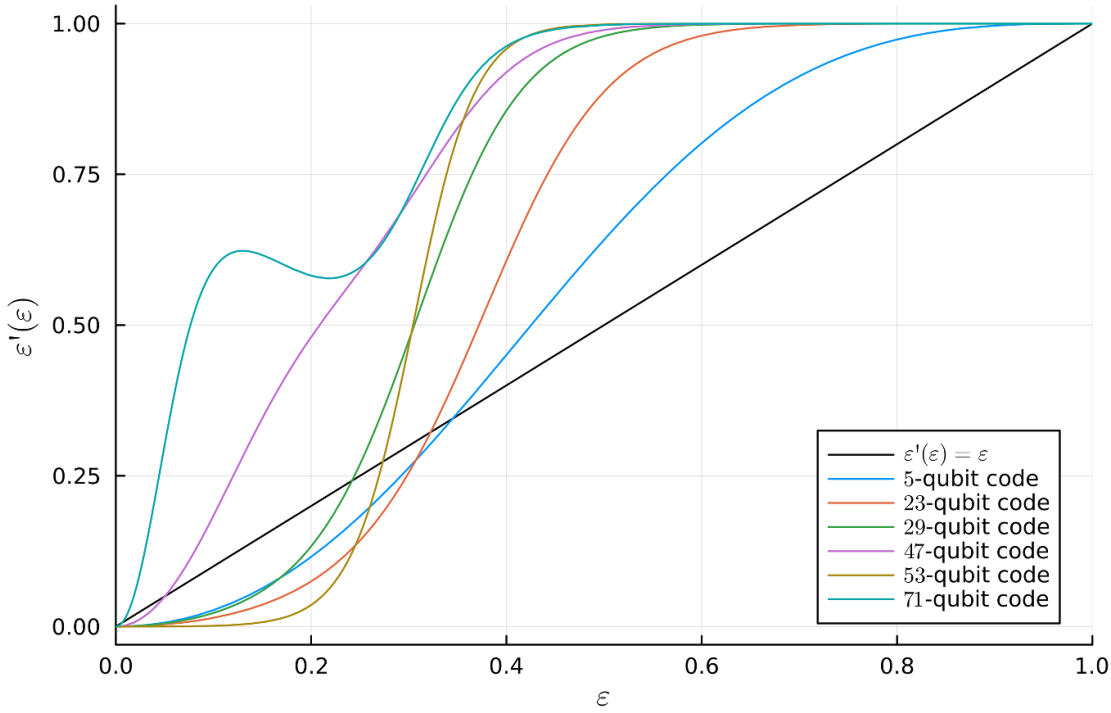


Figure 1. Post-distillation noise ϵ' as a function of pre-distillation noise ϵ for quantum quadratic residue codes that distill qubit T states. The 5-, 23-, 29-, 47-, 53-, and 71-qubit codes have non-trivial distillation thresholds.

There are $\lfloor n/6 \rfloor + 1$ coefficients $\{a_j\}$ in equation (91). There is exactly one vector in $\tilde{C}$ of weight 0, namely, the zero vector, and so $\tilde{A}_0 = 1$. This gives one linear constraint on the coefficients $\{a_j\}$, leaving $\lfloor n/6 \rfloor$ independent parameters. Computing $\lfloor n/6 \rfloor$ additional weight multiplicities, for example $\tilde{A}_2, \tilde{A}_4, \ldots, \tilde{A}_{2\lfloor n/6 \rfloor}$, determines $W_{\tilde{C}}$ uniquely by solving a linear system for the coefficients $\{a_j\}$.

To obtain these low-weight multiplicities $\tilde{A}_w$ without enumerating all codewords of $\tilde{C}$, we use the algorithm of [60], which computes the number of codewords of specified small weights from a complementary pair of generator matrices[4]. Using this strategy, we computed the weight enumerators of all quantum QR codes on qubits up to length 71. The weight enumerators are listed in table 4 in appendix C.

Using these weight enumerators, and the relation between weight enumerators and the performance of M_3 codes for distilling T states described in section 4.2.1, we find that the qubit quantum QR codes of length 5, 23, 29, 47, 53, and 71 distill T states with nontrivial thresholds. These thresholds are listed in table 2. The distillation functions $\epsilon'(\epsilon)$ for these codes are shown in figure 1, and their success probabilities $P_s(\epsilon)$ are shown in figure 2. An additional plot showing the functions $\epsilon'(\epsilon)$ for all qubit quantum QR codes up to length 71, not just those with non-trivial thresholds, is shown in figure 5 in appendix D.

[4] We have a Julia implementation of this algorithm that runs on Nvidia GPUs available at github.com/mzurel/WeightDistributionsGPU.jl.

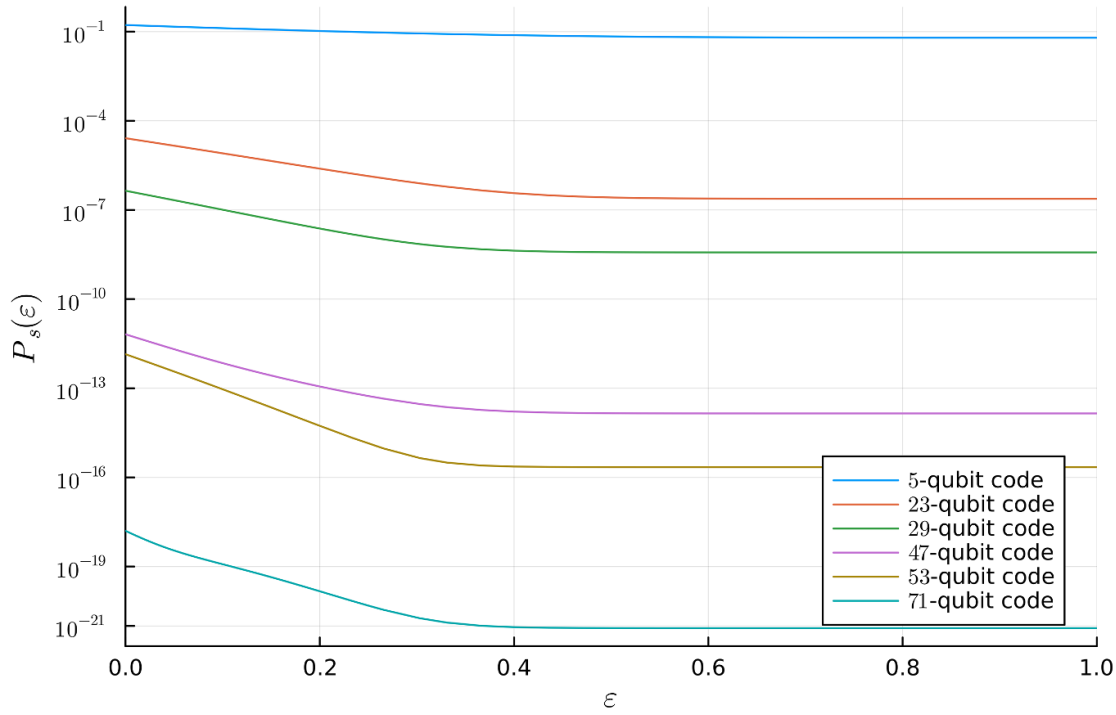


Figure 2. Success probability of magic state distillation protocols for qubit T states based on quantum quadratic residue codes.

The noise threshold is not the only figure of merit for a magic state distillation protocol. In appendix E, we compare the resource consumption of a single round of these protocols based on quantum QR codes against the baseline of protocols formed by concatenating the 5-qubit protocol, both in terms of the number of T states consumed by a successful fully post-selected run of the protocol and the expected number of raw input T states consumed taking into account the success probability of the protocols.

4.2.3. Infinitely many quantum QR codes distill T states

Table 2 shows that the qubit quantum QR codes of length 5, 23, 29, 47, 53, and 71 distill T states with non-trivial thresholds. Here we show that there are infinitely many code lengths with non-trivial threshold for T state distillation. This is the result of the following theorem, and corollary.

Theorem 4. *Let p be a prime with $p \equiv 5$ or $23 \mod 24$. Let $C \subset \mathbb{F}_4^p$ be the expurgated QR code of length p and let $I = \iota^{-1}(C)$ be the corresponding quantum QR code. Then there is a sign λ in equation (86) such that the map $\epsilon'(\epsilon)$ for the corresponding T state distillation protocol satisfies*

$$\epsilon'(\epsilon) = O\left(\epsilon^2\right), \quad (\epsilon \to 0). \tag{92}$$

In other words, the quantum QR code of length p distills T states with a non-trivial threshold. Furthermore, these protocols have a positive success probability for sufficiently small ϵ.

Proof of theorem 4. The congruence $p \equiv 5 \text{or } 23 \mod 24$ implies $p \equiv 5 \text{or } 7 \mod 8$, and so a quantum QR code of this length exists by corollary 1. It also implies that $p \equiv 2 \mod 3$.

Let $\tilde{C} \subset \mathbb{F}_4^{p+1}$ be the extended QR code. The expurgated QR code C is obtained from $\tilde{C}$ by shortening on any coordinate, and the corresponding quantum QR code is obtained via $\iota^{-1}(C)$. We use the expression equation (89) for the post-distillation noise parameter $\epsilon'(\epsilon)$ of the distillation protocol based on the code C. This expression involves evaluating partial derivatives of $W_{\tilde{C}}(x,y)$ at $(x,y) = (1, it(\epsilon))$ where $t(\epsilon) = (1-\epsilon)/\sqrt{3}$.

As stated above, $\tilde{C}$ is Hermitian self-dual with Hamming weight enumerator in the ring $\mathbb{C}[i_2, h_6]$. Let $n = p+1 = 6m$ (so $m = (p+1)/6 \in \mathbb{Z}$). Since $\deg(i_2) = 2$ and $\deg(h_6) = 6$, the weight enumerator of $\tilde{C}$ must have the form

$$W_{\tilde{C}}(x,y) = \sum_{j=0}^{m} a_j i_2(x,y)^{3j} h_6(x,y)^{m-j} = a_0 h_6(x,y)^m + i_2(x,y)^3 R(x,y) \tag{93}$$

for some coefficients $a_j \in \mathbb{C}$ and some homogeneous polynomial $R \in \mathbb{C}[i_2, h_6]$.

Let $u(\epsilon) := i_2(1, it(\epsilon)) = 1 - 3t^2 = 2\epsilon - \epsilon^2$ and $v(\epsilon) := h_6(1, it(\epsilon))$. Differentiating equation (93) for $W_{\tilde{C}}(x,y)$, we have

$$\begin{cases} \left.\frac{\partial W_{\tilde{C}}(x,y)}{\partial x}\right|_{(x,y)=(1,it(\epsilon))} = a_0 m v(\epsilon)^{m-1} \left.\frac{\partial h_6(x,y)}{\partial x}\right|_{(x,y)=(1,it(\epsilon))} + O\left(u(\epsilon)^2\right), \\ \left.\frac{\partial W_{\tilde{C}}(x,y)}{\partial y}\right|_{(x,y)=(1,it(\epsilon))} = a_0 m v(\epsilon)^{m-1} \left.\frac{\partial h_6(x,y)}{\partial y}\right|_{(x,y)=(1,it(\epsilon))} + O\left(u(\epsilon)^2\right). \end{cases} \tag{94}$$

Here the error terms are $O(u^2)$ because every term coming from differentiating $i_2^3 R$ carries at least a factor i_2^2. Therefore, for $\epsilon \ll 1$,

$$\left.\frac{\frac{\partial}{\partial y} W_{\tilde{C}}(x,y)}{\frac{\partial}{\partial x} W_{\tilde{C}}(x,y)}\right|_{(x,y)=(1,it(\epsilon))} = \left.\frac{\frac{\partial}{\partial y} h_6(x,y)}{\frac{\partial}{\partial x} h_6(x,y)}\right|_{(x,y)=(1,it(\epsilon))} + O\left(u(\epsilon)^2\right). \tag{95}$$

Computing these partial derivatives, we find

$$\Phi(t) := \left.\frac{\frac{\partial}{\partial y} h_6(x,y)}{\frac{\partial}{\partial x} h_6(x,y)}\right|_{(x,y)=(1,it(\epsilon))} = \mathrm{i}\frac{-30t^3 + 18t^5}{1 + 15t^4}. \tag{96}$$

At $\epsilon = 0$, we have $t = 1/\sqrt{3}$ and so $\Phi(1/\sqrt{3}) = -\mathrm{i}\sqrt{3}$. Furthermore,

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = \mathrm{i}\frac{90t^2\left(t^2+1\right)^2\left(3t^2-1\right)}{\left(1+15t^4\right)^2}, \tag{97}$$

so at $t = 1/\sqrt{3}$, $\mathrm{d}\Phi/\mathrm{d}t = 0$. Since $t - 1/\sqrt{3} = -\epsilon/\sqrt{3}$, this implies $\Phi(t) = \Phi(1/\sqrt{3}) + O(\epsilon^2) = -\mathrm{i}\sqrt{3} + O(\epsilon^2)$, and therefore

$$\left.\frac{\frac{\partial}{\partial y} W_{\tilde{C}}(x,y)}{\frac{\partial}{\partial x} W_{\tilde{C}}(x,y)}\right|_{(x,y)=(1,it(\epsilon))} = -\mathrm{i}\sqrt{3} + O\left(\epsilon^2\right). \tag{98}$$

Finally, we plug this into equation (89). With the choice of λ for which $\epsilon'(0) = 0$, the constant term cancels and we obtain $\epsilon'(\epsilon) = O(\epsilon^2)$. Thus, $\epsilon'(\epsilon)/\epsilon \to 0$ as $\epsilon \to 0$, so there exists $\epsilon_{\mathrm{th}}(p) > 0$ such that $\epsilon'(\epsilon) < \epsilon$ for all $\epsilon \in (0, \epsilon_{\mathrm{th}}(p))$, i.e. a non-trivial threshold.

