

Low Overhead Qutrit Magic State Distillation

Shiroman Prakash¹ and Tanay Saha^{1,2}

¹Department of Physics and Computer Science, Dayalbagh Educational Institute, Agra-282005, India

²Present Address: Department of Mathematics, Simon Fraser University, Burnaby, B.C., Canada

We show that using qutrits rather than qubits leads to a substantial reduction in the overhead cost associated with an approach to fault-tolerant quantum computing known as magic state distillation. We construct a family of $[[9m - k, k, 2]]_3$ triorthogonal qutrit error-correcting codes for any positive integers m and k with $k \leq 3m - 2$ that are suitable for magic state distillation. In magic state distillation, the number of ancillae required to produce a magic state with target error rate ϵ is $O(\log^\gamma \epsilon^{-1})$, where the yield parameter γ characterizes the overhead cost. For $k = 3m - 2$, our codes have $\gamma = \log_2(2 + \frac{6}{3m-2})$, which tends to 1 as $m \rightarrow \infty$. Moreover, the $[[20, 7, 2]]_3$ qutrit code that arises from our construction when $m = 3$ already has a yield parameter of 1.51 which outperforms all known qubit triorthogonal codes of size less than a few hundred qubits.

1 Introduction

To realize their full potential, quantum computers must achieve large-scale fault tolerance. Magic state distillation [1, 2] is one of the leading approaches to fault-tolerant quantum computation, in which many noisy magic states are distilled into fewer pure magic states, which are then used to achieve quantum universality via state injection. There has been some encouraging experimental progress in magic state distillation in recent years, e.g., [3–7]. However, the overhead associated with magic state distillation – i.e., the number of noisy ancillae required to obtain an ancilla in a sufficiently pure magic state – is high, and represents a significant portion of the cost of quantum computing (as estimated in e.g., [8]), though this cost has decreased over the years (e.g., [9–12]). Reducing the overhead of magic state distillation is crucial in order to realize large-scale fault-tolerant quantum computation within the foreseeable future. At the core of any magic state distillation routine is a quantum error-correcting code with desirable theoretical properties that enable it to distill a few high-fidelity magic states from many low-fidelity magic states – in this paper, we present a new family of small error-correcting codes based on qutrits that achieve this goal with low overhead.

A magic state distillation routine takes n noisy ancillae as input and, with probability P , yields k pure ancillae as output with infidelity suppression $\epsilon_{\text{out}} \sim \epsilon_{\text{in}}^\nu$, where ν is the noise-suppression exponent. To obtain an ancilla with output error rate ϵ_{out} less than ϵ , one requires $O\left(\log^\gamma \frac{1}{\epsilon}\right)$ noisy ancillae [1, 13], where

$$\gamma = \log_\nu \left(\frac{n}{Pk} \right). \quad (1)$$

Shiroman Prakash: sprakash@dei.ac.in

Tanay Saha: tanays@iitk.ac.in

γ is known as the *yield parameter* and characterizes the overhead cost of magic state distillation in the asymptotic limit of many rounds of distillation.

Of the two magic state distillation routines originally proposed in [1], distillation via the 15-qubit Reed-Muller code has the lowest overhead, given by yield parameter of $\gamma = \log_3 15 \approx 2.46$. This code was later shown to be part of a larger family of codes known as *triorthogonal codes* [13]. For distillation via triorthogonal codes, it can be shown that the probability of successful distillation, $P \rightarrow 1$ as $\epsilon_{\text{in}} \rightarrow 0$, and that $\nu = d$, where d is the distance of the quantum error correcting code. For all other known magic state distillation routines, $P_{\text{success}} \leq 1$ and $\nu \leq d$. Therefore, in order to construct distillation routines with the lowest possible yield parameters, it is natural to focus attention on triorthogonal codes.

Triorthogonal codes have given rise to several reductions to the overhead cost of magic state distillation. In particular, [13] constructed a family of magic state distillation routines with yield parameter approaching $\gamma = \log_2 3 \approx 1.58$. Triorthogonal codes have been used to construct qubit magic state distillation routines with even lower yield parameters – but these all seem to require an inordinately large number of qubits. Notably, [14] showed that one can construct qubit triorthogonal codes for which $\gamma < 1$, however, the smallest of these codes requires a block size of 2^{58} qubits. Other families of qubit distillation routines which achieve a yield parameter that approaches 1 are known [15–17], however these also seem to require an unwieldy number of qubits.¹ Do there exist low-overhead magic state distillation routines involving a reasonably small number of qubits? A recent computational search [21] for qubit triorthogonal codes suggest that the answer is no.

An alternative approach to lower the overhead is to construct magic state distillation routines for *qudits* – quantum systems of prime dimension p – rather than qubits, which are also of purely theoretical interest due to the direct connection between distillability and contextuality present only for qudits of odd dimension [22–25]. [26] provided the first explicit examples of qudit magic state distillation routines, demonstrating, in particular, that $[[5, 1, 3]]_3$ code can distill an eigenstate of the qutrit Hadamard gate. However, the distillation routines proposed in [26] (which are based on codes that are not triorthogonal) have the disadvantage of having noise suppression exponent $\nu = 1$, and therefore infinite yield parameter γ . Subsequent constructions qudit magic state distillation routines include both triorthogonal codes [27–29] and non-triorthogonal² codes [30–32]. Of these, the lowest known overheads arise from punctured quantum Reed-Solomon codes, which achieve a yield parameter $\gamma = \log_{\lfloor (p+1)/3 \rfloor} (p-1)$ with a single puncture [28], and sublogarithmic yields with multiple punctures [29] for sufficiently large p . While such codes can have modest block sizes $\sim O(p)$, they require qudits of rather unwieldy dimension $p \gtrsim 23$ to obtain $\gamma \sim 1$.

Given these results, many expect that low-overhead magic distillation, i.e., $\gamma \rightarrow 1$, though possible in theory, is not realizable in practice, unless one is willing to work with qudits of large dimension, or codes with large block sizes. Here we show that this expectation is incorrect, by constructing a family of relatively small triorthogonal codes for qutrits that gives rise to a yield parameter approaching 1.

Our paper is organized as follows. In Section 2 we review the construction of triorthogonal codes, which are codes that possess a transversal non-Clifford gate. In Section 3, we present our new family of qutrit triorthogonal codes. In Section 4 (supported by the appendix) we study the noise reduction and compute the threshold of our new family of

¹After submission of this work for publication, several works[18–20] appeared which construct asymptotically good qubit distillation routines with $\gamma \rightarrow 0$.

²While distillation via non-triorthogonal codes involves higher overheads, such codes do appear to have higher thresholds to noise.

codes. In Section 5 we compute the yield parameter and also estimate the practical overhead cost of our family of codes when only a few rounds of distillation are required. In Section 6, we conclude with discussion and open questions.

2 Triorthogonal Codes

We first review the general framework for constructing triorthogonal quantum codes [13], adapted to qudits [29] of dimension p . First, one must construct a κ -dimensional self-orthogonal linear subspace \mathcal{T} of \mathbb{F}_p^n , such that, for any three vectors $h^{(a)}, h^{(b)}, h^{(c)} \in \mathcal{T}$, the triple product

$$\sum_{i=1}^n h_i^{(a)} h_i^{(b)} h_i^{(c)} = 0 \pmod{p}. \quad (2)$$

Such a space is known as a ternary *triorthogonal space* of size n and dimension κ .

