



This is a repository copy of *Unlocking early fault-tolerant quantum computing with mitigated magic dilution*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/233252/>

Version: Published Version

Article:

Luthra, S. orcid.org/0009-0001-8029-5835, Moylett, A.E. orcid.org/0000-0003-0163-5262, Browne, D.E. orcid.org/0000-0003-3001-158X et al. (1 more author) (2025) Unlocking early fault-tolerant quantum computing with mitigated magic dilution. *Quantum Science and Technology*, 10 (4). 045066. ISSN: 2058-9565

<https://doi.org/10.1088/2058-9565/ae0aef>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

PAPER • OPEN ACCESS

Unlocking early fault-tolerant quantum computing with mitigated magic dilution

To cite this article: Surabhi Luthra *et al* 2025 *Quantum Sci. Technol.* **10** 045066

View the [article online](#) for updates and enhancements.

You may also like

- [Advancing optical spiking neural networks with exciton-polaritons](#)
Krzysztof Tyszką, Andrzej Opala and Barbara Pietka
- [A bending microactuator with large angular range for origami-inspired bidirectional folding](#)
Lena Seigner, Vincent Gottwald, Lars Bumke et al.
- [Properties of a definition of the SI second based on several optical transitions](#)
Jérôme Lodewyck and Tetsuya Ido



 **Delft Circuits**
Pioneering i/o for advanced technologies

Connect to your ambitions

Flexible, high-density, cryogenic i/o

Quantum Science and Technology



PAPER

OPEN ACCESS

RECEIVED

30 May 2025

REVISED

12 August 2025

ACCEPTED FOR PUBLICATION

24 September 2025

PUBLISHED

17 October 2025

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.



Unlocking early fault-tolerant quantum computing with mitigated magic dilution

Surabhi Luthra^{1,*} , Alexandra E Moylett^{2,4} , Dan E Browne¹ and Earl T Campbell^{2,3}

¹ Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

² Riverlane, St Andrews House, 59 St Andrews Street, Cambridge CB2 3BZ, United Kingdom

³ School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, S3 7RH, United Kingdom

⁴ Current address: Nu Quantum, 21 JJ Thompson Avenue, Cambridge, CB3 0FA, United Kingdom

* Author to whom any correspondence should be addressed.

E-mail: zcaplut@ucl.ac.uk

Keywords: quantum error mitigation, quantum error correction, quantum simulation, mitigated magic dilution

Abstract

As quantum computing progresses towards the early fault-tolerant regime, quantum error correction will play a crucial role in protecting qubits and enabling logical Clifford operations. However, the number of logical qubits will initially remain limited, posing challenges for resource-intensive tasks like magic state distillation. It is therefore essential to develop efficient methods for implementing non-Clifford operations, such as small-angle rotations, to maximise the computational capabilities of devices within these constraints. In this work, we introduce mitigated magic dilution (MMD) as an approach to synthesise small-angle rotations by employing quantum error mitigation techniques to sample logical Clifford circuits given noisy encoded magic states. We explore the utility of our approach for the simulation of the 2D Fermi–Hubbard model. We identify evolution time regimes where MMD outperforms state-of-the-art synthesis techniques in the number of noisy encoded magic states required for square lattices up to size 8×8 . Moreover, we demonstrate that our method can provide a practical advantage that is quantified by a better-than-quadratic improvement in the resource requirements for small-angle rotations over classical simulators. This work paves the way for early fault-tolerant demonstrations on devices supporting millions of quantum operations, the so-called MegaQuOp regime.

1. Introduction

Recent progress in experimental demonstrations of quantum error correction and logical computation [1–5] has encouraged research towards practical applications of early fault-tolerant quantum computers [6–9]. The Gottesman–Knill theorem [10] shows that universal quantum computation necessitates non-Clifford gates. Implementing these gates typically requires the preparation of magic states or non-stabiliser states where resource-intensive techniques such as magic state distillation [11] are used to improve their fidelity. Despite ongoing research in reducing the associated overhead of these factories [12–14], resource analysis of fault-tolerant implementations of large algorithms using magic state distillation require millions of physical qubits [15–18]. In the anticipated early fault-tolerant era, where useful magic state distillation factories cannot be accommodated due to physical qubit limitations, we investigate quantum error mitigation (QEM) as a promising alternative.