The last thing to show is that these protocols maintain a nonzero success probability for sufficiently small ϵ. As shown above, the success probability is proportional to $W_C(1, it(\epsilon))$. By Prange's theorem, we have

$$W_C(x,y) = \frac{1}{n}\frac{\partial}{\partial x} W_{\tilde{C}}(x,y) \tag{99}$$

and so it suffices to show that $\partial W_{\tilde{C}}(x,y)/\partial x$ evaluated at $(x,y) = (1, it(\epsilon))$ is nonzero in the limit $\epsilon \to 0$, or equivalently, that

$$D := \left.\frac{\partial W_{\tilde{C}}(x,y)}{\partial x}\right|_{(x,y)=\left(1,i/\sqrt{3}\right)} \neq 0. \tag{100}$$

Note that this also justifies the expansion equation (95) by showing that the denominator is nonzero in the small ϵ limit.

Since $\tilde{C}$ is Hermitian self-dual, every code word of $\tilde{C}$ has even Hamming weight [42]. Thus, evaluating at $(x,y) = (1, \mathrm{i}/\sqrt{3})$ introduces powers of $-1/3$ for terms with nonzero powers of the y coordinate, that is, $D \in \mathbb{Z}[1/3]$. Then since $p \neq 3$, reduction modulo p is well-defined in $\mathbb{Z}[1/3]$, and so it suffices to show that $D \not\equiv 0 \pmod{p}$.

Let σ be the cyclic shift of the p coordinates of $\tilde{C}$ that are shared with C which fixes the extension coordinate. This is an automorphism of $\tilde{C}$ since QR codes are cyclic codes. Then, since p is prime, every orbit of codewords of $\tilde{C}$ under the action of the group generated by σ has size 1 or p, and every word in an orbit has the same weight. Therefore, orbits of code words that are not fixed by σ contribute to the coefficients of $W_{\tilde{C}}(x,y)$ in multiples of p, and thus, $W_{\tilde{C}}(x,y)$ is congruent modulo p to the weight enumerator $W_{\tilde{C}^\sigma}(x,y)$ of the subcode $\tilde{C}^\sigma \subset \tilde{C}$ formed from code words of $\tilde{C}$ that are fixed by σ. Differentiating preserves congruences modulo p, so we also have

$$\frac{\partial W_{\tilde{C}}(x,y)}{\partial x} \equiv \frac{\partial W_{\tilde{C}^\sigma}(x,y)}{\partial x} \pmod{p}, \tag{101}$$

and so it suffices to show that the partial derivative of $W_{\tilde{C}^\sigma}(x,y)$ is nonzero at $(x,y) = (1, \mathrm{i}/\sqrt{3})$.

Code words of $\tilde{C}$ that are fixed by σ have the form $(a,\ldots,a,b)$ with $a,b \in \mathbb{F}_4$. Again, since the code words of $\tilde{C}$ all have even Hamming weight, there are no words of weight 1 or p, and so the code words of $\tilde{C}$ that are fixed by σ all have weight 0 or $n = p+1$. Thus,

$$\frac{\partial}{\partial x} W_{\tilde{C}^\sigma} = n x^{n-1}. \tag{102}$$

Evaluating this at $x = 1$ gives

$$\frac{\partial}{\partial x} W_{\tilde{C}}(x,y)\Big|_{(x,y)=\left(1,\mathrm{i}/\sqrt{3}\right)} \equiv n \equiv 1 \pmod{p}. \tag{103}$$

Therefore, the success probability of these protocols is nonzero in the small ϵ limit, and by continuity, it remains nonzero for all sufficiently small $\epsilon > 0$.

□

Corollary 6. *There are infinitely many quantum QR codes that distill T states with non-trivial threshold.*

Proof of corollary 6. There exists a quantum QR code for every prime length $p \equiv 5 \text{or } 7 \mod 8$. By Theorem 4, they are guaranteed to distill T states whenever $p \equiv 5 \text{ or } 23 \mod 24$. By Dirichlet's theorem, there are infinitely many prime lengths p satisfying this condition. □

4.3. Qutrit Strange state distillation

As evidenced by the 11-qutrit Golay code example given in section 4.1, some quantum QR codes are useful for distilling qutrit Strange states [16, 21, 61]. The Strange state is the single-qutrit state $|S\rangle := (|1\rangle - |2\rangle)/\sqrt{2}$. Prakash and Singhal showed that the performance of a magic state distillation protocol for distilling Strange states is determined by the weight distribution of the corresponding code, and relies on a relation between the weight enumerator of the code and the Wigner function of the state [19]. For completeness, we include a derivation of this fact here.

4.3.1. Wigner functions and weight enumerators

The phase space point operators of the Wigner function [30, 62] on n d-dimensional qudits are

$$A_u = \frac{1}{d^n} \sum_{v \in \mathbb{F}_d^{2n}} \omega^{[u,v]} T_v, \qquad \forall u \in \mathbb{F}_d^{2n}. \tag{104}$$

These operators are Hermitian, have unit trace, and together they form an orthogonal basis for the real vector space of Hermitian operators on the n-qudit Hilbert space satisfying the orthogonality relation $Tr(A_u A_v) = d^n \delta_{u,v}$. The Wigner function of a state ρ is given by the coefficients of ρ when expanded in this basis,

$$\rho = \sum_{u \in \mathbb{F}_d^{2n}} W_\rho(u) A_u, \tag{105}$$

which, by orthogonality, can be extracted by $W_\rho(u) = Tr(\rho A_u)/d^n$. Since the phase space point operators are Hermitian, the Wigner function is real, and taking a trace of equation (105), we get the normalization condition, $1 = \sum_{u \in \mathbb{F}_d^{2n}} W_\rho(u)$. A summary of how the Wigner function transforms under the stabilizer operations is given in appendix B.

The Wigner function of the Strange state has a particularly simple form. As a density matrix, the pure Strange state is

$$|S\rangle\langle S| = \frac{1}{3}\mathbb{1} - \frac{1}{2}\left(A_0 - \frac{1}{3}\mathbb{1}\right) = -\frac{1}{3}A_0 + \frac{1}{6}\sum_{u \in \mathbb{F}_3^2 \setminus \{0\}} A_u, \tag{106}$$

and therefore has Wigner function $W_{|S\rangle\langle S|}(0) = -1/3$ and $W_{|S\rangle\langle S|}(u) = 1/6$ for all $u \neq 0 \in \mathbb{F}_3^2$. When subject to depolarizing noise, the noisy Strange state,

$$\rho_S(\epsilon) := (1-\epsilon)|S\rangle\langle S| + \frac{\epsilon}{3}\mathbb{1}, \tag{107}$$

has Wigner function

$$W_{\rho_S(\epsilon)}(u) = \begin{cases} \frac{4\epsilon-3}{9} & \text{if} u = 0, \\ \frac{3-\epsilon}{18} & \text{if} u \neq 0. \end{cases} \tag{108}$$

We will assume that the resource states for magic state distillation are depolarized Strange states of the form equation (107). No generality is lost in making this assumption since we can always apply a group twirl with respect to the subgroup of the Clifford group that fixes A_0 to get a state that lies on the line running through $|S\rangle\langle S|$ and the maximally mixed state. Concretely, let

$$G_0 = \left\{ g \in \mathcal{C}\ell \mid gA_0g^\dagger = A_0 \right\} \cong \mathrm{Sp}\left(\mathbb{F}_3^2\right). \tag{109}$$

The group twirl

$$\mathcal{T}_S(\rho) = \frac{1}{|G_0|} \sum_{g \in G_0} g\rho g^\dagger \tag{110}$$

is a stabilizer operation that does not change the value of the Wigner function at the origin of phase space, $W_{\mathcal{T}_S(\rho)}(0) = W_\rho(0)$, but makes the Wigner function $W_{\mathcal{T}_S(\rho)}(u)$ uniform on all other phase space points. With the normalization of the Wigner function, $\sum_u W_\rho(u) = 1$, the Wigner function of any state after the twirl is determined uniquely by the value of the Wigner function at the origin, or equivalently, by the noise parameter ϵ in equation (107).

Let $x(\epsilon) = (4\epsilon - 3)/9$ and $y(\epsilon) = (3-\epsilon)/18$ be the values of the Wigner function of the state $\rho_S(\epsilon)$ in equation (108). Then the initial state used for magic state distillation, an n-fold tensor product of noisy Strange states, can be written as

$$\rho_S(\epsilon)^{\otimes n} = \left(x(\epsilon) A_0 + y(\epsilon) \sum_{u \in \mathbb{F}_3^2 \setminus \{0\}} A_u \right)^{\otimes n} = \sum_{u \in \mathbb{F}_3^{2n}} x(\epsilon)^{n - swt(u)} y(\epsilon)^{swt(u)} A_u, \tag{111}$$

and so the Wigner function of this state is

$$W_{\rho_S(\epsilon)^{\otimes n}}(u) = x(\epsilon)^{n-swt(u)} y(\epsilon)^{swt(u)}. \tag{112}$$

Fix an n-to-1 qutrit distillation protocol defined by an $[\![n, 1]\!]_3$ stabilizer code $I \subset \mathbb{F}_3^{2n}$, post-selection on measurement outcomes $r : I \to \mathbb{F}_3$, and decoding Clifford D. The recovery map of the protocol is

$$\mathcal{R}(A) := Tr_{2,3,\ldots,n}\left(D\Pi_I^r A \Pi_I^r D^\dagger \right). \tag{113}$$

When applied to a resource state ρ, the success probability of the protocol is given by $Tr(\mathcal{R}(\rho^{\otimes n}))$, and the post-distillation state looks like

$$\rho' = \frac{\mathcal{R}(\rho^{\otimes n})}{Tr(\mathcal{R}(\rho^{\otimes n}))} = \frac{Tr_{2,3,\ldots,n}\left(D\Pi_I^r \rho^{\otimes n} \Pi_I^r D^\dagger \right)}{Tr(D\Pi_I^r \rho^{\otimes n} \Pi_I^r D^\dagger)}. \tag{114}$$

Post-selecting on measurement outcomes $r(u) = 0$ for all $u \in I$, with Lemma 3 in appendix B, we have

$$\Pi_I^0 A_v \Pi_I^0 = \delta_{v \in I^\perp} \frac{1}{|I|} \sum_{w \in v + I} A_w, \tag{115}$$

and so using the expansion $\rho^{\otimes n} = \sum_{v \in \mathbb{F}_3^{2n}} W_{\rho^{\otimes n}}(v) A_v$, the subnormalized post-measurement state $\Pi_I^0 \rho^{\otimes n} \Pi_I^0$ has Wigner function

$$W_{\Pi_I^0 \rho^{\otimes n} \Pi_I^0}(u) = \begin{cases} \frac{1}{|I|} \sum_{v \in u + I} W_{\rho^{\otimes n}}(v) & \text{if} u \in I^\perp, \\ 0 & \text{otherwise.} \end{cases} \tag{116}$$

The decoding D identifies the cosets $I^\perp / I \cong \mathbb{F}_3^2$ with the phase space of the first physical qutrit, and so the Wigner function of the normalized post-distillation state is

$$W_{\rho'}(u) = \frac{\sum_{v \in \bar{u} + I} W_{\rho^{\otimes n}}(v)}{\sum_{u' \in \mathbb{F}_3^2} \sum_{v \in \bar{u}' + I} W_{\rho^{\otimes n}}(v)} = \frac{\sum_{v \in \bar{u} + I} W_{\rho^{\otimes n}}(v)}{\sum_{v \in I^\perp} W_{\rho^{\otimes n}}(v)}, \tag{117}$$

where $\bar{u} \in I^{\perp}$ is a representative of the coset of $u \in \mathbb{F}_3^2$, and the denominator is determined by imposing normalization of the Wigner function, $\sum_{u\in\mathbb{F}_3^2} W_{\rho'}(u) = 1$.