We next puncture³ the triorthogonal space at k different coordinates to form a space generated by a $\kappa \times (n - k)$ dimensional matrix H , known as a triorthogonal matrix. We partition H as

$$H = \begin{pmatrix} H_1 \\ H_0 \end{pmatrix}, \quad (3)$$

where H_0 consists of those rows of H that are self-orthogonal, and H_1 consists of those rows of H that are not. H is called a triorthogonal matrix, and is used to construct a CSS code, known as a triorthogonal code, with stabilizers \mathcal{S}_x spanned by H_0 and \mathcal{S}_z spanned by H^\perp , and logical operators determined by the rows of H_1 . Any such code possesses a transversal gate from the third level of the Clifford hierarchy, which, for qutrits, can be taken to be [33]

$$T = \sum_k e^{2\pi i k/9} |k\rangle \langle k|. \quad (4)$$

This gate and its generalizations have been studied extensively, see, e.g., [34–38].

As a consequence of this transversal non-Clifford gate, a triorthogonal code can be used for magic state distillation of a particular qutrit magic state analogous to the $|H\rangle$ state of [1], (sometimes referred to as an equatorial magic state [39]), which we denoted as:

$$|M_0\rangle = \sum_j e^{2\pi i j/9} |j\rangle. \quad (5)$$

The noise suppression of the resulting routine is, schematically, $\epsilon_{\text{out}} \sim \epsilon_{\text{in}}^d$, where d is the code distance⁴ [1, 27]. Note that, because H_0 is self-orthogonal, $\mathcal{S}_x \subset \mathcal{S}_z$, and therefore, \mathcal{S}_x determines the distance of the code. For this reason, it is natural to also demand that the triorthogonal spaces \mathcal{T} be *maximal* – i.e., no additional basis vector can be added to \mathcal{T} while preserving triorthogonality. The reason for this is as follows. Suppose that there exist two triorthogonal spaces \mathcal{T}' and \mathcal{T} , such that $\mathcal{T}' \subset \mathcal{T}$ – puncturing \mathcal{T}' and \mathcal{T} at the same locations would give rise to stabilizers \mathcal{S}'_x and \mathcal{S}_x that satisfy $\mathcal{S}'_x \subset \mathcal{S}_x$. Therefore the triorthogonal code obtained from \mathcal{T}' would have distance $d' \leq d$, the distance of the code obtained from \mathcal{T} .

³Recall that puncturing a code at coordinate j means deleting the j th column of the generator matrix. Each time a code is punctured, n decreases by 1.

⁴Let us point out, when codes that are not triorthogonal are used for magic state distillation, the noise suppression exponent, ν defined via $\epsilon_{\text{out}} = \epsilon_{\text{in}}^\nu$, is not the same as the code distance d .

As an example, we present the well-known family of qubit triorthogonal codes were proposed in [13]. A representative member of this family correspond to the triorthogonal space $\mathcal{T}^{(2)}$ generated by the matrix $T^{(2)}$ given by

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \quad (6)$$

This can be punctured at coordinates 1 and 5 to give rise to a triorthogonal matrix of the form:

$$H = \begin{pmatrix} H_1 \\ H_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad (7)$$

that defines a $[[14, 2, 2]]_2$ triorthogonal code with yield parameter $\gamma = \log_2 7 \approx 2.81$.

3 Our Construction

We define a triorthogonal space denoted as \mathcal{T}_m of size $n = 9m$ and dimension $3m$, spanned by basis vectors w and $\{v^{(1)}, v^{(2)}, \dots, v^{(3m-1)}\}$ given by

$$w = (0, 1, 2, 0, 1, 2, \dots) \quad (8)$$

and

$$v_i^{(a)} = \begin{cases} 0 & 0 < i \leq 3(a-1), \\ 1 & 3(a-1) < i \leq 3a, \\ 0 & 3a < i \leq n-3, \\ 2 & n-3 < i \leq n. \end{cases} \quad (9)$$

By direct computation, it is easy to check that \mathcal{T}_m is triorthogonal and maximal.

Explicitly, for $m = 2$, the generator matrix for \mathcal{T}_m is given by

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}. \quad (10)$$

The 9-qutrit triorthogonal space that arises from this construction when $m = 1$ coincides with a linear Reed-Muller code in two variables, which can be punctured to give rise to the $[[8, 1, 2]]_3$ code used for distillation in [27]. The 27-qutrit triorthogonal space that arises from this construction when $m = 3$ contains the linear Reed-Muller code in three variables as a proper subspace.

One can puncture \mathcal{T}_m once per block in any $3m - 2$ blocks of three trits to obtain a $[[9m - k, k, 2]]_3$ CSS code for $k \leq 3m - 2$. Explicitly, let us choose to puncture at

coordinates $3j + 1$, for all $0 \leq j < 3m - 2$. For $m = 2$, this leads to a triorthogonal matrix H given by

$$H = \begin{pmatrix} H_1 \\ H_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 2 & 2 & 2 \\ 1 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}. \quad (11)$$

For any m , the two stabilizers in \mathcal{S}_x will be of the form $(1, 2, 1, 2, \dots, 0, 1, 2, 0, 1, 2)$ and $(0, \dots, 0, 1, 1, 1, 2, 2, 2)$, which allow us to detect, but not correct, a single Z -error, leading us to conclude that the code has distance 2.

4 Threshold

Let us first compute the threshold to noise for these distillation routines. Note that, although the threshold does not affect the computation of the yield parameter γ , it does have a direct impact on the actual overhead cost of distillation in the limit when only a few rounds of distillation are used, as is discussed in the next section.

To describe a general mixed qutrit state, one requires 8 independent noise parameters. However, after twirling, as described in [27] and [39], noisy qutrit $|M_0\rangle$ magic states can be expressed in terms of only two noise parameters, ϵ_1 and ϵ_2 , via

$$\rho(\epsilon_1, \epsilon_2) = (1 - \epsilon_1 - \epsilon_2) |M_0\rangle \langle M_0| + \epsilon_1 |M_1\rangle \langle M_1| + \epsilon_2 |M_2\rangle \langle M_2|, \quad (12)$$

where $|M_j\rangle = Z^j |M_0\rangle$. The performance of a qutrit magic state distillation routine is thus characterized by two functions $\epsilon'_1(\epsilon_1, \epsilon_2)$ and $\epsilon'_2(\epsilon_1, \epsilon_2)$, which determine a basin of attraction in the space of all convex combinations of the $|M_i\rangle \langle M_i|$ that distill to the desired magic state.

To obtain the highest thresholds (i.e., largest basins of attraction) possible using our construction, we set $k = 1$. We computed both $\epsilon'_1(\epsilon_1, \epsilon_2)$ and $\epsilon'_2(\epsilon_1, \epsilon_2)$ for small values of m following [27], and found that $m = 2$ gives the largest basin of attraction. Details of this computation are provided in Appendix A. Explicitly, up to order, ϵ_i^3 , we found, for $m = 2$,

$$\epsilon'_1(\epsilon_1, \epsilon_2) = \epsilon_1^2 + 2 \left(\epsilon_1^3 + \epsilon_2 \epsilon_1^2 + 15 \epsilon_2^2 \epsilon_1 \right) + O(\epsilon_i^4) \quad (13)$$

$$\epsilon'_2(\epsilon_1, \epsilon_2) = \epsilon_2^2 + 2 \left(\epsilon_2^3 + \epsilon_1 \epsilon_2^2 + 15 \epsilon_1^2 \epsilon_2 \right) + O(\epsilon_i^4). \quad (14)$$

We plot the basins of attraction of the $[[17, 1, 2]]_3$ code in the two-dimensional space of qutrit state space in Figure 1. As can be seen from this figure, the region that distills to $|M_0\rangle$ extends fairly close to the Wigner polytope [40, 41] which sets the theoretical bound for the best possible magic state distillation routine, and is tied to the onset of (state-dependent) contextuality [22, 25].