QEM is generally considered in the context of estimating the expectation value of an observable using a quantum circuit, for example in variational quantum eigensolvers [19] and statistical versions of phase estimation [20–24]. QEM methods improve the accuracy of the estimated expectation value by reducing the noise-induced bias in the circuit [25, 26]. This is done through post-processing of the measurement outcomes from an ensemble of circuit implementations and so is distinct from quantum error correction which reduces the logical error rate for each individual circuit run [26]. Consequently, most error mitigation

techniques require additional sampling that increases with the noise in the circuit, resulting in asymptotically unfavourable scaling [26, 27].

Despite this, QEM techniques could thrive in the early fault-tolerant regime when applied at the logical level by focusing on their pre-asymptotic behaviour prior to the exponential scaling becoming impractical [6, 28]. As per the standard magic state model [11], we assume that logical Clifford operations are ideal; these may be implemented transversely depending on the error-correcting code and underlying hardware. Meanwhile, non-Clifford operations are subject to noise due to imperfect encoded magic states. Given this model, we leverage methods from QEM to simulate single-qubit Z -rotation gates in this work, which are an integral part of many quantum algorithms.

QEM ideas have previously been proposed to simulate circuits containing T gates such as in [27]. A single-qubit Z -rotation gate could then be realised by synthesising into error-corrected Clifford gates and encoded noisy T gates, for example, using the Ross–Selinger method [29]. This leads to a T -count (total number of T gates) that scales as $\mathcal{O}(\log_2(1/\epsilon_{\text{synth}}))$ with accuracy ϵ_{synth} of the resultant rotation. The sampling overhead also increases dramatically with the inverse accuracy. This is counter-intuitive as magic resource theory [30–34] shows that a small-angle Z -rotation gate is a less powerful resource than a T gate, and yet they are more costly to implement. This perspective hints towards a much more efficient approach to QEM where small-angle Z -rotation gates have a low gate and sampling overhead assuming error-corrected Clifford gates with sufficiently low noise levels. This direction has been explored in the noisy-intermediate scale quantum (NISQ) setting [35] but not in the context of early fault-tolerant quantum computing.

In this paper, we introduce a framework that applies the quasiprobability method [31, 36] to explore the advantages of decomposing single-qubit Z -rotation gates into gates from the Clifford hierarchy [37]. A conceptual starting point is to consider the following two-step process: first, use magic state dilution [38] to convert a high magic resource into many low magic resources; and second, use QEM to reduce any inherent noise or noise introduced in the dilution process. While each of these two steps could be cast as separate convex optimisation problems, it is more elegant and optimal to compress them. Therefore, rather than implementing these processes independently, our approach unifies them into a single optimisation problem, which we introduce as mitigated magic dilution (MMD). Specifically, we use convex optimisation to find the optimal sample complexity of performing small-angle single-qubit rotations from noisy encoded magic states.

There has been significant progress in the preparation of encoded magic states. For example, in [39], it was shown that encoded magic states can be prepared with a logical error rate of $\sim 0.4 \times 10^{-3}$ under the assumption of ideal single-qubit operations and depolarised two-qubit gates with 0.1% error rate using post-selection. Moreover, there are several improvements to the state preparation of non-stabiliser states for arbitrary small-angle rotations that could offer further advantages to the results presented in this paper, which we examine in section 5.

We compare our framework against a baseline classical approach where small-angle single-qubit rotations are decomposed into Clifford operations, similar to [30–34]. Notably, we demonstrate a polynomial advantage, the magnitude of which depends on the quality of initial magic state preparation. We quantify this advantage in terms of the polynomial degree of magic resource saving of our method. For encoded magic states prepared with 1% dephasing noise, this saving is better than cubic, while for 0.1% dephasing noise we find a saving of degree approximately 11.43. Moreover, we show that our method can improve upon a state-of-the-art classical simulator, the sum-over-Cliffords stabiliser extent method [34], offering a 2.37 degree of magic resource saving.

To evaluate the practical benefits of this approach, we study the resource requirements to simulate the time evolution of the 2D Fermi–Hubbard model. The 2D Fermi–Hubbard model [40, 41] is of notable interest in the early fault-tolerant regime [6, 42] due to its importance in condensed matter physics (e.g. to understand high-temperature superconductivity [43] and the Mott metal–insulator transition [44]), as well as its simplicity arising from its highly regular lattice structure. Our analysis in section 4 demonstrates that MMD is a more resource-efficient method than direct gate synthesis, requiring fewer magic states in total. Moreover, the expected number of magic states per sample is significantly smaller, and therefore MMD is particularly amenable to early fault-tolerant devices.