If we twirl by $\mathcal{T}_S$ after the protocol, the result is a depolarized Strange state $\rho_S(\epsilon')$ of the form equation (107) with noise parameter ϵ'. The value of ϵ' is determined by the value of the Wigner function at phase space point $0 \in \mathbb{F}_3^2$, via the relation $W_{\rho_S(\epsilon')}(0) =: x(\epsilon') = (4\epsilon' - 3)/9$. Solving this for ϵ', we get the expression

$$\epsilon' = \frac{3}{4}\left(1 + 3W_{\rho_S(\epsilon')}(0)\right) = \frac{3}{4}\left(1 + 3\frac{\sum_{v\in I} W_{\rho^{\otimes n}}(v)}{\sum_{v\in I^{\perp}} W_{\rho^{\otimes n}}(v)}\right). \tag{118}$$

If the resource state is $\rho = \rho_S(\epsilon)$, then the numerator in this expression is exactly the symplectic weight enumerator polynomial $W_I(x,y)$ of the code I evaluated at $(x,y) = (x(\epsilon), y(\epsilon))$, and the denominator is the symplectic weight enumerator polynomial $W_{I^{\perp}}(x,y)$ of $I^{\perp}$ evaluated at the same point. That is, we get the expression for the post-distillation noise ϵ' as a function of the pre-distillation noise ϵ,

$$\epsilon'(\epsilon) = \frac{3}{4}\left(1 + 3\frac{W_I(x(\epsilon), y(\epsilon))}{W_{I^{\perp}}(x(\epsilon), y(\epsilon))}\right). \tag{119}$$

Using the MacWilliams identity to write the symplectic weight enumerator of $I^{\perp}$ in terms of that of I, we get the final expression

$$\epsilon'(\epsilon) = \frac{3}{4}\left(1 + 3^n \frac{W_I\left(\frac{4\epsilon-3}{9}, \frac{3-\epsilon}{18}\right)}{W_I\left(1, \frac{\epsilon-1}{2}\right)}\right), \tag{120}$$

where we use the identities $x(\epsilon) + 8y(\epsilon) = 1$ (from normalization of the Wigner function), and $x(\epsilon) - y(\epsilon) = (\epsilon - 1)/2$.

The success probability of the protocol applied to resource states $\rho_S(\epsilon)$ is

$$P_s(\epsilon) = Tr\left(\mathcal{R}\left(\rho^{\otimes n}\right)\right) = \sum_{v\in I^{\perp}} W_{\rho_S(\epsilon)^{\otimes n}}(v) = W_{I^{\perp}}(x(\epsilon), y(\epsilon)). \tag{121}$$

See [19] for a more complete derivation of these expressions for $\epsilon'(\epsilon)$ and $P_s(\epsilon)$.

4.3.2. Distilling Strange states with quantum QR codes

The result of the previous section is that the performance of a magic state distillation protocol for distilling Strange states is determined entirely by the symplectic weight enumerator of the stabilizer code I. Via the map ι of section 2.3, this weight enumerator is exactly the Hamming weight enumerator of the corresponding $\mathbb{F}_9$ code $\iota(I)$. Here we will use these results to analyze the performance of quantum QR codes for distilling Strange states. Throughout this section, let C be a $[p, (p-1)/2]_9$ Hermitian self-orthogonal expurgated QR code, and let $I = \iota^{-1}(C)$ be the corresponding $[\![p, 1]\!]_3$ quantum QR code.

Our strategy for computing these weight enumerators will be the same as in section 4.2.2. We use Prange's theorem to relate the weight enumerator of C to that of $\tilde{C}$, the length $n = p + 1$ extended QR code via equation (34). From invariant theory, the Hamming weight enumerators of Hermitian self-dual codes over $\mathbb{F}_9$ live in the ring

$$\mathbb{C}[r_2, r_4] \oplus r_6\mathbb{C}[r_2, r_4] \tag{122}$$

with syzygy

$$r_6^2 = \frac{3}{4}r_2^4 r_4 - \frac{3}{2}r_2^2 r_4^2 - \frac{1}{4}r_4^3 - r_2^3 r_6 + 3r_2 r_4 r_6 \tag{123}$$

where

$$\begin{cases} r_2 = x^2 + 8y^2 \\ r_4 = x^4 + 32xy^3 + 48y^4 \\ r_6 = x^6 + 16x^3y^3 + 72x^2y^4 + 288xy^5 + 352y^6 \end{cases} \tag{124}$$

are the weight enumerators of certain length 2, 4, and 6 codes [59, §5.8]. Therefore, the weight enumerator $W_{\tilde{C}}(x,y)$ can be written as

$$W_{\tilde{C}}(x,y) = \sum_{j=0}^{\lfloor n/4 \rfloor} a_j r_2(x,y)^{\frac{n}{2}-2j} r_4(x,y)^j + r_6(x,y) \sum_{j=0}^{\lfloor (n-6)/4 \rfloor} b_j r_2(x,y)^{\frac{n-6}{2}-2j} r_4(x,y)^j. \tag{125}$$

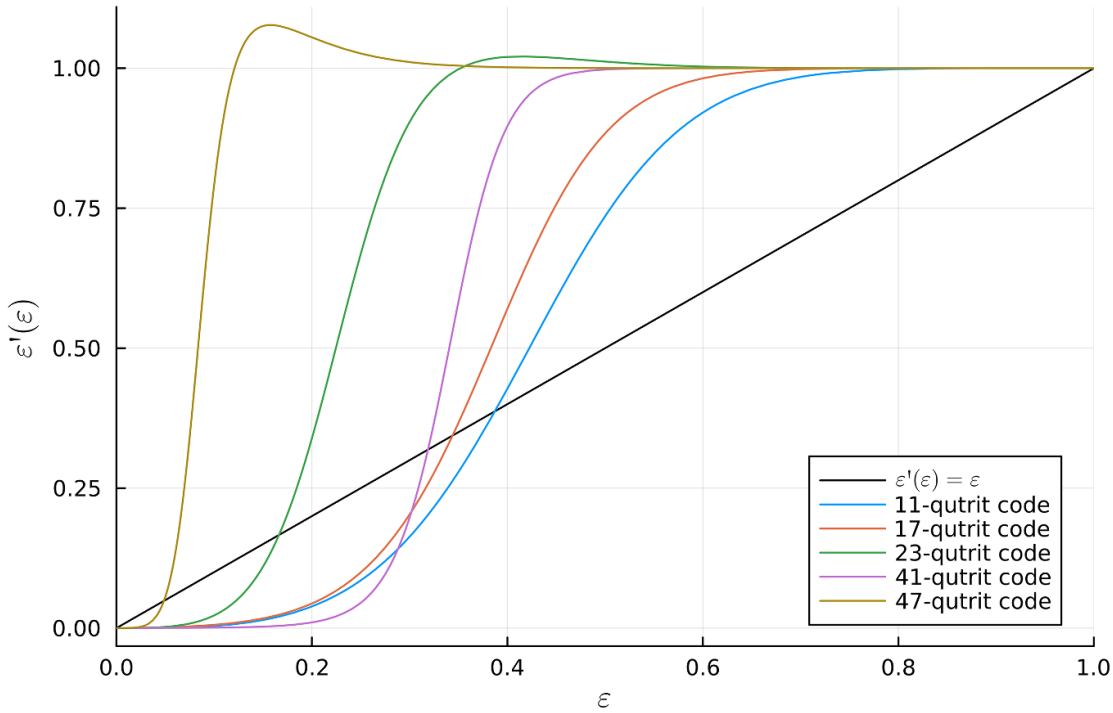


Figure 3. Post-distillation noise as a function of pre-distillation noise for codes that distill qutrit Strange states. The 11-qutrit Golay code, as well as the 17-, 23-, 41-, and 47-qutrit codes have non-trivial distillation thresholds.

Table 3. Noise thresholds of magic state distillation protocols for qutrit Strange states based on quadratic residue codes over $\mathbb{F}_9$.

Code	Threshold
$[11, 5, 6]_9$	≈ 0.38715
$[17, 8, 8]_9$	≈ 0.34394
$[23, 11, 9]_9$	≈ 0.16636
$[41, 20, 14]_9$	≈ 0.31877
$[47, 23, 15]_9$	≈ 0.04864

The coefficients a_j and b_j are obtained by imposing a number of linear constraints obtained by computing the numbers of vectors of certain low weights. These numbers are obtained using the algorithm of [60][5].

We computed the weight enumerators for all qutrit quantum QR codes up to length 47. The weight enumerators are listed in table 5 in appendix C. Using the weight enumerators found, together with the expression for the distillation function $\epsilon'(\epsilon)$ in terms of the weight enumerators of the corresponding code, we find that the length 11, 17, 23, 41, and 47 qutrit quantum QR codes distill Strange states with non-trivial thresholds. The distillation functions $\epsilon'(\epsilon)$ for these codes are plotted in figure 3, and the thresholds of these codes are listed in table 3. A plot showing the success probability of these protocols is given in figure 4. Additional plots showing the functions $\epsilon'(\epsilon)$ for all codes up to length 47 (not just those with non-trivial thresholds) are given in figure 6 in appendix D.

5. Discussion

In this paper, we showed that quantum QR codes are useful for distilling magic states. This includes showing that some codes which were previously known to distill qubit H and T states and qutrit Strange states are equivalent to certain quantum QR codes. We also found new examples of quantum QR codes

[5] A Julia implementation of this algorithm for codes over $\mathbb{F}_9$ that runs on Nvidia GPUs is also available at github.com/mzurel/WeightDistributionsGPU.jl.