To condense the information contained in the basin of attraction plotted in Figure 1, it is often customary to assume depolarizing noise, by demanding that $\epsilon_1 = \epsilon_2 = \delta/3$. This allows us to characterize the distillation performance by a single function $\delta'(\delta)$, and the basin of attraction by a single inequality, $\delta < \delta_*$, where δ_* is the threshold to depolarizing noise. We found that, for $m \geq 2$ and $k = 1$,

$$\delta' = \frac{1}{3} \delta^2 + \frac{8 + 9(m-1)}{9} 2\delta^3 + \frac{41 + 54(m-1)}{27} 2\delta^4 + O(\delta^5). \quad (15)$$

The thresholds to depolarizing noise for codes with $k = 1$ and $m < 8$ are plotted in Figure 2. We find that the $[[17, 1, 2]]_3$ code obtained from our construction when $m = 2$ and $k = 1$ has the highest depolarizing noise threshold, $\delta_* = 0.353$, exceeding that of the $[[8, 1, 2]]_3$ punctured Reed-Muller code of [27], whose noise threshold is $\delta_* = 0.317$.⁵ The bound for the best possible threshold to depolarizing noise determined by the Wigner polytope is $\delta_*^{\max} = 0.467$.

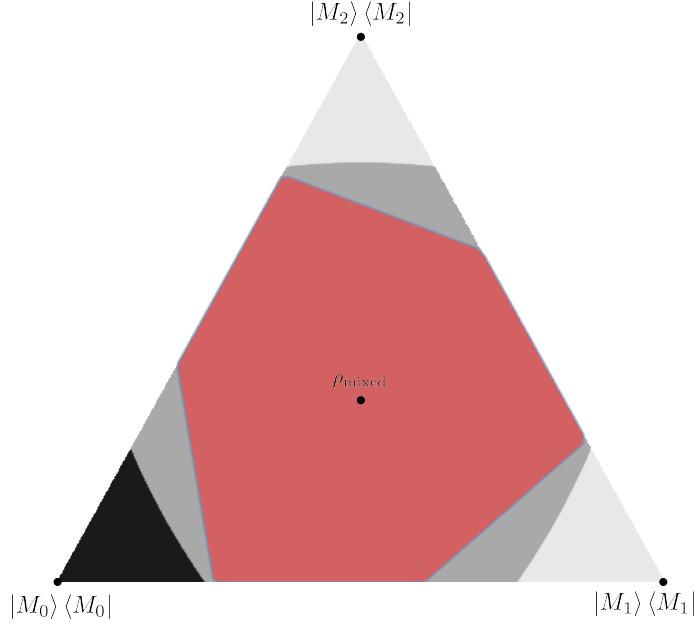


Figure 1: Any noisy magic state can be twirled to lie in the triangle spanned by convex combinations of $|M_i\rangle$, pictured above. The region of state space that distills to the magic state $|M_0\rangle$ via the $[[17, 1, 2]]_3$ code is shown in black. The light gray regions distill to magic states $|M_1\rangle$ and $|M_2\rangle$. The red and dark gray regions distill to the maximally mixed state, at the centre of the triangle. (The red region is inside the Wigner polytope, and cannot be distilled to a useful magic state by any code.)

Let us also discuss the case of codes with $k > 1$, restricting to the case of depolarizing noise for simplicity. We find that, for the $[[9m - k, k, 2]]_3$ code, each individual output qutrit has depolarizing noise δ' given by

$$\delta' = \frac{A_2(k, m)}{6} \delta^2 + O(\delta^3) \quad (16)$$

where

$$A_2 = \begin{cases} 2k^2 & k < 3m - 2 \\ 2k^2 + 6 & k = 3m - 2 \end{cases} \quad (17)$$

is the number of logical Z or Z^2 operators of weight 2 that act on any one individual output qutrit. For the case of $k = 3m - 2$, A_2 is computed in the Appendix; the case $k < 3m - 2$ can be determined in a similar manner. We see that the threshold for large k is $\delta_* \approx \frac{6}{A_2} \sim \frac{3}{k^2}$. The threshold decreases quadratically with $1/k$ rather than linearly, as it does for the case of qubits [13].

⁵For distillation of certain other qutrit magic states, slightly higher thresholds are possible using the 11-qutrit Golay code [32] or [30], but at a cost of a much lower yield.

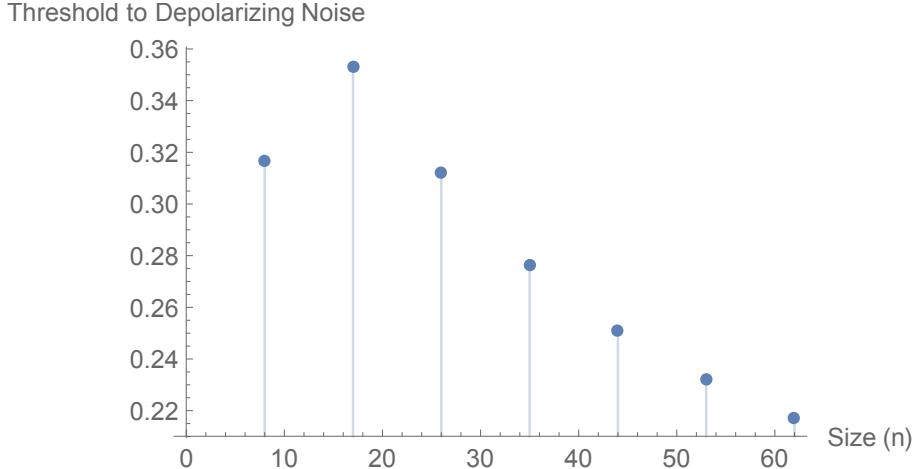


Figure 2: The threshold to depolarizing noise for distillation via the $[[9m-1, 1, 2]]_3$ codes, as a function of code size. The $[[17, 1, 2]]_3$ code has the highest threshold, $\delta_* = 0.353$.

The success probability is given by

$$P_s(k, m) = 1 - \frac{18m - 2k}{3}\delta + O(\delta^2). \quad (18)$$

5 Overhead

How should one measure the overhead of a magic state distillation protocol? For both qubits and qudits, Magic state distillation, as proposed in [1, 26, 27], starts from the assumption that one already has a fault-tolerant quantum computer able to implement Clifford unitaries and prepare/measure stabilizer states without any error⁶. Following [1, 13], we define the (theoretical) overhead cost of magic state distillation simply as the number of low-fidelity magic states needed to produce a high-fidelity magic state, assuming free and perfect Clifford operations. Explicitly, if N_{in} with noise rate δ_{in} are needed to produce N_{out} magic states with noise rate δ_{out} , following [13, 15, 45], we define the cost of distillation to be

$$C(\delta_{\text{in}}, \delta_{\text{out}}) = \frac{N_{\text{in}}}{N_{\text{out}}}. \quad (19)$$

Is it possible to compare fault-tolerant schemes for qutrits to qubits? In general there are more noise channels for qutrits, and a noisy qutrit may be harder to produce than a noisy qubit. One might model these differences by introducing “conversion constants” $K > 1$ and $K' > 1$, such that $n_{\text{qutrit}} = Kn_{\text{qubit}}$ and $\epsilon_{\text{qutrit}} = K'\epsilon_{\text{qubit}}$. However, the yield parameter γ measures how the number of noisy magic states scales with the logarithm of the target error rate in the asymptotic limit of many rounds of distillation, and is clearly independent of K and K' – therefore it appears meaningful to compare yield parameters across qudits of different dimensionalities. In subsection 5.1 we compute the yield parameter for our codes and compare to yield parameters from similar qubit codes.