Even when quantum computers have an asymptotic speedup over classical computing they can fail to have an in-practice speedup for example problems of a relevant size [45], and so it is crucial to make such comparison. Here, we consider the example problem of a Fermi–Hubbard model with a 6×6 square lattice and evolution time $t = 0.25$ as a strong candidate for quantum advantage. We find that our MMD method requires a circuit with only 1037 non-Clifford gate teleportations (on average) and 5.34×10^6 samples. Even with error correction overheads, a sample per second is a conservative runtime estimate, taking a single quantum computer 62 days for all samples. For the same calculation, the sum-over-Cliffords stabiliser extent method would require a 5.18×10^{17} seconds runtime, which would take 1.64×10^4 years to complete when

assuming a million classical processors in parallel. Thereby, our MMD protocol facilitates a significant reduction in runtime over the the sum-over-Cliffords classical simulator for conservative time evolution and lattice sizes of interest. Combined with a Pauli-based model of computation [46], such an application would require logical error rates of approximately 1 part in a million, the so-called MegaQuOp regime [9].

This paper is structured as follows. In section 2 we briefly summarise the quasiprobability method. In section 3, we introduce our framework in two steps, working in the channel representation of all unitary gates U throughout, which are denoted as $\mathcal{U}(\cdot) = U(\cdot)U^\dagger$. We first present the application of the quasiprobability method to target $\mathcal{Z}$ rotation channels over diagonal Clifford + $\mathcal{T}$ channels in section 3.1 using a linear combination of channels decomposition. We then generalise this to decompositions over diagonal Clifford + $\mathcal{T}^{\frac{1}{n}}$ channels in section 3.2. Finally, in section 4, we demonstrate the practical applicability of our framework for a second-order Trotter simulation of the Fermi–Hubbard model with comparison to gate synthesis.

2. Quasiprobability method

The motivation behind the quasiprobability method is that an ideal quantum operation can be decomposed into a basis set of noisy operations. Let $\mathcal{U}^t(\rho) = U^t \rho U^{t\dagger}$ be the ideal target unitary channel of a quantum operation U^t and let $\{\mathcal{U}_i^n\}$ be a set of channels corresponding to the noisy operations that can be performed on a given quantum hardware. The noise on these operations would typically be characterised using a tomography procedure [26, 47]. We can then write the following decomposition, hereafter referred to as the linear combinations of channels (LCC) decomposition

$$\mathcal{U}^t = \sum_i x_i \mathcal{U}_i^n \quad (1)$$

where x_i are real coefficients that can be positive or negative such that equation (1) is a quasiprobability representation.

It follows that the expectation value of an observable O for the target operation can be written in terms of the expectation value of the noisy operations as

$$\begin{aligned} \text{Tr} [O \mathcal{U}^t(\rho)] &= \text{Tr} \left[O \sum_i x_i \mathcal{U}_i^n(\rho) \right] \\ &= \sum_i x_i \text{Tr} [O \mathcal{U}_i^n(\rho)]. \end{aligned} \quad (2)$$

The expectation values $\text{Tr} [O \mathcal{U}_i^n(\rho)]$ associated with noisy channels $\mathcal{U}_i^n$ can be estimated using Monte Carlo sampling. Each noisy operation $\mathcal{U}_i^n$, specified by index i , is applied with probability $|x_i|/\sum_i |x_i|$, and the resulting expectation value is multiplied by $\text{sign}(x_i)$ [26, 31]. From this, an estimate of the ideal expectation value can be calculated up to an accuracy ϵ and probability greater than $1 - \delta$, where

$$\delta = 2 \exp \left(\frac{-N\epsilon^2}{2(\sum_i |x_i|)^2} \right) \quad (3)$$

and the number of samples N is determined by Hoeffding's inequality to be

$$N = \frac{2}{\epsilon^2} \left(\sum_i |x_i| \right)^2 \ln \left(\frac{2}{\delta} \right). \quad (4)$$