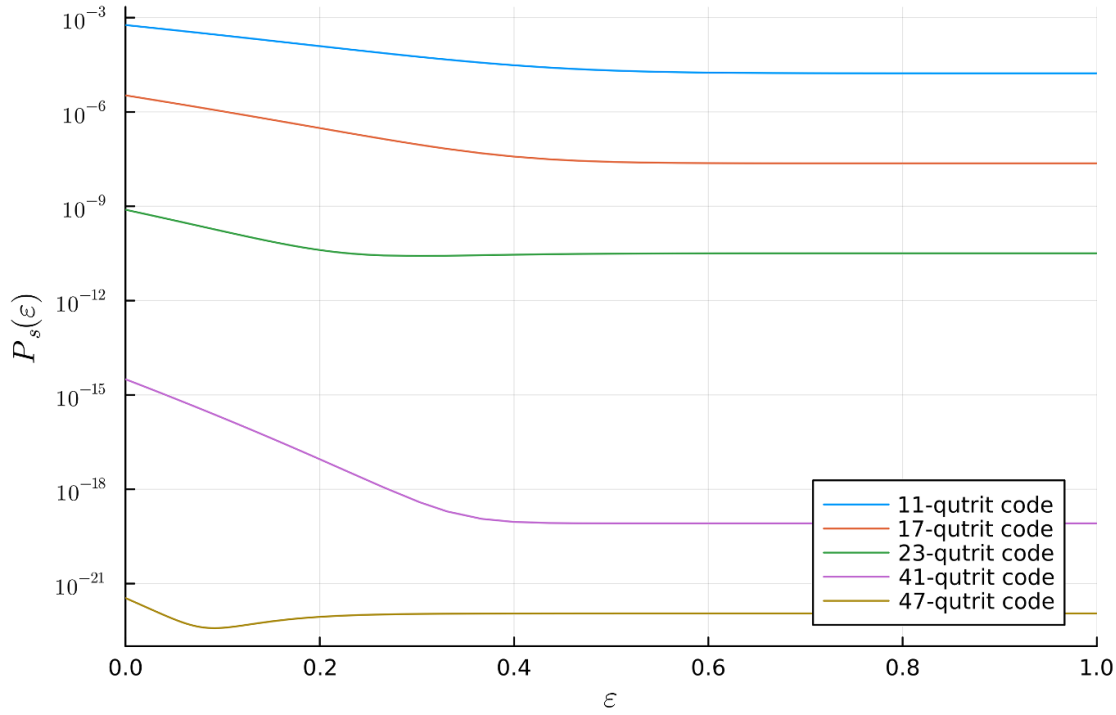


Figure 4. Success probability as a function of pre-distillation noise for codes that distill qutrit Strange states.

that distill qubit T states and qutrit Strange states. Finally, we showed that there are infinitely many quantum QR codes on qubits that distill T states with a non-trivial threshold.

A strength of the quantum QR code family is that many particularly special codes are unified under a single umbrella. This includes the codes with the highest currently known thresholds for T and Strange state distillation, namely the 5-qubit perfect code [1] and the 11-qutrit Golay code [16], as well as the 7-qubit Steane code and 23-qubit Golay code, which achieve the highest possible threshold for H state distillation [20]. Therefore, quantum QR codes provide a common algebraic framework for defining magic state distillation protocols for a variety of magic states on qubits and on qutrits.

In addition to reproducing previous magic state distillation protocols, we find a number of new distillation protocols, including some fairly large codes that distill qubit T states and qutrit Strange states with fairly high thresholds. This is contrary to the common expectation that smaller codes tend to have better distillation thresholds. In fact, from tables 2 and 3, we can see that for quantum QR codes the distillation thresholds are not monotonic either with the length of the code or with the code distance. Rather, the thresholds depend on the full weight distributions of the codes.

There is an infinite sequence of quantum QR codes of increasing size and increasing distance for both qubits and qutrits, as well as more generally for prime-dimensional qudits. Theorem 4 shows that there are infinitely many quantum QR codes that distill T states with a non-trivial threshold. Although this falls short of a resolution of the conjecture stated in the introduction, it does suggest that the capacity of these codes to distill T states is not merely incidental for some small examples. Some deeper algebraic structure of these codes makes them genuinely well-suited to magic state distillation.

Taken together, the above-mentioned results make quantum QR codes natural candidates for resolving the two conjectures stated in the introduction on which states are universal for quantum computation. That said, the performance of these codes for distilling magic states seems to depend on the entire weight distribution of the codes, and a complete analytic characterization of these weight distributions is likely out of reach with present methods. The weight distributions of the corresponding classical QR codes are not known in general either, and characterizing them has been an open problem since the codes were first introduced several decades ago.

In light of this limitation, one possible avenue for future work could be to extend the numeric results obtained here to larger code sizes. Extending the numerics to larger code lengths would also extend the resource usage analysis in appendix E, where we compare the performance of single-round protocols based on quantum QR codes against the baseline of multi-round protocols formed by concatenating the 5-qubit protocol with itself. For the accessible protocol sizes, the concatenated protocols seem to consume fewer raw input states for a given target output fidelity. However, there are reasons to expect

that larger single-round protocols could have benefits over multi-round protocols when the protocols" success probabilities are boosted, for example, by error correcting rather than post-selecting on a single syndrome. These are detailed in appendix E.

Furthermore, Theorem 4 shows that some asymptotic results are possible even without a complete characterization of the weight distributions of these codes. Another avenue could involve extending Theorem 4, for example, by showing that there exists a subsequence of qubit quantum QR codes with thresholds that remain bounded away from 0, or even which increase monotonically. An equivalent of Theorem 4 could also be shown for qutrit Strange state distillation. Finally, quantum QR codes are defined on qudits of any prime Hilbert space dimension, not just qubits and qutrits. They could lead to distillation protocols for ququint magic states and magic states on higher-dimensional qudits. Overall, the present results suggest that quantum QR codes are a promising testing ground for both finite computational results and asymptotic questions in magic state distillation.

Acknowledgments

We thank Amolak Ratan Kalra, Shiroman Prakash, and Mark Howard for insightful discussions. M. Z. is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC). S. J. is funded by the NSERC-European Commission project FoQaCiA. N. D. acknowledges support from the Canada Research Chair program, NSERC Discovery Grant RGPIN-2022-03103, the NSERC-European Commission project FoQaCiA, and the Faculty of Science of Simon Fraser University. This work was also supported by NSERC Alliance International Catalyst Quantum Grant (ALLRP 592647-24). Finally, this work was aided by discussions that took place at the Mathematical Foundations of Quantum Advantage workshop held at Simon Fraser University in May of 2025 and funded by the Pacific Institute for the Mathematical Sciences (PIMS).

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. The structure of magic state distillation protocols for qudits

In the main text, we use the Campbell-Browne result [8], which says that in analyzing the performance of n-to-1 qudit magic state distillation protocols (as measured by the conditional output fidelity or achievable noise threshold) it suffices to consider stabilizer reductions, i.e. protocols of the form

$$\rho \longmapsto \mathcal{R}\left(\rho^{\otimes n}\right) := Tr_{2,\ldots,n}\left(D\Pi_I^r \rho^{\otimes n} \Pi_I^r D^\dagger\right), \tag{126}$$

where $I \subset \mathbb{F}_d^{2n}$ is an isotropic subspace, $r : I \to \mathbb{F}_d$ is a consistent value assignment (see equation (16)), Π_I^r is the stabilizer projector (equation (14)), and D is a decoding Clifford unitary. In [8], this was proven only for qubits. In this appendix, we show that this result also holds for magic state distillation on odd-prime-dimensional qudits.

Theorem. *Let d be an odd prime. Consider any n-to-1 qudit stabilizer protocol Φ on d-dimensional qudits (a completely positive, trace non-increasing map built from Clifford unitaries, Pauli measurements with postselection, and partial trace). Fix an input state ρ and a target pure state $|M\rangle$. Then either (i) there is a stabilizer state whose fidelity with respect to $|M\rangle$ is at least that of $\Phi(\rho^{\otimes n})$, or (ii) there exists an n-to-1 qudit stabilizer protocol $\mathcal{R}$ whose conditional output fidelity with respect to $|M\rangle$ is at least that of Φ on input state $\rho^{\otimes n}$ and which has the form of a stabilizer reduction, for some $(n-1)$-dimensional isotropic subspace $I \subset \mathbb{F}_d^{2n}$, consistent assignment $r : I \to \mathbb{F}_d$, and decoding Clifford D.*

Proof sketch. Any stabilizer protocol Φ has a decomposition into a number of branches, with each branch determined by the outcomes of the Pauli measurements and any classical randomness used. For the fixed input state $\rho^{\otimes n}$, write the unnormalized output state of the protocol Φ as

$$\Phi\left(\rho^{\otimes n}\right) = \sum_\alpha \Pi_\alpha\left(\rho^{\otimes n}\right), \tag{127}$$

where α ranges over indices of accepted branches, and each channel Π_α is simply a sequence of Clifford gates $g_i \in \mathcal{C}\ell$ and Pauli projectors $\Pi_{I_j}^{r_j}$, followed by a partial trace of qudits 2 through n. Let $p_\alpha = Tr(\Pi_\alpha(\rho^{\otimes n}))$, and let $\rho_\alpha = \Pi_\alpha(\rho^{\otimes n})/p_\alpha$ be the normalized output state of each branch. The normalized output of Φ is

$$\rho' = \frac{\Phi\left(\rho^{\otimes n}\right)}{Tr\left(\Phi\left(\rho^{\otimes n}\right)\right)} = \sum_\alpha \frac{p_\alpha}{\sum_\beta p_\beta}\rho_\alpha. \tag{128}$$

The fidelity of the post-distillation state with respect to $|M\rangle$ is

$$\mathrm{Tr}\left(|M\rangle\langle M|\,\rho'\right) = \sum_\alpha \frac{p_\alpha}{\sum_\beta p_\beta}\mathrm{Tr}\left(|M\rangle\langle M|\,\rho_\alpha\right) \leqslant \max_\alpha \mathrm{Tr}\left(|M\rangle\langle M|\,\rho_\alpha\right), \tag{129}$$

where the first equality follows from linearity of the trace, and the last inequality follows from the fact that the coefficients are nonnegative. Therefore, there is a branch α^* such that the fidelity of ρ_{α^*} with respect to $|M\rangle$ is at least the fidelity of ρ' with respect to $|M\rangle$.

The channel Π_{α^*} has the form

$$\Pi_{\alpha^*}\left(\rho^{\otimes n}\right) = Tr_{2,\ldots,n}\left(\cdots \Pi_{I_2}^{r_2} g_2 \Pi_{I_1}^{r_1} g_1 \rho^{\otimes n} g_1^\dagger \Pi_{I_1}^{r_1} g_2^\dagger \Pi_{I_2}^{r_2} \cdots\right). \tag{130}$$

with $g_1, g_2, \cdots \in \mathcal{C}\ell$, $I_1, I_2, \cdots \subset \mathbb{F}_d^{2n}$ isotropic subspaces, and $r_1, r_2, \ldots$ consistent value assignments for the subspaces. This can be reduced to the form of a stabilizer reduction using standard techniques from Pauli-based computation (see, for example, [63–65]). In particular, to start the Clifford gates can be propagated past the projectors, conjugating the Pauli projectors into other Pauli projectors, resulting in a map of the form

$$\Pi_{\alpha^*}\left(\rho^{\otimes n}\right) = Tr_{2,\ldots,n}\left(g \cdots \Pi_{I_2'}^{r_2'} \Pi_{I_1'}^{r_1'} \rho^{\otimes n} \Pi_{I_1'}^{r_1'} \Pi_{I_2'}^{r_2'} \cdots g^\dagger\right). \tag{131}$$

Then using techniques from Pauli-based computation (see [65]), the sequence of Pauli projectors can be reduced to a single Pauli projector Π_I^r to get a map of the form

$$\Pi_{\alpha^*}\left(\rho^{\otimes n}\right) = Tr_{2,\ldots,n}\left(g \Pi_I^r \rho^{\otimes n} \Pi_I^r g^\dagger\right). \tag{132}$$

Note that this step holds for qudits of all prime dimensions, but not necessarily for qudits of composite dimensions, because it assumes the Clifford group acts transitively on the sets of Pauli projectors Π_i^r of each dimension.

The last step involves tracing out the last $n-1$ qudits. This is equivalent to measuring observables $Z_2, \ldots, Z_n$ on these qudits and then discarding the outcomes. Let $J = \langle z_2, \ldots, z_n\rangle$ and $s : J \to \mathbb{F}_d$ be the labels of the measurement outcomes for measurements J. Then

$$\Pi_{\alpha^*}\left(\rho^{\otimes n}\right) = \sum_{s \in J^*} p(s)\, Tr_{2,\ldots,n}\left(\Pi_J^s g \Pi_I^r \rho^{\otimes n} \Pi_I^r g^\dagger \Pi_J^s\right). \tag{133}$$

with nonnegative coefficients $p(s)$. By the same argument as before, without decreasing the output fidelity (but reducing the success probability), we can post-select on one of the outcomes $s : J \to \mathbb{F}_d$.

Now, propagating the new projector Π_J^s past the Clifford group element g, and absorbing it into the old projector, we get

$$\Pi_J^s g \Pi_I^{\ r} \propto g \Pi_{I'}^{r'}, \tag{134}$$

where $\Pi_{I'}^{r'}$ has rank 1 or d. Here we have two cases. If $\Pi_{I'}^{r'}$ has rank 1, then the post-distillation state is a stabilizer state and we are in case 1 of the theorem statement. Otherwise, $\Pi_{I'}^{r'}$ has rank d, we are still in the second case of the theorem statement, and we can assume $\dim(I') = n-1$. Also, by construction, g is a decoding Clifford for $\Pi_{I'}^{r'}$, it maps the code space onto the first physical qudit.