In practice, however, only a few rounds of distillation may be needed, so the yield parameter is not necessarily meaningful. We study the distillation cost as a function of target noise rate in section 5.2.

⁶Qudit generalizations of the color code and surface code that allow for fault-tolerant implementations of the qudit Clifford group are discussed in [42–44].

The distillation cost is a very natural measure of overhead cost associated with magic state distillation – indeed, the entire resource theory of magic (e.g., [22, 40, 41, 46–49]) seeks to place theoretical bounds on this measure. However, in practice, the cost of fault-tolerant implementation of the Clifford group must also be taken into consideration when estimating the total overhead cost of quantum computing via magic states, and this requires a more careful study of the circuits used for implementing the distillation protocol. While we expect that distillation with codes with low theoretical overhead cost may give rise to lower overheads in practice, a detailed estimate of the total overhead cost of fault-tolerant quantum computation via magic state distillation using the new family of qutrit error-correcting codes proposed here is beyond the scope of the present work.⁷

5.1 Yield Parameter

Let us first compute the yield parameter, which, captures the overhead cost in the limit of many rounds of distillation.

To obtain triorthogonal codes with the maximal yield parameter, we set $k = 3m - 2$, which is the maximum value for which the code has distance 2. The resulting $[[6m + 2, 3m - 2, 2]]_3$ triorthogonal codes have yield parameter given by,

$$\gamma = \log_2\left(2 + \frac{6}{3m - 2}\right), \quad (20)$$

which approaches 1 as $m \rightarrow \infty$. The yield parameter as a function of code size n is plotted in Figure 3. The advantages over qubit codes of comparable size is striking. Let us explicitly list the yields for the smallest codes in our construction, obtained by setting $k = 3m - 2$.

- For $m = 2$, we obtain a $[[14, 4, 2]]_3$ code with a yield parameter of 1.81, which is better than the best known qubit triorthogonal codes of size less than 50 [13, 21].
- For $m = 3$, we obtain a $[[20, 7, 2]]_3$ code, whose yield parameter is $1.51 < \log_2 3$, outperforming all the codes constructed in [13], which, to our knowledge, are the best known qubit triorthogonal codes with block size less than a few hundred qubits.
- For $m = 4$, we obtain a $[[26, 10, 2]]_3$ code, whose yield parameter is 1.38.
- For $m = 5$, we obtain a $[[32, 13, 2]]_3$ code, whose yield parameter is 1.30.
- For $m = 6$, we obtain a $[[38, 16, 2]]_3$ code, whose yield parameter is 1.25.
- For $m = 7$, we obtain a $[[44, 19, 2]]_3$ code, whose yield parameter is 1.21.
- For $m = 8$, we obtain a $[[50, 22, 2]]_3$ code with yield parameter 1.18. To the best of our knowledge, to get a comparable yield parameter, one requires a code with ~ 1000 qubits, using triorthogonal codes constructed from punctured binary Reed-Muller codes in [14, 17].

Note also that previous constructions of qudit-based triorthogonal codes in [28, 29] require qudits of dimension at least 17 to obtain comparable improvements in yields.

⁷Indeed, a variety of creative techniques and optimizations have been proposed in the literature for magic state preparation and distillation that can greatly reduce the total overhead cost in practice. Other techniques, such as combining magic state distillation with gate synthesis [9] can further reduce the overhead cost for fault-tolerant implementation of a given algorithm. All these ideas can, presumably, also be applied to qutrits, but the details have not been worked out in the literature.

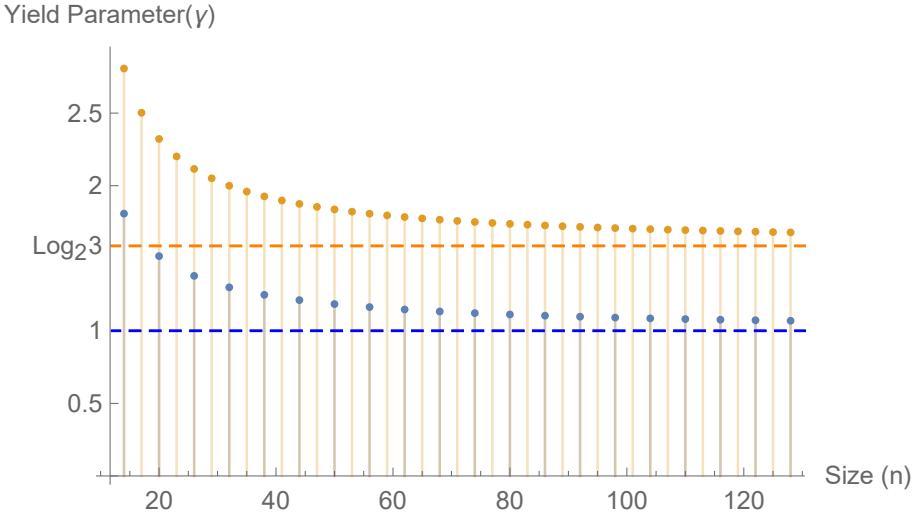


Figure 3: Yield parameters γ of the lowest overhead $[[n, k, 2]]_3$ triorthogonal code produced by our construction are shown in blue, as a function of code size n . As $n \rightarrow \infty$, $\gamma \rightarrow 1$, shown by the blue line. For comparison, yield parameters for binary $[[n, k, 2]]_2$ triorthogonal codes in [13] with the same n are plotted in orange on the same graph – for these codes $\gamma \rightarrow \log_2 3$ as $n \rightarrow \infty$, shown by the orange dashed line.

5.2 Practical Estimates

In this section, we study the cost of magic state distillation when only a few rounds of distillation are needed, which is, perhaps, more relevant in practice than the yield parameter γ .

To see that our codes may still offer an advantage, let us compare the distillation cost using our family of codes to distillation using the $[[15, 1, 3]]_2$ code of [1]. Of course, it is not clear how to fairly compare qubit noise to qudit noise – following [28] for simplicity, we work with the depolarizing noise δ . Expressed in terms of depolarizing noise, the noise reduction of the $[[15, 1, 3]]_2$ code for distillation is

$$\delta' = \frac{35}{4} \delta^3 + O(\delta^4). \quad (21)$$

The probability of successful projection is $P_s = 1 - \frac{15}{2} \delta + O(\delta^2)$.