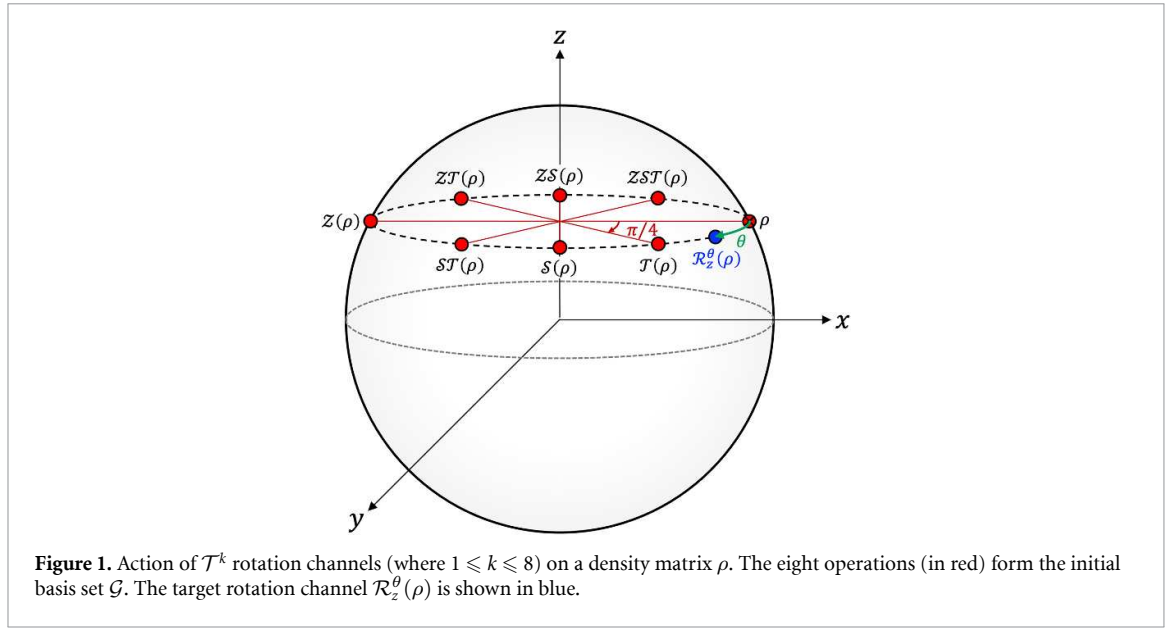
Thus, the number of samples scale as $\mathcal{O}(\lambda^2/\epsilon^2)$ up to logarithmic factors, with $\lambda = \sum_i |x_i|$. The quantity λ^2 is referred to as the *sampling overhead* of using this method over finding the expectation value of the ideal target channel directly [26, 27].

3. Methodology

3.1. LCC decomposition

We first demonstrate how the quasiprobability method can be used to decompose target $\mathcal{Z}$ -rotation channels. In the simplest case, we consider a diagonal Cliffords + $\mathcal{T}$ channel decomposition, forming an overcomplete basis set to ensure that we can optimise this decomposition with respect to the sampling overhead.

Consider the action of a small-angle $\mathcal{Z}$ -rotation channel $\mathcal{R}_z^\theta(\rho) = R_z(\theta)\rho(R_z(\theta))^\dagger$ acting on a state represented by density matrix ρ in figure 1, where $R_z(\theta) = \exp(-i(\theta/2)Z)$ is a gate representing a rotation



about the Z -axis with angle θ . In this geometric picture, we label eight channels corresponding to the action of a basis set of gates

$$G = \{T^k : 1 \leq k \leq 8\} = \{I, T, S, ST, Z, ZT, ZS, ZST\}, \quad (5)$$

which forms a cyclic group $(G, \times)$ generated by the T gate under multiplication. We represent these gates in their channel representation to form a set $\mathcal{G}$ such that

$$\mathcal{G} = \{\mathcal{T}^k(\rho) = (T^k)\rho(T^k)^\dagger : 1 \leq k \leq 8\}. \quad (6)$$

The LCC decomposition is given by

$$\mathcal{R}_z^\theta(\cdot) = \sum_{U \in G} x_U U(\cdot) U^\dagger = \sum_{\mathcal{U} \in \mathcal{G}} x_{\mathcal{U}} \mathcal{U}(\cdot), \quad (7)$$

where $\mathcal{R}_z^\theta(\cdot)$ represents the single-qubit Z -rotation channel for the $R_z(\theta)$ gate. This is a quasiprobability decomposition such that the coefficients x_U are real and satisfy $\sum_{U \in G} x_U = 1$. The quantity λ is defined as the l_1 -norm of the coefficients in this decomposition, that is

$$\lambda = \|x\|_1 = \sum_{U \in G} |x_U|, \quad (8)$$

such that λ^2 is the sampling overhead of each decomposition from Hoeffding's inequality [25, 27].