The result is a protocol of the form

$$\rho \longmapsto \mathcal{R}\left(\rho^{\otimes n}\right) = Tr_{2,\ldots,n}\left(g \Pi_{I'}^{r'} \rho^{\otimes n} \Pi_{I'}^{r'} g^\dagger\right), \tag{135}$$

where $I' \subset \mathbb{F}_d^{2n}$ is a $n-1$-dimensional isotropic subspace, $r' : I' \to \mathbb{F}_d$ is a consistent value assignment for I', and g is a Clifford that maps the code space of I' onto the first physical qudit. This is exactly a stabilizer reduction. By construction, this stabilizer reduction has conditional output fidelity at least that of the original protocol Φ on $\rho^{\otimes n}$.

□

Appendix B. Stabilizer operations in phase space

In this section, we assume we have a system of n qudits, each with odd-prime Hilbert space dimension d. It is helpful to understand how stabilizer operations act in the phase space picture. This section is reproduced here for convenience from parts of section 2.4 of [66].

First, we consider Clifford unitaries. For odd-prime d, the Clifford group splits as

$$\mathcal{C}\ell \cong \mathrm{Sp}\left(\mathbb{F}_d^{2n}\right) \ltimes \mathbb{F}_d^{2n}, \tag{136}$$

and Clifford operations act on the phase space as affine symplectic transformations,

$$gA_u g^\dagger = A_{S_g u + a_g} \qquad \forall g \in \mathcal{C}\ell \quad \forall u \in \mathbb{F}_d^{2n}. \tag{137}$$

where $S_g \in \mathrm{Sp}(\mathbb{F}_d^{2n})$ and $a_g \in \mathbb{F}_d^{2n}$ [30, theorem 7]. The Pauli group is a subgroup of the Clifford group corresponding to translations of the phase space, $T_a A_u T_a^\dagger = A_{u+a}$. As a corollary of equation (137), the Wigner function is Clifford covariant, i.e. for any quantum state ρ and any Clifford group element $g \in \mathcal{C}\ell$,

$$W_{g\rho g^\dagger}\left(S_g u + a_g\right) = W_\rho(u). \tag{138}$$

The effect of Pauli measurements on the phase space is described by the following lemma.

Lemma 3. *For any isotropic subspace $I \subset \mathbb{F}_d^{2n}$, any value assignment $r : I \to \mathbb{F}_d$, and any phase space point $u \in \mathbb{F}_d^{2n}$,*

(i) *if $r(a) = -[u,a]$ for all $a \in I$, then $Tr(\Pi_I^r A_u) = 1$ and*

$$\Pi_I^r A_u \Pi_I^r = \frac{1}{|I|} \sum_{v \in u+I} A_v, \tag{139}$$

(ii) *if $r(a) \neq -[u,a]$ for some $a \in I$, then $Tr(\Pi_I^r A_u) = 0$ and $\Pi_I^r A_u \Pi_I^r = 0$.*

Proof of lemma 3. For any isotropic subspace $I \subset \mathbb{F}_d^{2n}$, any value assignment $r : I \to \mathbb{F}_d$, and any phase space point $u \in \mathbb{F}_d^{2n}$,

$$\Pi_I^r A_u \Pi_I^r = \frac{1}{|I|^2 d^n} \sum_{a,c \in I} \sum_{b \in \mathbb{F}_d^{2n}} \omega^{-r(a)-r(c)+[u,b]} T_a T_b T_c \tag{140}$$

$$= \frac{1}{|I|^2 d^n} \sum_{a,c \in I} \sum_{b \in \mathbb{F}_d^{2n}} \omega^{-r(a+c)+[u,b]+[b,c]} T_{a+c} T_b \tag{141}$$

$$= \frac{1}{|I|^2 d^n} \sum_{a \in I} \sum_{b \in \mathbb{F}_d^{2n}} \omega^{-r(a)+[u,b]} \left[\sum_{c \in I} \omega^{[b,c]}\right] T_a T_b \tag{142}$$

$$= \frac{1}{|I| d^n} \sum_{a \in I} \sum_{b \in I^\perp} \omega^{-r(a)+[u,b]} T_a T_b. \tag{143}$$

Taking a trace of this equation, we get

$$Tr\left(\Pi_I^r A_u\right) = \frac{1}{|I| d^n} \sum_{a \in I} \sum_{b \in I^\perp} \omega^{-r(a)+[u,b]} d^n \delta_{b,-a} \tag{144}$$

$$= \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)-[u,a]} \tag{145}$$

$$= \begin{cases} 1 & \text{if} r(a) = -[u,a] \ \forall a \in I, \\ 0 & \text{otherwise.} \end{cases} \tag{146}$$

Here the last line follows from orthogonality of characters. Now, assuming $r(a) = -[u,a]\ \forall a \in I$, we have

$$\Pi_I^r A_u \Pi_I^r = \frac{1}{|I| d^n} \sum_{a \in I} \sum_{b \in I^\perp} \omega^{[u,a+b]} T_{a+b}. \tag{147}$$

Since I is isotropic, $I \subset I^\perp$, and we can write $I^\perp$ as $I^\perp = \bigsqcup_{a \in I^\perp/I} I + a$. Then $I + I^\perp = |I| \cdot I^\perp$, and

$$\Pi_I^r A_u \Pi_I^r = \frac{1}{d^n} \sum_{a \in I^\perp} \omega^{[u,a]} T_a. \tag{148}$$

It is easy to check that for any $v \in u + I$, we have $[v,w] = [u,w]\ \forall w \in I^\perp$. Then

$$\frac{1}{|I|} \sum_{v \in u+I} A_v = \frac{1}{|I| d^n} \sum_{v \in u+I} \sum_{b \in \mathbb{F}_d^{2n}} \omega^{[v,b]} T_b \tag{149}$$

$$= \frac{1}{|I| d^n} \sum_{b \in \mathbb{F}_d^{2n}} \left[\sum_{v \in u+I} \omega^{[v,b]} \right] T_b \tag{150}$$

$$= \frac{1}{d^n} \sum_{b \in I^\perp} \omega^{[u,b]} T_b = \Pi_I^r A_u \Pi_I^r. \tag{151}$$

This proves the first statement of the lemma. It remains to show that in the case that there exists some $a \in I$ such that $r(a) \neq -[u,a]$, we have $\Pi_I^r A_u \Pi_I^r = 0$ as an operator. To see this, consider the inner product $Tr(T_a \Pi_I^r A_u \Pi_I^r)$ for any Pauli operator $T_a,\ a \in \mathbb{F}_d^{2n}$. Here we have three cases.

(I) If $a \in I$, then $T_a \Pi_I^r = \omega^{r(a)} \Pi_I^r$, and so $Tr(T_a \Pi_I^r A_u \Pi_I^r) = \omega^{r(a)} Tr(\Pi_I^r A_u) = 0$.

(II) If $a \in I^\perp \setminus I$, then by lemma 5 of [67], $\Pi_I^r = \sum_{r' \in \Gamma} \Pi_{\langle a,I \rangle}^{r'}$, where Γ is the set of all value assignments on $\langle a, I \rangle$ which agree with r on I. Then

$$Tr(\Pi_I^r A_u) = \sum_{r' \in \Gamma} Tr\left(\Pi_{\langle a,I \rangle}^{r'} A_u \right). \tag{152}$$

Here the left hand side is zero by assumption, and on the right hand side we have a sum of nonnegative terms. Therefore, each term in the sum on the right hand side is zero. The spectral decomposition of T_a is

$$T_a = \sum_{k \in \mathbb{Z}_d} \omega^k \Pi_a^k. \tag{153}$$

Then,

$$Tr(T_a \Pi_I^r A_u \Pi_I^r) = \sum_{k \in \mathbb{Z}_d} \omega^k Tr\left(\Pi_a^k \Pi_I^r A_u \Pi_I^r \right) \tag{154}$$

$$= \sum_{k \in \mathbb{Z}_d} \omega^k Tr\left(\Pi_{\langle a,I \rangle}^{r_k} A_u \right) = 0. \tag{155}$$

(III) If $a \notin I^\perp$, then

$$\Pi_I^r T_a \Pi_I^r = \frac{1}{|I|^2} \sum_{b,c \in I} \omega^{-r(b)-r(c)} T_b T_a T_c \tag{156}$$

$$= \frac{1}{|I|^2} T_a \sum_{b,c \in I} \omega^{-r(b+c)+[b,a]} T_{b+c} \tag{157}$$

$$= \frac{1}{|I|^2} T_a \Pi_I^r \sum_{b \in I} \omega^{[b,a]}. \tag{158}$$

By character orthogonality, the sum in the final expression vanishes. Therefore,

$$Tr(T_a \Pi_I^r A_u \Pi_I^r) = Tr(\Pi_I^r T_a \Pi_I^r A_u) = 0. \tag{159}$$

Thus, $Tr(T_a \Pi_I^r A_u \Pi_I^r) = 0$ for all Pauli operators $T_a, a \in \mathbb{F}_d^{2n}$, and so $\Pi_I^r A_u \Pi_I^r = 0$ as an operator. □

Appendix C. Weight enumerator tables

Table 4. Weight enumerator polynomials of expurgated quadratic residue codes over $GF(4)$.