Using the $[[15, 1, 3]]_2$ code, starting with a $\delta_{\text{in}} = .001$, a single round of distillation would give $\delta_{\text{out}} \approx 10^{-8}$, with a distillation cost of 15.2. Now consider distillation with the $[[32, 13, 2]]_3$ code from our family of codes. Starting with $\delta_{\text{in}} = .001$, after three rounds of distillation with the $[[32, 13, 2]]_3$ code, using Eq. (16) one finds, $\delta_{\text{out}} \approx \left(\frac{A_2(5, 13)}{6}\right)^7 \delta_{\text{in}}^8 \approx 2 \times 10^{-12}$, with a distillation cost of $C \approx \left(\frac{32}{13}\right)^3 / P_s(.001) \approx 15.2$. Although depolarizing noise for qubits may not be directly comparable to depolarizing noise for qudits, for the same distillation cost, qudits with $\delta \sim 10^{-12}$ are presumably preferable to qubits with $\delta \sim 10^{-8}$.

One also has the freedom to choose a different code (i.e., a code with a different value of m) during each round, as described in [13, 45]. Suppose we distill with an $[[n_1, k_1, 2]]_3$ code on round 1, an $[[n_2, k_2, 2]]_3$ on round 2 and an $[[n_3, k_3, 2]]_3$ on round 3. Assuming $\delta_{\text{in}} \ll A_2/6$, we have,

$$\delta_{\text{out}} \approx A_2(k_3) A_2(k_2)^2 A_2(k_1)^4 \frac{\delta_{\text{in}}^8}{6^7}, \quad (22)$$

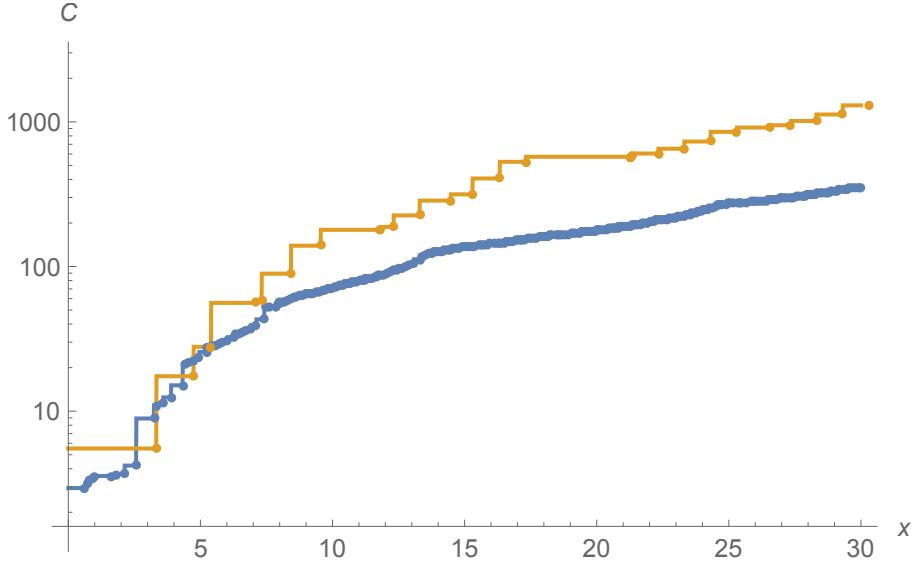


Figure 4: We plot the distillation cost $C = \frac{N_{\text{in}}}{N_{\text{out}}}$ needed to produce magic states a target depolarizing noise rate of $\delta_{\text{out}} = 10^{-x}$, starting from $\delta_{\text{in}} = 0.02$ for optimal distillation sequences of from our family of qutrit triorthogonal codes in blue. Though a comparison may not be fair, we also plot, in orange, similar results for optimal sequences of small qubit triorthogonal codes ($n \leq 128$), taken from [13]. (These optimal sequences typically involve one round of distillation with the $[[15, 1, 3]]_2$ code, followed by additional rounds with other triorthogonal codes with higher rates.)

with cost of distillation

$$C \approx \left(\frac{n_3}{k_3} \right) \left(\frac{n_2}{k_2} \right) \left(\frac{n_1}{k_1} \right). \quad (23)$$

Because triorthogonal codes with larger k typically have better rates, but smaller thresholds, there is some optimization required in the choice of distillation sequence. Heuristically, it appears advantageous to use distillation protocols with smaller k for early rounds of distillation, and protocols with larger k for later rounds.

Using a randomized search we determined optimal sequences for qutrit distillation routines using our family of codes. (The optimal sequences were computed subject to the following simplifying assumptions: for any m , we always choose $k = k_{\max}$; we restricted to codes with $m \leq 12$; we restricted the total rounds of distillation to be 5 or less.) The results are shown in Figure 4, which presents the optimal cost C as a function of target depolarizing noise rate $\delta_{\text{target}} = 10^{-x}$. For comparison, we plot distillation costs arising from optimal sequences of distillation using small qubit triorthogonal codes (including $[[15, 1, 3]]_2$ code, the $[[10, 2, 2]]_2$ code of [45] and the block $[[3k + 8, k, 2]]_2$ codes in [13]), which were computed in [13] for $\delta_{\text{in}} = 0.02$. The plot suggests an advantage in overhead costs persists for qutrits for target noise rates better than 10^{-10} .

Do these advantages hold up in practice, when one takes into account the cost and imperfections of Clifford operations? While we leave this for future work, let us make a few comments about it below. The $[[9m - k, k, 2]]_3$ qutrit triorthogonal codes constructed in this paper offer similar noise reduction to $[[3k + 8, k, 2]]_2$ block triorthogonal codes of [13] but have better rates. Overhead costs of distillation with the $[[3k + 8, k, 2]]_2$ block codes have been compared to the $[[15, 1, 3]]_2$ code, with mixed results [8, 50]. At least in noise regimes where distillation via $[[3k + 8, k, 2]]_2$ block codes is favored, our qutrit codes have a chance to give rise to lower overhead costs in practice, as evidenced by the computations in this section. However, for distillation with qutrits, no small triorthogonal

code with distance 3, analogous to the $[[15, 1, 3]]_2$ code, is known at present. Therefore, unless such a code is discovered, it appears that, in order to distill qutrit magic states with low overhead, one must make do with the family of codes presented in this paper.

6 Discussion

Several authors [27–29] have noticed theoretical advantages in the overhead cost of magic state distillation with qudits rather than qubits. However, none of these constructions provided any overhead reductions for qutrits. Indeed, prior to this work, the only practically useful code for qutrit distillation was the $[[8, 1, 2]]_3$ code of [27] – which does not appear to offer any substantial advantage over the $[[15, 1, 3]]_2$ qubit code of [1].⁸ We provided a simple construction of a new family of qutrit triorthogonal codes that have yield parameter $\gamma \rightarrow 1$ in the limit $m \rightarrow \infty$. The yield parameters of even the smallest codes in our family, involving as few as 20 qutrits, are substantially better than all known constructions involving hundreds of qubits or fewer. Even when only a few rounds of distillation are needed, these codes can outperform the $[[15, 1, 3]]_2$ code, or codes of [13], in terms of overhead cost, assuming perfect and free Clifford unitaries and stabilizer measurements.

This work indicates that the space of qutrit error-correcting codes remains relatively unexplored. Another natural question is whether there exist other qutrit triorthogonal codes with better yields, or higher distances? In particular, what is the smallest qutrit triorthogonal code with distance 3? We carried out an preliminary computer search over triorthogonal spaces of small dimension. The only non-trivial triorthogonal spaces of size $n \leq 21$ and dimension $\kappa \geq \lfloor n/3 \rfloor$ correspond to the spaces produced by our construction with $m = 1$ and $m = 2$. We therefore believe that the codes presented in this paper are somewhat distinguished, at least amongst very small qutrit triorthogonal codes.