It can be noted that the decomposition in equation (7) is not unique over a chosen basis set and so it is important to minimise the sampling overhead over different solutions. We define the optimal decomposition for a given Z -rotation channel as the decomposition into elements of $\mathcal{G}$ with the minimum λ , which we denote as $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta)$. This can be written as

$$\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot)) = \min \left\{ \lambda = \|x_{\mathcal{U}}\|_1 \left| \mathcal{R}_z^\theta(\cdot) = \sum_{\mathcal{U} \in \mathcal{G}} x_{\mathcal{U}} \mathcal{U}(\cdot) \right. \right\}. \quad (9)$$

By definition, $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta) \geq 1$ where $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta) = 1$ if $\mathcal{R}_z^\theta \in \mathcal{G}$ for some group $\mathcal{G}$ representing the basis set of channels. The solution to equation (9) is equivalent to minimising the sampling overhead λ^2 over the basis set of channels, thus $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot))$ finds the optimal decomposition of $\mathcal{R}_z^\theta(\cdot)$ for the chosen group $\mathcal{G}$. This can be further expressed as a convex optimisation problem to solve a linear system given by

$$\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot)) = \min \|x\|_1 \text{ subject to } Ax = b. \quad (10)$$

Here, vector b represents m elements of the target channel $\mathcal{R}_z^\theta$ expressed in a vectorised form, and A is an $m \times n$ matrix, with n columns for the basis set of channels in the decomposition and m rows of elements that specify each channel (see equations (A1) and (A2) as an example). The j th column of matrix A can be generated by determining the vectorised form of the j th channel $\mathcal{U}_j(\cdot) = U_j(\cdot)U_j^\dagger$ for $\mathcal{U}_j \in \mathcal{G}$ and $j \in \{1, 2, \dots, n\}$. Where possible, we also make a concerted effort to minimise the number of non-zero elements in x whilst keeping the optimal sampling overhead constant.

For the ideal basis set of channels $\mathcal{G}$ defined in equation (6), we use CVXPY [48, 49] to solve equation (10) and find that decompositions of the form

$$\mathcal{R}_z^\theta(\cdot) = x_I \mathcal{I}(\cdot) + x_T \mathcal{T}(\cdot) + x_Z \mathcal{Z}(\cdot) \quad (11)$$

are optimal, with

$$\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot)) = (\sqrt{2} - 1) \sin(\theta) + \cos(\theta) \quad (12)$$

for $0 \leq \theta \leq \frac{\pi}{4}$. Additionally, considering subsets of $\mathcal{G}$ with only 2 elements, we find that a solution to the optimisation problem cannot be found, and therefore 3 contributions are required (see appendix A).

As a classical baseline, we consider the subset of $\mathcal{G}$ that are Clifford operations, denoted as $\mathcal{C}$. In that case, the optimal decomposition consists of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$ channels and the corresponding l_1 -norm $\Lambda_{\mathcal{C}}$ is

$$\Lambda_{\mathcal{C}}(\mathcal{R}_z^\theta(\cdot)) = \sin(\theta) + \cos(\theta). \quad (13)$$

We compare this baseline to our quantum protocol by considering $(\Lambda_{\mathcal{G}}^\gamma)^\gamma = \Lambda_{\mathcal{C}}^\gamma$, which implicitly defines γ where

$$\gamma(\theta) = \frac{\ln(\Lambda_{\mathcal{C}}(\mathcal{R}_z^\theta))}{\ln(\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta))}. \quad (14)$$

Consequently, we interpret γ as the polynomial degree of magic resource saving of the optimal decomposition over $\mathcal{G}$ compared to the optimal decomposition over $\mathcal{C}$, for a specified target rotation channel. As $\theta \rightarrow 0$, γ approaches $\sqrt{2} + 1 \approx 2.41 > 2$, indicating a slightly better-than-quadratic advantage of decomposing a (very) small-angle rotation into a quasiprobability decomposition of $\{\mathcal{I}, \mathcal{T}, \mathcal{Z}\}$ channels instead of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$ channels. From this, we can now proceed to generalise these findings to achieve a better advantage for small-angle rotations by our choice of $\mathcal{G}$ from the Clifford hierarchy.