Code	Weight enumerator
$[5,2,4]_4$	x^5+15xy^4
$[7,3,4]_4$	$x^7+21x^3y^4+42xy^6$
$[13,6,6]_4$	$x^{13}+156x^7y^6+1053x^5y^8+2028x^3y^{10}+858xy^{12}$
$[23,11,8]_4$	$x^{23}+1518x^{15}y^8+110124x^{11}y^{12}+425040x^9y^{14}+1298649x^7y^{16}+1530144x^5y^{18}$ $+743820x^3y^{20}+85008xy^{22}$
$[29,14,12]_4$	$x^{29}+71253x^{17}y^{12}+613872x^{15}y^{14}+5618025x^{13}y^{16}+24701040x^{11}y^{18}+65315250x^9y^{20}$ $+91040976x^7y^{22}+62560134x^5y^{24}+17276112x^3y^{26}+1238793xy^{28}$
$[31,15,8]_4$	$x^{31}+1395x^{23}y^8+58590x^{19}y^{12}+468720x^{17}y^{14}+6072807x^{15}y^{16}+37914240x^{13}y^{18}$ $+135749124x^{11}y^{20}+296335200x^9y^{22}+345151365x^7y^{24}+200987136x^5y^{26}$ $+48076350x^3y^{28}+2926896xy^{30}$
$[37,18,12]_4$	$x^{37}+12987x^{25}y^{12}+163836x^{23}y^{14}+4190472x^{21}y^{16}+49694700x^{19}y^{18}$ $+402823773x^{17}y^{20}+2140777080x^{15}y^{22}+7315768908x^{13}y^{24}+15818285880x^{11}y^{26}$ $+20702769285x^9y^{28}+15425347212x^7y^{30}+5876290827x^5y^{32}+942953436x^3y^{34}$ $+40398339xy^{36}$
$[47,23,12]_4$	$x^{47}+38916x^{35}y^{12}+1070190x^{31}y^{16}+9988440x^{29}y^{18}+231731808x^{27}y^{20}$ $+3350122776x^{25}y^{22}+32422518540x^{23}y^{24}+226165393080x^{21}y^{26}+1134417096120x^{19}y^{28}$ $+4010049018360x^{17}y^{30}+9894334090905x^{15}y^{32}+16673213458440x^{13}y^{34}$ $+18572482193952x^{11}y^{36}+13080263144520x^9y^{38}+5432928694380x^7y^{40}$ $+1192386006504x^5y^{42}+113540737860x^3y^{44}+2948872872xy^{46}$
$[53,26,16]_4$	$x^{53}+196365x^{37}y^{16}+1008696x^{35}y^{18}+94594188x^{33}y^{20}+1497103296x^{31}y^{22}$ $+25024672920x^{29}y^{24}+272515859616x^{27}y^{26}+2301852043920x^{25}y^{28}$ $+14237883206976x^{23}y^{30}+65437002577947x^{21}y^{32}+220371245649360x^{19}y^{34}$ $+538407879739080x^{17}y^{36}+937381604347968x^{15}y^{38}+1135693969571784x^{13}y^{40}$ $+925961040090528x^{11}y^{42}+484514894731920x^9y^{44}+151674555714240x^7y^{46}$ $+25413417241719x^5y^{48}+1867132096440x^3y^{50}+38016923532xy^{52}$
$[61,30,18]_4$	$x^{61}+201300x^{43}y^{18}+9791964x^{41}y^{20}+297188340x^{39}y^{22}+7311487755x^{37}y^{24}$ $+134130915792x^{35}y^{26}+1902476580870x^{33}y^{28}+20779092892368x^{31}y^{30}$ $+175323674956545x^{29}y^{32}+1141965426170520x^{27}y^{34}+5726086694463408x^{25}y^{36}$ $+21992184927250200x^{23}y^{38}+64200120673894854x^{21}y^{40}+140927247768490320x^{19}y^{42}$ $+229267347842000340x^{17}y^{44}+271133644941905040x^{15}y^{46}+227146495371493950x^{13}y^{48}$ $+130168863213334452x^{11}y^{50}+48592446257133540x^9y^{52}+11002063150069140x^7y^{54}$ $+1350253440272367x^5y^{56}+73516463568960x^3y^{58}+1121442784950xy^{60}$
$[71,35,12]_4$	$x^{71}+7455x^{59}y^{12}+499485x^{55}y^{16}+44730x^{53}y^{18}+42565068x^{51}y^{20}+51290400x^{49}y^{22}$ $+2044474110x^{47}y^{24}+24392610900x^{45}y^{26}+478715797065x^{43}y^{28}+8385401713080x^{41}y^{30}$ $+123970358688225x^{39}y^{32}+1477842682144710x^{37}y^{34}+14063872023855282x^{35}y^{36}$ $+107143099467071640x^{33}y^{38}+652728283699886100x^{31}y^{40}$ $+3172457555098365720x^{29}y^{42}+12254266094922768945x^{27}y^{44}$ $+37402445047293925872x^{25}y^{46}+89525811715214684355x^{23}y^{48}$ $+166409387015909428710x^{21}y^{50}+237190163461701506280x^{19}y^{52}$ $+255090582785520762000x^{17}y^{54}+202747446918460600590x^{15}y^{56}$ $+115908476245380852180x^{13}y^{58}+45970393268864625351x^{11}y^{60}$ $+12033457215125864280x^9y^{62}+1933977124137515550x^7y^{64}$ $+170403391711501914x^5y^{66}+6731583805555770x^3y^{68}+75335382697656xy^{70}$

Table 5. Weight enumerator polynomials of expurgated quadratic residue codes over $GF(9)$.

Code	Weight enumerator
$[5,2,4]_9$	$x^5+40xy^4+40y^5$
$[11,5,6]_9$	$x^{11}+528x^5y^6+7920x^3y^8+11000x^2y^9+23760xy^{10}+15840y^{11}$
$[17,8,8]_9$	$x^{17}+2720x^9y^8+77248x^7y^{10}+220864x^6y^{11}+1214752x^5y^{12}+3179680x^4y^{13}$ $+7936960x^3y^{14}+12206272x^2y^{15}+12404152xy^{16}+5804072y^{17}$
$[23,11,9]_9$	$x^{23}+10120x^{14}y^9+366528x^{11}y^{12}+2003760x^{10}y^{13}+13115520x^9y^{14}+62298720x^8y^{15}$ $+241544160x^7y^{16}+803325600x^6y^{17}+2154216064x^5y^{18}+4512831840x^4y^{19}$ $+7226651520x^3y^{20}+8268193824x^2y^{21}+6005888064xy^{22}+2090613888y^{23}$
$[29,14,12]_9$	$x^{29}+43848x^{17}y^{12}+82824x^{16}y^{13}+1714944x^{15}y^{14}+13739040x^{14}y^{15}+90658176x^{13}y^{16}$ $+571797408x^{12}y^{17}+3028883424x^{11}y^{18}+13993514304x^{10}y^{19}+56123954040x^9y^{20}$ $+192285362808x^8y^{21}+559241011200x^7y^{22}+1362023628000x^6y^{23}$ $+2723779515240x^5y^{24}+4357956668520x^4y^{25}+5363900914848x^3y^{26}$ $+4767732820928x^2y^{27}+2724475922400xy^{28}+751572223008y^{29}$
$[41,20,14]_9$	$x^{41}+13120x^{27}y^{14}+298480x^{25}y^{16}+984000x^{24}y^{17}+51010560x^{23}y^{18}$ $+241860640x^{22}y^{19}+3178539760x^{21}y^{20}+21320311600x^{20}y^{21}+169521814400x^{19}y^{22}$ $+1076372094720x^{18}y^{23}+6584910839040x^{17}y^{24}+35468786005600x^{16}y^{25}$ $+175512737265408x^{15}y^{26}+778023868378880x^{14}y^{27}+3116202323611840x^{13}y^{28}$ $+11168036013571840x^{12}y^{29}+35749106898555904x^{11}y^{30}+101465555251461632x^{10}y^{31}$ $+253683676686494280x^9y^{32}+553469380198545480x^8y^{33}+1041846451913296000x^7y^{34}$ $+1666936669433112672x^6y^{35}+2222593774949021136x^5y^{36}$ $+2402798170045927120x^4y^{37}+2023411266902614720x^3y^{38}$ $+1245175538958225280x^2y^{39}+498070323942465632xy^{40}+97184444550609056y^{41}$
$[47,23,15]_9$	$x^{47}+1141536x^{32}y^{15}+50417840x^{29}y^{18}+55935264x^{27}y^{20}+2167776864x^{26}y^{21}$ $+12302333472x^{25}y^{22}+115066828800x^{24}y^{23}+957374289216x^{23}y^{24}$ $+7020700962528x^{22}y^{25}+47575787881248x^{21}y^{26}+296009366479904x^{20}y^{27}$ $+1691060003624640x^{19}y^{28}+8862795495027936x^{18}y^{29}+42539887522934208x^{17}y^{30}$ $+186634650694261440x^{16}y^{31}+746558999374719168x^{15}y^{32}$ $+2714712409349744760x^{14}y^{33}+8942498372112933936x^{13}y^{34}$ $+26572195657617214176x^{12}y^{35}+70859323050286695456x^{11}y^{36}$ $+168529756696023632448x^{10}y^{37}+354799566202158364320x^9y^{38}$ $+655015292493578207136x^8y^{39}+1048023930654922368576x^7y^{40}$ $+1431446993858724464160x^6y^{41}+1635940023559588443840x^5y^{42}$ $+1521804500115282894720x^4y^{43}+1106766736976396745216x^3y^{44}$ $+590275754871576999040x^2y^{45}+205313252389258020096xy^{46}$ $+34946943259659753984y^{47}$

Appendix D. Post-distillation noise plots

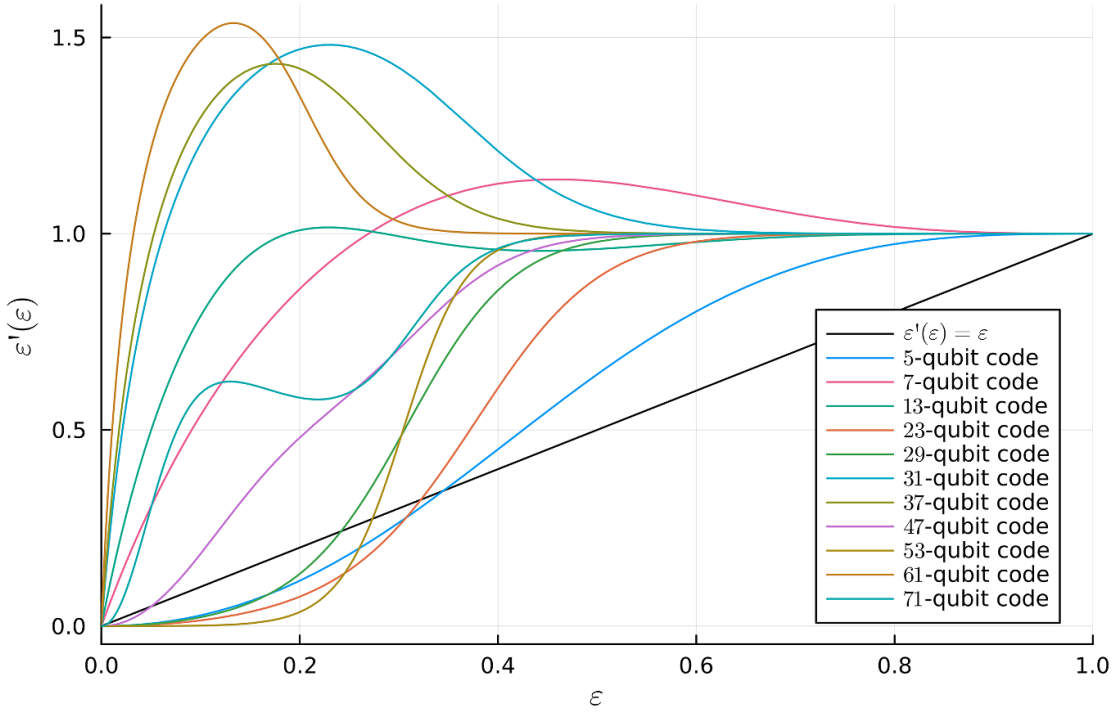


Figure 5. Post-distillation noise as a function of predistillation noise for magic state distillation protocols based on expurgated quadratic residue codes for qubit T states. The length 5, 23, 29, 47, 53, and 71 codes have nonzero threshold. The other code lengths tested have a fixed point at the pure magic state but cannot distill them from noisy magic states.

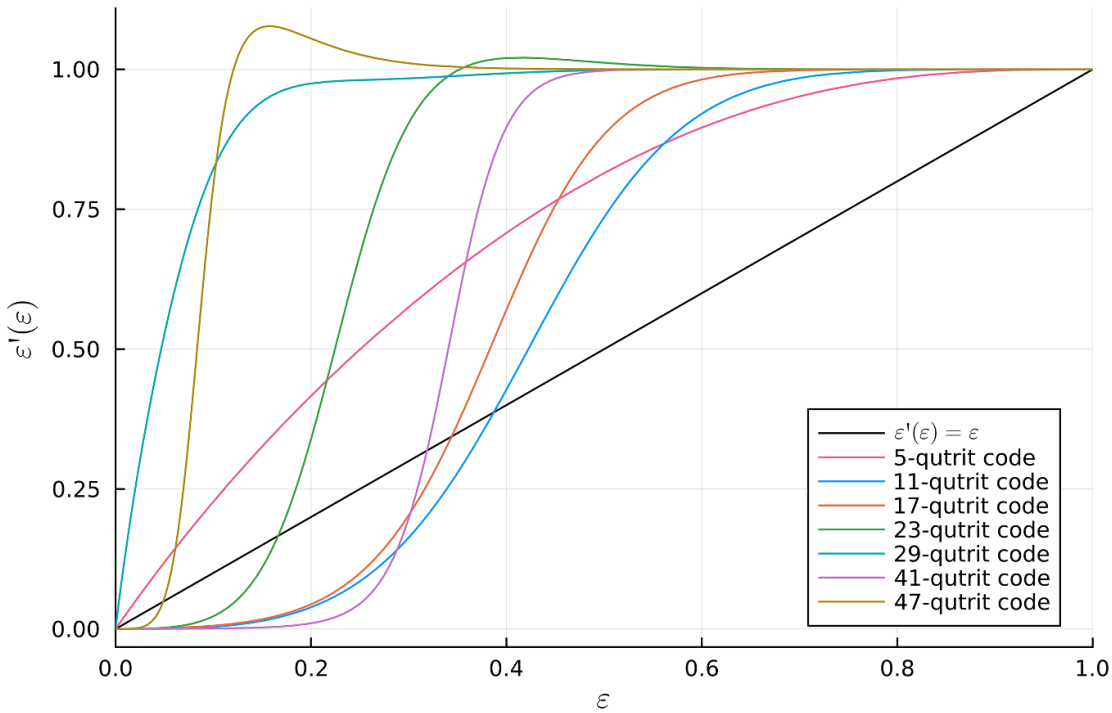


Figure 6. Post-distillation noise as a function of predistillation noise for magic state distillation protocols based on expurgated quadratic residue codes for qutrit Strange states. The length 11, 17, 23, 41, and 47 codes have a nonzero threshold. The other code lengths tested have a fixed point at the pure magic state but cannot distill them from noisy magic states.