The existence of the family of codes presented here demonstrates a concrete theoretical advantage to pursuing fault-tolerant quantum computation with qutrits rather than qubits. A number of promising experimental realizations of qutrit-based quantum computers exist (e.g., [51–60]), and if one has already made the decision to work with qutrits, the magic state distillation routines presented in this paper appear to be the best currently available. However, if one wishes to assess whether or not the possible advantages for large-scale fault-tolerance in the long term outweigh the increased complexities associated with control of qutrits in the short-term, more work needs to be done. In particular, it is important to know how well do these theoretical advantages hold up in practice, when the cost of (imperfect) Clifford unitaries is taken into account.

Acknowledgements

SP thanks Prof. P.S. Satsangi for inspiration, guidance, and emphasizing to him the importance of qudits of odd prime dimension. The authors also thank Noah Goss, Prahladh Harsha, Rishabh Kothary, Jeongwan Haah and Luca Giuzzi for discussions on related topics. Some of the computations in this paper were performed with assistance of the software package MAGMA [61]. The authors acknowledge the support of MeitY QCAL, Amazon Braket and DST-SERB grant (CRG/2021/009137).

⁸Other qutrit distillation protocols, such as those of [26, 30, 32], while theoretically important, give rise to distillation routines with impractically high overhead costs.

A Threshold Computations

A.1 General Analysis

In this Appendix, we directly extend the analysis of section 4 of [13] to qutrits. A similar analysis, restricted to $k = 1$, appears in [27].

The magic state we wish to distill is $|M_0\rangle = T|+\rangle$, where $|+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle)$ and T is the qudit version of the $\pi/8$ gate defined in [27, 33]. A triorthogonal code is a CSS-code with a transversal T gate, constructed from a triorthogonal matrix, of the form given in Eq. (3). Let $\mathcal{H}_0 = \text{span}(H_0)$ and $\mathcal{H} = \text{span}(H)$. The X -stabilizers are given by $\mathcal{S}_x = \mathcal{H}_0$ and the Z -stabilizers are given by \mathcal{H}^\perp . The rows of H_1 , which we denote by $f^{(a)}$ for $a \in \{1, \dots, k\}$, define the logical Z operators of the code as in [13] via

$$\bar{Z}^{(a)} = Z^{f_1^{(a)}} \otimes Z^{f_2^{(a)}} \otimes \dots \otimes Z^{f_n^{(a)}}. \quad (24)$$

A twirled noisy magic state can be written in the form,

$$\rho(\epsilon_1, \epsilon_2) = (1 - \epsilon_1 - \epsilon_2) |M_0\rangle \langle M_0| + \epsilon_1 |M_1\rangle \langle M_1| + \epsilon_2 |M_2\rangle \langle M_2|, \quad (25)$$

where $|M_j\rangle = Z^j |M_0\rangle$. As in the case of qubits, all noise is in the form of Z errors on input qutrits, which translate into Z errors on the output magic state. However, unlike the case of qubits, both Z and Z^2 errors may occur with *a priori* independent probabilities – ϵ_1 is the probability of a Z error and ϵ_2 is the probability of a Z^2 error. For simplicity, it is often convenient to assume depolarizing noise, in which case we can set $\epsilon_1 = \epsilon_2 = \frac{\delta}{3}$.

The performance of the magic state distillation routine via a triorthogonal code can be expressed in terms of weight enumerators \mathcal{S}_x and \mathcal{S}_z and their duals. However, unlike the case of qubits, one must work with complete weight enumerators (e.g., [62]), rather than simple weight enumerators. To this end, given a ternary vector $u \in F_3^n$ we define $|u|_1$ to be the number of entries of u equal to 1, and $|u|_2$ to be the number of entries of u equal to 2. The Hamming weight of u is $|u| = |u|_1 + |u|_2$. For example, if $u = (1, 2, 2, 0, 0, 0, 2, 2)$, $|u|_1 = 1$, $|u|_2 = 4$ and $|u| = 5$.

Projecting $\rho(\epsilon_1, \epsilon_2)^{\otimes n}$ onto the codespace of a triorthogonal code is successful with probability,

$$P_s = \sum_{u \in \mathcal{H}_0^\perp} (1 - \epsilon_1 - \epsilon_2)^{n - |u|_1 - |u|_2} \epsilon_1^{|u|_1} \epsilon_2^{|u|_2}. \quad (26)$$

If we restrict to depolarizing noise, we have

$$P_s \sum_{u \in \mathcal{H}_0^\perp} \left(1 - \frac{2\delta}{3}\right)^{n - |u|} \left(\frac{\delta}{3}\right)^{|u|} = \frac{1}{|\mathcal{H}_0|} \sum_{f \in \mathcal{H}_0} (1 - \delta)^{|u|}, \quad (27)$$

where we used the MacWilliams identity for the simple weight enumerator of a classical ternary code [62] in the last equality.

Our code encodes k logical qutrits. After projection and decoding, the output state is

$$\rho_{\text{out}} = \sum_{x \in F_3^k} P_{\text{out}}(x) |M_x\rangle \langle M_x| \quad (28)$$

where $|M_x\rangle = |M_{x_1}\rangle \otimes |M_{x_2}\rangle \otimes \dots \otimes |M_{x_k}\rangle$, and

$$P_{\text{out}}(x) = \sum_{f \in \mathcal{S}_Z + x_1 f^{(1)} + \dots + x_k f^{(k)}} (1 - \epsilon_1 - \epsilon_2)^{1 - |f|} \epsilon_1^{|f|_1} \epsilon_2^{|f|_2}. \quad (29)$$

The reduced density matrix describing the a th output qutrit can be written as,

$$\rho = (1 - \bar{\epsilon}_1^{(a)} - \bar{\epsilon}_2^{(a)}) |M_0\rangle \langle M_0| + \bar{\epsilon}_1^{(a)} |M_1\rangle \langle M_1| + \bar{\epsilon}_2^{(a)} |M_2\rangle \langle M_2| \quad (30)$$

where

$$\bar{\epsilon}_j^{(a)} = \frac{1}{P_s} \sum_{x: x_a=j} P_{\text{out}}(x). \quad (31)$$

If we restrict to depolarizing noise, demanding that $\bar{\epsilon}_1^{(a)} = \bar{\epsilon}_2^{(a)} = \frac{\bar{\delta}^{(a)}}{3}$, then one can write,

$$\begin{aligned} \frac{2}{3} \bar{\delta}^{(a)} &= 1 - \frac{1}{P_s} \sum_{x: x_a=0} P_{\text{out}}(x) \\ &= 1 - \frac{\sum_{u \in (H_0 \oplus f^{(a)})^\perp} (1 - \frac{2\delta}{3})^{n-|u|} \left(\frac{\delta}{3}\right)^{|u|}}{\sum_{u \in H_0^\perp} (1 - \frac{2\delta}{3})^{n-|u|} \left(\frac{\delta}{3}\right)^{|u|}}. \end{aligned} \quad (32)$$

Note that, again using the ternary MacWilliams identity, one can write

$$\sum_{u \in (H_0 \oplus f^{(a)})^\perp} (1 - \frac{2\delta}{3})^{n-|u|} \left(\frac{\delta}{3}\right)^{|u|} = \frac{1}{|\mathcal{H}_0 \oplus f^{(a)}|} \sum_{u \in \mathcal{H}_0 \oplus f^{(a)}} (1 - \delta)^{|u|}. \quad (33)$$

Thus the performance of triorthogonal qutrit codes can be characterized in terms of the classical weight enumerators of \mathcal{H}_0 and $\mathcal{H}_0 \oplus f^{(a)}$. Using these formulae one can reproduce the various results given in the main text.

A.2 Our codes

As an example, let us illustrate how the formulae from the previous subsection apply to our codes, for the case of maximum number of punctures $k = 3m - 2$. For this case, H_0 consists of two rows, and is of the form:

$$H_0 = \begin{pmatrix} 1 & 2 & 1 & 2 & \dots & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}. \quad (34)$$

where $\begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ repeats k times.

The success probability is expressed in terms of the simple weight enumerator of the classical ternary code generated by H_0 , and is easily seen to be,

$$\begin{aligned} P_s &= \frac{1}{9} \sum_{f \in \mathcal{H}_0} (1 - \delta)^{|u|} \\ &= \frac{1}{9} \left(1 + 2(1 - \delta)^6 + 6(1 - \delta)^{2k+4} \right), \\ &\approx 1 - \left(4 + \frac{4k}{3} \right) \delta + \frac{2}{3} \left(11 + 7k + 2k^2 \right) \delta^2 + O(\delta^3). \end{aligned} \quad (35)$$

For any choice of a , $\mathcal{H}_0 \oplus f^{(a)}$ is generated by,

$$H_0 \oplus f^{(a)} = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 2 & \dots & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \end{pmatrix}. \quad (36)$$

The numerator of the fraction in Eq. (32) is determined by the simple classical weight enumerator of the classical code determined by $H_0 \oplus f^{(a)}$ and is seen to be,

$$\begin{aligned} & \frac{1}{27} \sum_{f \in H_0 \oplus f^a} (1 - \delta)^{|u|} \\ &= \frac{1}{27} \left(1 + 2(1 - \delta)^6 + 12(1 - \delta)^{4+k} + 4(1 - \delta)^{3+2k} + 6(1 - \delta)^{2k+4} + 2(1 - \delta)^{6+2k} \right). \end{aligned} \quad (37)$$

Putting these together, we find,

$$\bar{\delta}^{(a)} \approx \left(1 + \frac{k^2}{3} \right) \delta^2 + \left(2 + 2k + \frac{k^2}{3} + \frac{k^3}{9} \right) \delta^3 + O(\delta^4). \quad (38)$$

References

- [1] Sergey Bravyi and Alexei Kitaev. “Universal quantum computation with ideal Clifford gates and noisy ancillas”. *Phys. Rev. A* **71**, 022316 (2005).
- [2] E. Knill. “Fault-tolerant postselected quantum computation: Schemes” (2004). [arXiv:quant-ph/0402171](https://arxiv.org/abs/quant-ph/0402171).
- [3] Alexandre M. Souza et al. “Experimental magic state distillation for fault-tolerant quantum computing”. *Nature Commun.* **2**, 1–5 (2011).
- [4] Lukas Postler et al. “Demonstration of fault-tolerant universal quantum gate operations”. *Nature* **605**, 675–680 (2022).
- [5] Yangsen Ye et al. “Logical magic state preparation with fidelity beyond the distillation threshold on a superconducting quantum processor”. *Phys. Rev. Lett.* **131**, 210603 (2023).
- [6] Riddhi S. Gupta et al. “Encoding a magic state with beyond break-even fidelity”. *Nature* **625**, 259–263 (2024).
- [7] Pedro Sales Rodriguez et al. “Experimental demonstration of logical magic state distillation” (2024). [arXiv:2412.15165](https://arxiv.org/abs/2412.15165).
- [8] Joe O’Gorman and Earl T. Campbell. “Quantum computation with realistic magic-state factories”. *Phys. Rev. A* **95**, 032338 (2017).
- [9] Earl T. Campbell and Mark Howard. “Unifying gate synthesis and magic state distillation”. *Phys. Rev. Lett.* **118**, 060501 (2017).
- [10] Daniel Litinski. “Magic state distillation: Not as costly as you think”. *Quantum* **3**, 205 (2019).
- [11] Christopher Chamberland and Kyungjoo Noh. “Very low overhead fault-tolerant magic state preparation using redundant ancilla encoding and flag qubits”. *npj Quantum Information* **6**, 91 (2020).
- [12] Héctor Bombín, Mihir Pant, Sam Roberts, and Karthik I Seetharam. “Fault-tolerant postselection for low-overhead magic state preparation”. *PRX Quantum* **5**, 010302 (2024).
- [13] Sergey Bravyi and Jeongwan Haah. “Magic-state distillation with low overhead”. *Phys. Rev. A* **86**, 052329 (2012).
- [14] Matthew B. Hastings and Jeongwan Haah. “Distillation with sublogarithmic overhead”. *Phys. Rev. Lett.* **120**, 050504 (2018).
- [15] Cody Jones. “Multilevel distillation of magic states for quantum computing”. *Phys. Rev. A* **87**, 042305 (2013).

[16] Jeongwan Haah, Matthew B. Hastings, D. Poulin, and D. Wecker. “Magic state distillation with low space overhead and optimal asymptotic input count”. *Quantum* **1**, 31 (2017).

[17] Jeongwan Haah and Matthew B. Hastings. “Codes and Protocols for Distilling T , controlled- S , and Toffoli Gates”. *Quantum* **2**, 71 (2018).

[18] Adam Wills, Min-Hsiu Hsieh, and Hayata Yamasaki. “Constant-overhead magic state distillation” (2024). [arXiv:2408.07764](https://arxiv.org/abs/2408.07764).

[19] Louis Golowich and Venkatesan Guruswami. “Asymptotically good quantum codes with transversal non-Clifford gates” (2024). [arXiv:2408.09254](https://arxiv.org/abs/2408.09254).

[20] Quynh T. Nguyen. “Good binary quantum codes with transversal CCZ gate” (2024). [arXiv:2408.10140](https://arxiv.org/abs/2408.10140).

[21] Sepehr Nezami and Jeongwan Haah. “Classification of small triorthogonal codes”. *Phys. Rev. A* **106**, 012437 (2022).

[22] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. “Contextuality supplies the ‘magic’ for quantum computation”. *Nature* **510**, 351–355 (2014).

[23] Dayal Pyari Srivastava, Vishal Sahni, and Prem Saran Satsangi. “Modelling microtubules in the brain as n-qudit quantum Hopfield network and beyond”. *Int. J. Gen. Syst.* **45**, 41–54 (2016).

[24] Dayal Pyari Srivastava, Vishal Sahni, and Prem Saran Satsangi. “From n-qubit multi-particle quantum teleportation modelling to n-qudit contextuality based quantum teleportation and beyond”. *Int. J. Gen. Syst.* **46**, 414–435 (2017).

[25] Shiroman Prakash and Aashi Gupta. “Contextual bound states for qudit magic state distillation”. *Phys. Rev. A* **101**, 010303 (2020).

[26] Hussain Anwar, Earl T Campbell, and Dan E Browne. “Qutrit magic state distillation”. *New J. Phys.* **14**, 063006 (2012).