Appendix E. Resource analysis for distillation with concatenated codes

The main figure of merit in this paper is the noise threshold of magic state distillation protocols. However, as long as the input noise parameter ϵ_{in} is below the threshold of a given protocol, the protocol can be applied repeatedly, allowing the noise parameter to be made arbitrarily small. If this condition is satisfied, then the practically relevant figure of merit becomes the resource consumption to achieve a given output noise parameter ϵ_{out}. There are many ways of measuring this, for example, the number of distillation rounds required, the number of input states required to reach a given noise parameter, and taking the success probability of the protocol into account, the expected number of raw input states required to achieve a given noise parameter.

Two properties of the quantum quadratic residue (QR) code family make them potentially favourable for this sort of resource usage comparison. First, a protocol formed by concatenating a fixed n-to-1 qubit protocol with itself is possible only at sizes n^r for $r \in \mathbb{N}$. For the case of the concatenated 5-qubit protocol, these protocols are available at sizes 5, 25, 125, 625, …. If a target output error lies between the errors achieved by the r-round and $r+1$-round concatenated protocols, the $r+1$-round protocol must be used, potentially requiring a large number of additional input states, as well as additional latency from running an additional round. On the other hand, qubit quantum QR codes that distill T states exist for every prime length p such that $p \equiv 5$ or $23 \pmod{24}$ (see theorem 4). The result is a much denser distribution of distillation protocol sizes than those formed from concatenating small protocols, potentially saving resource usage. For example, between the length 25 and 125 concatenated 5-qubit protocols, there are quantum QR code-based protocols of length 29, 47, 53, 71, and 101.

Second, an r-round concatenated protocol requires r sequential rounds of encoding, measurement, post-selection, decoding, and classical processing between rounds, whereas a single-round protocol can reach the desired output noise parameter with a single round of measurement. This paper assumes ideal stabilizer operations throughout, so this effect is not seen in the plots here, but on realistic near-term hardware, fewer rounds of distillation can itself be advantageous, especially on hardware with limited coherence times.

In this appendix, we compare the performance of the quantum QR codes for distilling qubit T states against the baseline of protocols formed from concatenating the 5-qubit protocol [1]. Concatenation is modeled by iterating the distillation map $\epsilon'(\epsilon)$. That is, the initial noise parameter is $\epsilon_0 = \epsilon_{\text{in}}$, and the noise parameter after r rounds is $\epsilon_r = \epsilon'(\epsilon_{r-1})$. A successful, fully post-selected run of the protocol will consume n^r qudits.

Figure 7 plots the output noise parameter ϵ_{out} against the number of input T states used, assuming we post-select on the desired measurement outcomes. The protocols based on quantum QR codes fill some of the gaps between the concatenated 5-qubit protocols. For example, at $\epsilon_{\text{in}} = 0.1$, a single round of the 23-qubit protocol reaches an output noise parameter intermediate between one and two rounds of the 5-qubit protocol, and a single round of the 53-qubit protocol reaches output noise parameter intermediate between two and three rounds of the 5-qubit protocol. That said, the concatenated 5-qubit protocol achieves the fastest error suppression of the protocols tested, especially with $\epsilon_{\text{in}} = 0.01$.

Figure 8 shows a similar plot, but without assuming post-selection on the desired measurement outcomes. Instead, it takes into account the success probability of the protocols, and plots the output noise parameter against the expected number of raw input T states. For r rounds of a n-to-1 qubit protocol, the expected number of T states is

$$N_r = \prod_{j=0}^{r-1} \frac{n}{P_{\text{s}}\left(\epsilon_j\right)}, \tag{160}$$

since producing one accepted output at round $j+1$ requires on average $n/P_{\text{s}}(\epsilon_j)$ states from round j. Both plots include a subplot that assumes the raw input T states have initial noise parameter $\epsilon_{\text{in}} = 0.1$, and another subplot for $\epsilon_{\text{in}} = 0.01$.

Figure 8 shows that the efficiency of larger protocols suffers when the success probability is taken into account since the success probability seems to fall off rapidly with the size of the protocol (see figure 2 in the main text). The expected raw resource state consumption of these larger protocols is dominated by rejected runs, and the concatenated 5-qubit protocols seem to consume far fewer raw resource states per accepted output state for a given target output noise parameter.

There are two limitations of this resource usage analysis to keep in mind. First, our numerical results are limited to codes of length $pl10^{71}$ as a result of the cost of computing the weight enumerators of

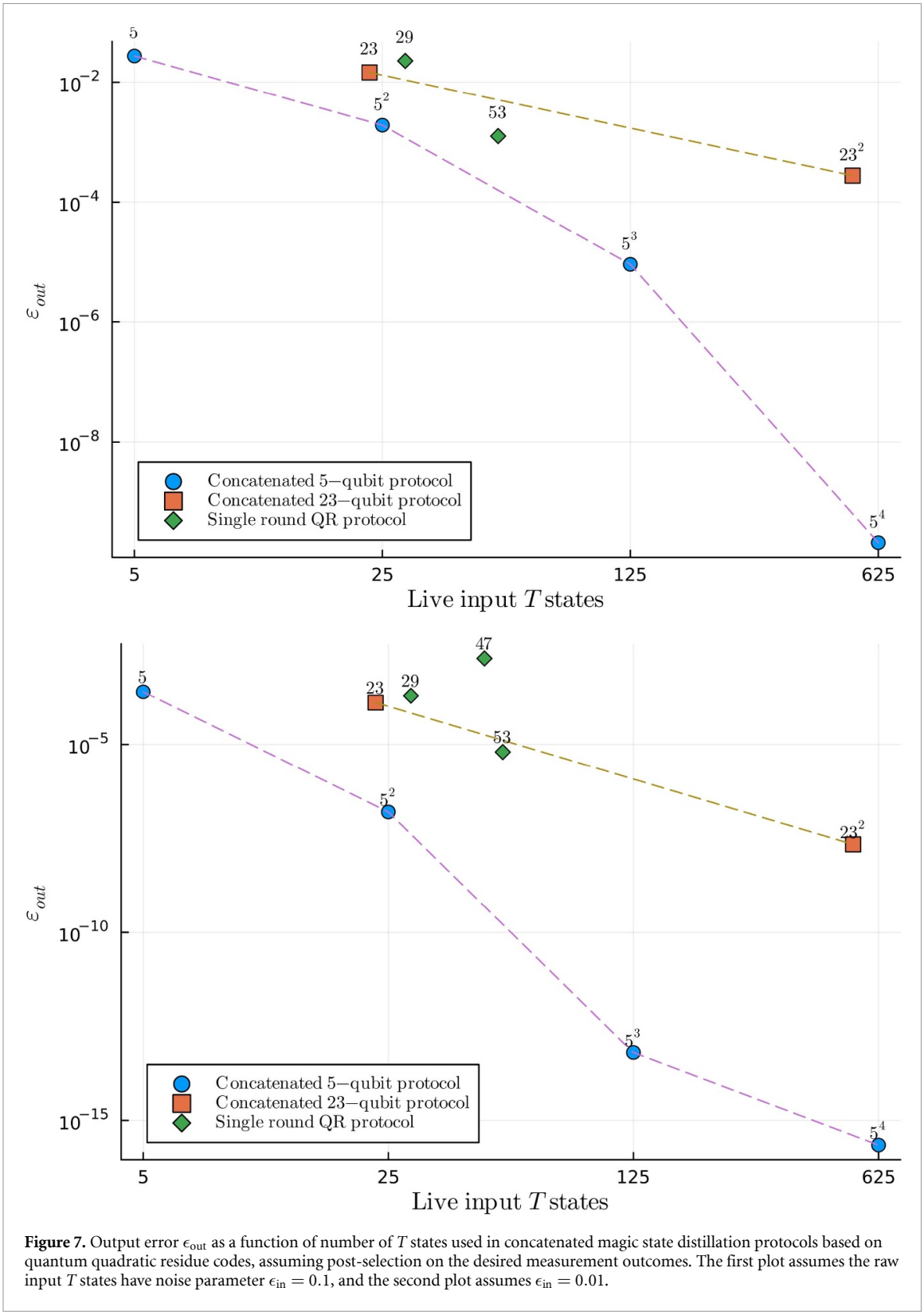


Figure 7. Output error ϵ_{out} as a function of number of T states used in concatenated magic state distillation protocols based on quantum quadratic residue codes, assuming post-selection on the desired measurement outcomes. The first plot assumes the raw input T states have noise parameter $\epsilon_{\text{in}} = 0.1$, and the second plot assumes $\epsilon_{\text{in}} = 0.01$.

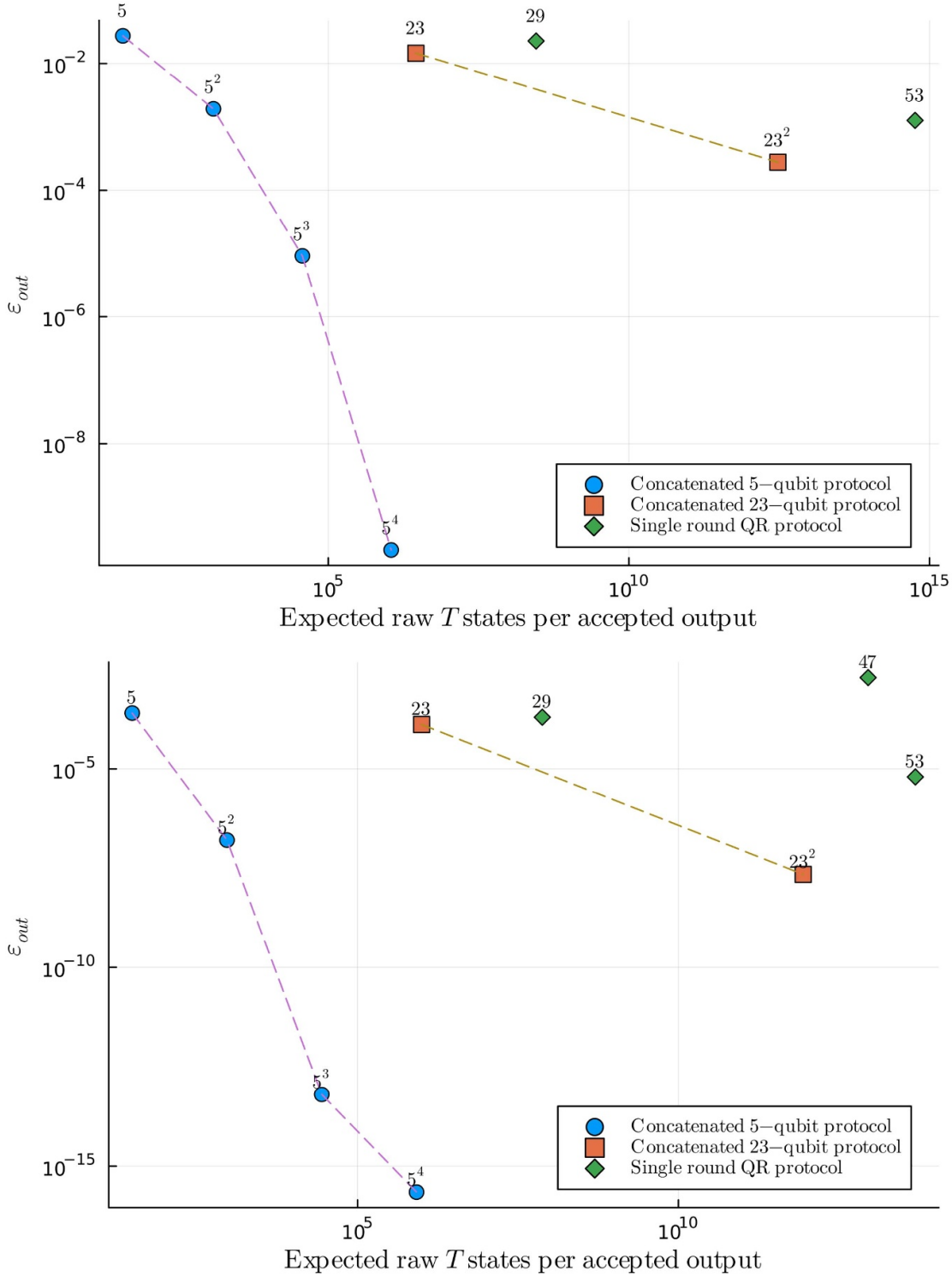


Figure 8. Output error parameter ϵ_{out} as a function of the expected number of raw input T states consumed by concatenated magic state distillation protocols based on quantum quadratic residue codes. The first plot assumes the raw input T states have noise parameter $\epsilon_{\text{in}} = 0.1$, and the second plot assumes $\epsilon_{\text{in}} = 0.01$.

the relevant codes. Therefore, the above argument on the relative density of protocol sizes for single-round QR code protocols versus concatenated protocols could become more significant for larger qubit numbers.