[27] Earl T. Campbell, Hussain Anwar, and Dan E. Browne. “Magic-state distillation in all prime dimensions using quantum Reed-Muller codes”. *Phys. Rev. X* **2**, 041021 (2012).

[28] Earl T. Campbell. “Enhanced fault-tolerant quantum computing in d -level systems”. *Phys. Rev. Lett.* **113**, 230501 (2014).

[29] Anirudh Krishna and Jean-Pierre Tillich. “Towards low overhead magic state distillation”. *Phys. Rev. Lett.* **123**, 070507 (2019).

[30] Hillary Dawkins and Mark Howard. “Qutrit magic state distillation tight in some directions”. *Phys. Rev. Lett.* **115**, 030501 (2015).

[31] Mark Howard and Hillary Dawkins. “Small codes for magic state distillation”. *Eur. Phys. J. D* **70**, 55 (2016).

[32] Shiroman Prakash. “Magic state distillation with the ternary Golay code”. *Proceedings of the Royal Society A* **476**, 20200187 (2020).

[33] Mark Howard and Jiri Vala. “Qudit versions of the qubit $\pi/8$ gate”. *Phys. Rev. A* **86**, 022316 (2012).

[34] Wim van Dam and Mark Howard. “Noise thresholds for higher-dimensional systems using the discrete Wigner function”. *Phys. Rev. A* **83**, 032310 (2011).

[35] Shiroman Prakash, Akalank Jain, Bhakti Kapur, and Shubangi Seth. “Normal form for single-qutrit Clifford+ T operators and synthesis of single-qutrit gates”. *Phys. Rev. A* **98**, 032304 (2018).

[36] Andrew N. Glaudell, Neil J. Ross, and Jacob M. Taylor. “Canonical forms for single-qutrit Clifford+ T operators”. *Annals of Physics* **406**, 54–70 (2019).

[37] Shawn X. Cui, Daniel Gottesman, and Anirudh Krishna. “Diagonal gates in the Clifford hierarchy”. *Phys. Rev. A* **95**, 012329 (2017).

[38] Amolak Ratan Kalra, Dinesh Valluri, and Michele Mosca. “Synthesis and arithmetic of single qutrit circuits” (2024). [arXiv:2311.08696](https://arxiv.org/abs/2311.08696).

[39] Akalank Jain and Shiroman Prakash. “Qutrit and ququint magic states”. *Phys. Rev. A* **102**, 042409 (2020).

[40] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. “Negative quasi-probability as a resource for quantum computation”. *New J. Phys.* **14**, 113011 (2012).

[41] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. “The resource theory of stabilizer quantum computation”. *New J. Phys.* **16**, 013009 (2014).

[42] Hussain Anwar, Benjamin J. Brown, Earl T. Campbell, and Dan E. Browne. “Fast decoders for qudit topological codes”. *New J. Phys.* **16**, 063038 (2014).

[43] Fern H. E. Watson, Hussain Anwar, and Dan E. Browne. “Fast fault-tolerant decoder for qubit and qudit surface codes”. *Phys. Rev. A* **92**, 032309 (2015).

[44] Fern H. E. Watson, Earl T. Campbell, Hussain Anwar, and Dan E. Browne. “Qudit color codes and gauge color codes in all spatial dimensions”. *Phys. Rev. A* **92**, 022312 (2015).

[45] Adam M. Meier, Bryan Eastin, and Emanuel Knill. “Magic-state distillation with the four-qubit code” (2012). [arXiv:1204.4221](https://arxiv.org/abs/1204.4221).

[46] A. Mari and J. Eisert. “Positive Wigner functions render classical simulation of quantum computation efficient”. *Phys. Rev. Lett.* **109**, 230503 (2012).

[47] Mark Howard and Earl Campbell. “Application of a resource theory for magic states to fault-tolerant quantum computing”. *Phys. Rev. Lett.* **118**, 090501 (2017).

[48] Xin Wang, Mark M. Wilde, and Yuan Su. “Efficiently computable bounds for magic state distillation”. *Phys. Rev. Lett.* **124**, 090505 (2020).

[49] Nikolaos Koukoulekidis and David Jennings. “Constraints on magic state protocols from the statistical mechanics of Wigner negativity”. *npj Quantum Information* **8**, 42 (2022).

[50] Austin G. Fowler, Simon J. Devitt, and Cody Jones. “Surface code implementation of block code state distillation”. *Scientific Reports* **3**, 1939 (2013).

[51] B. P. Lanyon, T. J. Weinhold, N. K. Langford, J. L. O’Brien, K. J. Resch, A. Gilchrist, and A. G. White. “Manipulating biphotonic qutrits”. *Phys. Rev. Lett.* **100**, 060504 (2008).

[52] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, and A. Wallraff. “Control and tomography of a three level superconducting artificial atom”. *Phys. Rev. Lett.* **105**, 223601 (2010).

[53] M.A. Yurtalan, J. Shi, G.J.K. Flatt, and A. Lupascu. “Characterization of multilevel dynamics and decoherence in a high-anharmonicity capacitively shunted flux circuit”. *Phys. Rev. Appl.* **16**, 054051 (2021).

[54] M. Kononenko, M. A. Yurtalan, S. Ren, J. Shi, S. Ashhab, and A. Lupascu. “Characterization of control in a superconducting qutrit using randomized benchmarking”. *Phys. Rev. Res.* **3**, L042007 (2021).

[55] Martin Ringbauer, Michael Meth, Lukas Postler, Roman Stricker, Rainer Blatt, Philipp Schindler, and Thomas Monz. “A universal qudit quantum processor with trapped ions”. *Nature Phys.* **18**, 1053–1057 (2022). [arXiv:2109.06903](https://arxiv.org/abs/2109.06903).

[56] Noah Goss et al. “High-fidelity qutrit entangling gates for superconducting circuits”. *Nature Commun.* **13**, 7481 (2022). [arXiv:2206.07216](https://arxiv.org/abs/2206.07216).

[57] Jan Schütz, Alexander Martin, Sanah Laschinger, and Gerhard Birkl. “Coherent dynamics in a five-level atomic system”. *J. Phys. B* **55**, 234004 (2022). [arXiv:2210.11893](https://arxiv.org/abs/2210.11893).

[58] Joseph Lindon, Arina Tashchilina, Logan W. Cooke, and Lindsay J. LeBlanc. “Complete unitary qutrit control in ultracold atoms”. *Phys. Rev. Appl.* **19**, 034089 (2023).

- [59] Noah Goss, Samuele Ferracin, Akel Hashim, Arnaud Carignan-Dugas, John Mark Kreikebaum, Ravi K. Naik, David I. Santiago, and Irfan Siddiqi. “Extending the computational reach of a superconducting qutrit processor” (2023). [arXiv:2305.16507](https://arxiv.org/abs/2305.16507).
- [60] Mahadevan Subramanian and Adrian Lupascu. “Efficient two-qutrit gates in superconducting circuits using parametric coupling” (2023). [arXiv:2309.05766](https://arxiv.org/abs/2309.05766).
- [61] Wieb Bosma, John Cannon, and Catherine Playoust. “The Magma algebra system I: The user language”. *J. Symb. Comput.* **24**, 235–265 (1997).
- [62] F.J. MacWilliams and N.J.A. Sloane. “Chapter 5: Dual codes and their weight distribution”. In *The Theory of Error-Correcting Codes*. Volume 16 of North-Holland Mathematical Library, pages 125–154. Elsevier (1977).