Second, since the noise threshold is our main figure of merit for the bulk of the paper, we make a number of assumptions that simplify the analysis without impacting the threshold. These assumptions could, however, have a significant effect on the current resource usage estimates. For example, in the protocols defined in section 4, we always post-select on a single syndrome measurement. This has a huge effect on the success probability. It is possible that allowing post-selection on other syndromes can boost

the success probability of larger protocols without significant impact on other figures of merit like the noise threshold. We expect this to hold in the case of quantum QR codes, since corollary 3 shows that the distance of the codes grows with the square root of the length. A distance δ code can correct errors up to weight $\lfloor(\delta-1)/2\rfloor$, so a protocol in which low-weight nontrivial syndromes are corrected instead of discarded could significantly boost the success probability. The analysis of error-corrected protocols is more complicated, and we leave it to future work.

The framework presented in this appendix, and in this paper broadly, can continue to be applied as the weight enumerators of larger codes become practically computable. The above mentioned properties of quantum QR codes make them promising objects of study for improved magic state distillation protocols.

ORCID iDs

Michael Zurel 0000-0003-0333-1174
Santanil Jana 0009-0006-5243-0653
Nadish de Silva 0000-0001-7904-306X

References

[1] Bravyi S and Kitaev A 2005 Universal quantum computation with ideal clifford gates and noisy ancillas *Phys. Rev. A* **71** 022316
[2] Campbell E T, Terhal B M and Vuillot C 2017 Roads towards fault-tolerant universal quantum computation *Nature* **549** 172–9
[3] Gottesman D 1997 Stabilizer codes and quantum error correction. *PhD Thesis* California Institute of Technology
[4] Gottesman D 1998 The heisenberg representation of quantum computers *Proc. XXII Int. Coll. on Group Theoretical Methods in Physics* (*Cambridge*), ed S P Corney, R Delbourgo and P D Jarvis (International Press of Boston Inc) pp 32–43
[5] Gottesman D 1998 Fault-tolerant quantum computation with higher-dimensional systems *NASA Int. Conf. on Quantum Computing and Quantum Communications* (Springer) pp 302–13
[6] Aaronson S and Gottesman D 2004 Improved simulation of stabilizer circuits *Phys. Rev. A* **70** 052328
[7] Eastin B and Knill E 2009 Restrictions on transversal encoded quantum gate sets *Phys. Rev. Lett.* **102** 110502
[8] Campbell E T and Browne D E 2009 On the Structure of Protocols for Magic State Distillation *Theory of Quantum Computation, Communication and Cryptography. Volume 5906 of Lecture Notes in Computer Science.*, ed A Childs and M Mosca (Springer)
[9] Reichardt B W 2009 Quantum universality by state distillation *Quantum Inf. Comput.* **9** 1030–52
[10] Anwar H, Campbell E T and Browne D E 2012 Qutrit magic state distillation *New J. Phys.* **14** 063006
[11] Bravyi S and Haah J 2012 Magic-state distillation with low overhead *Phys. Rev. A* **86** 052329
[12] Campbell E T, Anwar H and Browne D E 2012 Magic-state distillation in all prime dimensions using quantum reed-muller codes *Phys. Rev. X* **2** 041021
[13] Dawkins H and Howard M 2015 Qutrit magic state distillation tight in some directions *Phys. Rev. Lett.* **115** 030501
[14] Howard M and Dawkins H 2016 Small codes for magic state distillation *Eur. Phys. J.* **70** 55
[15] Haah J, Matthew B H, Poulin D and Wecker D 2017 Magic state distillation with low space overhead and optimal asymptotic input count *Quantum* **1** 31
[16] Prakash S 2020 Magic state distillation with the ternary golay code *R. Soc. A* vol **476** p 20200187
[17] Prakash S and Saha T 2025 Low overhead qutrit magic state distillation *Quantum* **9** 1768
[18] Wills A, Hsieh M-H and Yamasaki H 2025 Constant-overhead magic state distillation *Nat. Phys.* **21** 1842–6
[19] Prakash S and Singhal R 2026 Search for high-threshold qutrit magic-state distillation routines *Phys. Rev.* **113** 042404
[20] Reichardt B W 2005 Quantum universality from magic states distillation applied to CSS codes *Quantum Inf. Process.* **4** 251–64
[21] Veitch V, Mousavian A H, Gottesman D and Emerson J 2014 The resource theory of stabilizer quantum computation *New J. Phys.* **16** 013009
[22] Howard M and Campbell E 2017 Application of a resource theory for magic states to fault-tolerant quantum computing *Phys. Rev. Lett.* **118** 090501
[23] Mari A and Eisert J 2012 Positive Wigner functions render classical simulation of quantum computation efficient *Phys. Rev. Lett.* **109** 230503
[24] Veitch V, Ferrie C, Gross D and Emerson J 2012 Negative quasi-probability as a resource for quantum computation *New J. Phys.* **14** 113011
[25] Prakash S and Gupta A 2020 Contextual bound states for qudit magic state distillation *Phys. Rev. A* **101** 010303
[26] Hudson R L 1974 When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.* **6** 249–52
[27] Kenfack A and Życzkowski K 2004 Negativity of the Wigner function as an indicator of non-classicality *J. Opt. B: Quantum Semiclass. Opt.* **6** 396
[28] Galvão E F 2005 Discrete Wigner functions and quantum computational speedup *Phys. Rev. A* **71** 042302
[29] Cormick C, Galvão E F, Gottesman D, Paz J P and Pittenger A O 2006 Classicality in discrete Wigner functions *Phys. Rev. A* **73** 012301
[30] Gross D 2006 Hudson's theorem for finite-dimensional quantum systems *J. Math. Phys.* **47** 122107
[31] Gross D 2007 Non-negative Wigner functions in prime dimensions *Appl. Phys. B* **86** 367–70
[32] Spekkens R W 2008 Negativity and contextuality are equivalent notions of nonclassicality *Phys. Rev. Lett.* **101** 020401
[33] Howard M, Wallman J, Veitch V and Emerson J 2014 Contextuality supplies the magic for quantum computation *Nature* **510** 351–5
[34] Delfosse N, Okay C, Bermejo-Vega J, Browne D E and Raussendorf R 2017 Equivalence between contextuality and negativity of the Wigner function for qudits *New J. Phys.* **19** 123024

[35] Schmid D, Haoxing D, Selby J H and Pusey M F 2022 Uniqueness of noncontextual models for stabilizer subtheories *Phys. Rev. Lett.* **129** 120403
[36] Raussendorf R, Okay C, Zurel M and Feldmann P 2023 The role of cohomology in quantum computation with magic states *Quantum* **7** 979
[37] Campbell E T and Browne D E 2010 Bound states for magic state distillation in fault-tolerant quantum computation *Phys. Rev. Lett.* **104** 030503
[38] Rall P 2017 Signed quantum weight enumerators characterize qubit magic state distillation (arXiv:1702.06990)
[39] Kalra A R and Prakash S 2025 Invariant theory and magic state distillation (arXiv:2501.10163)
[40] MacWilliams F J 1962 Combinatorial problems of elementary abelian groups *PhD Thesis* Radcliffe College
[41] MacWilliams J 1963 A theorem on the distribution of weights in a systematic code *Bell Syst. Tech. J.* **42** 79–94
[42] MacWilliams F J and Sloane N J A 1977 *The Theory of Error-Correcting Codes. North Holland.* (Amsterdam)
[43] Vera S P, Huffman W C and Richard A B 1998 *An Introduction to Algebraic Codes.* (Elsevier)
[44] Dicuangco L, Moree P and Solé P 2007 The lengths of hermitian self-dual extended duadic codes *J. Pure Appl. Algebra* **209** 223–37
[45] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge University Press) 10th Anniversary edn
[46] Gottesman D 2024 Surviving as a quantum computer in a classical world
[47] Calderbank A R, Rains E M, Shor P M and Sloane N J A 1998 Quantum error correction via codes over gf(4) *IEEE Trans. Inf. Theory* **44** 1369–87
[48] Ashikhmin A and Knill E 2001 Nonbinary quantum stabilizer codes *IEEE Trans. Inf. Theory* **47** 3065–72
[49] Ketkar A, Klappenecker A, Kumar S and Sarvepalli P K 2006 Nonbinary stabilizer codes over finite fields *IEEE Trans. Inf. Theory* **52** 4892–914
[50] Rains E M 1999 Nonbinary quantum codes *IEEE Trans. Inf. Theory* **45** 1827–32
[51] Niven I, Zuckerman H S and Montgomery H L 1991 *An Introduction to the Theory of Numbers* (New York: Wiley)
[52] Mattson H F and Assmus E F 1964 Research program to extend the theory of weight distribution and related problems for cyclic error-correcting codes, report AFCRL-64-605 *Technical Report.* Air Force Cambridge Research Laboratories, Bedford, MA
[53] Blahut R E 1991 The gleason-prange theorem *IEEE Trans. Inf. Theory* **37** 1269–73
[54] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed-state entanglement and quantum error correction *Phys. Rev. A* **54** 3824–51
[55] Laflamme R, Miquel C, Paz J P and Zurek W H 1996 Perfect quantum error correcting code *Phys. Rev. Lett.* **77** 198–201
[56] Bosma W, Cannon J and Playoust C 1997 The Magma algebra system. I. The user language *J. Symb. Comput.* **24** 235–65
[57] Steane A 1996 Multiple-particle interference and quantum error correction *Proc. R. Soc. A* vol **452** pp 2551–77
[58] Dasu S and Burton S 2025 A classification of transversal clifford gates for qubit stabilizer codes (arXiv:2507.10519)
[59] Nebe G, Rains E M and Sloane N J A 2006 *Self-Dual Codes and Invariant Theory.* (Springer)
[60] Gaborit P, Nedeloaia C-S and Wassermann A 2005 On the weight enumerators of duadic and quadratic residue codes *IEEE Trans. Inf. Theory* **51** 402–7
[61] Jain A and Prakash S 2020 Qutrit and ququint magic states *Phys. Rev. A* **102** 042409
[62] Gross D 2008 Computational power of quantum many-body states and some results on discrete phase spaces *PhD thesis* Imperial College London (available at: www.thp.uni-koeln.de/gross/files/diss.pdf)
[63] Bravyi S, Smith G and Smolin J A 2016 Trading classical and quantum computational resources *Phys. Rev. X* **6** 021043
[64] Peres F C R and Galvão E F 2023 Quantum circuit compilation and hybrid computation using pauli-based computation *Quantum* **7** 1126
[65] Peres F C R 2023 Pauli-based model of quantum computation with higher-dimensional systems *Phys. Rev. A* **108** 032606
[66] Zurel M 2024 Classical descriptions of quantum computations: foundations of quantum computation via hidden variable models, quasiprobability representations, and classical simulation algorithms *PhD Thesis* University of British Columbia
[67] Zurel M, Okay C, Raussendorf R and Heimendahl A 2024 Hidden variable model for quantum computation with magic states on qudits of any dimension *Quantum* **8** 1323