3.2. Climbing the Clifford hierarchy

From equations (12) and (13), it can be seen that the optimal channel decompositions over our choice of $\mathcal{G}$ and $\mathcal{C}$ respectively are of the form $\{\mathcal{I}, \mathcal{R}_z^\phi, \mathcal{Z}\}$. Specifically, $\phi = \pi/2$ for $\mathcal{R}_z^\phi = \mathcal{S}$ and $\phi = \pi/4$ for $\mathcal{R}_z^\phi = \mathcal{T}$ with optimality valid within the range $0 \leq \theta \leq \phi$. Therefore, for smaller target rotations, we can minimise $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot))$ further by including $\mathcal{R}_z^\phi$ channels with smaller ϕ in our choice of $\mathcal{G}$.

First, we let $i \in \mathbb{Z}_{\geq 0}$ and define $n = 2^{i-1}$. We then choose $\mathcal{G}$ to be

$$\mathcal{G} = \left\{ \mathcal{T}_n^k(\rho) = \left(T_n^k\right) \rho \left(T_n^k\right)^\dagger : 1 \leq k \leq 8n \right\}, \quad (15)$$

where $T_n^{\frac{1}{n}}$ is an n th-root T gate from the $(i+2)$ -th level of the Clifford hierarchy, and $\phi = \pi/4n$. Therefore, $i=0$ represents the S gate, $i=1$ represents the T gate, and so on. These $T_n^{\frac{1}{n}}$ gates can be implemented by a generalised gate teleportation circuit as shown in figure 2.

We find that the LCC decomposition for equation (10) with $\mathcal{G}$ as equation (15) that minimises the sampling overhead is given by

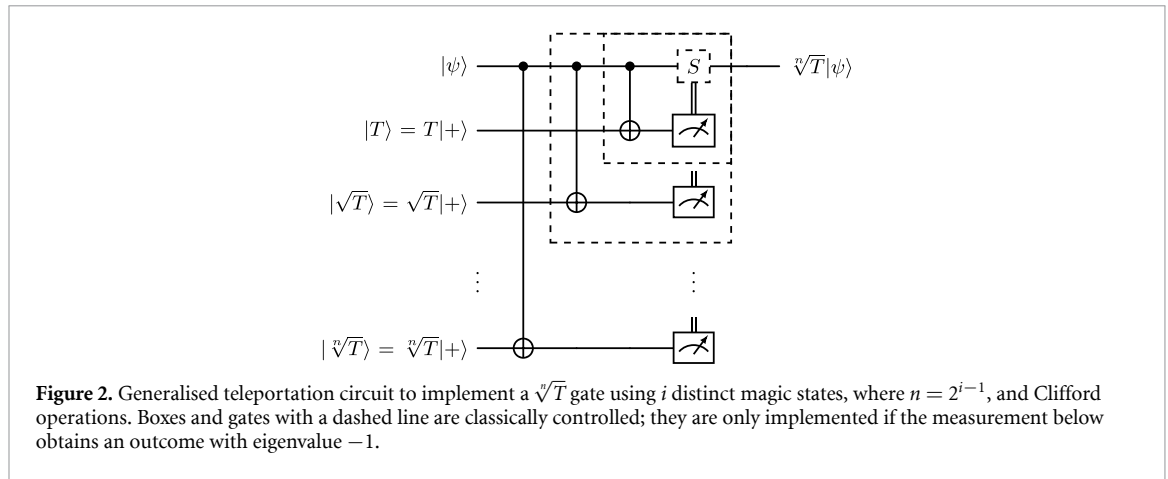
$$\mathcal{R}_z^\theta(\cdot) = x_I \mathcal{I}(\cdot) + x_n \mathcal{T}_n^{\frac{1}{n}}(\cdot) + x_Z \mathcal{Z}(\cdot), \quad (16)$$

with corresponding $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot))$ as

$$\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot)) = \cos(\theta) + \frac{\sin(\theta)}{\sin(\phi)} (1 - \cos(\phi)). \quad (17)$$

For small-angle rotations, $\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot))$ can be approximated to be

$$\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta(\cdot)) \approx 1 + \theta \left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right), \quad (18)$$



with a degree of resource saving compared to a Clifford decomposition of

$$\gamma = \frac{\sin(\phi)}{1 - \cos(\phi)} = \cot\left(\frac{\phi}{2}\right) \quad (19)$$

in the limit of small θ (see appendix B.1). Therefore in the $\{\mathcal{I}, \mathcal{T}_n^{\frac{1}{n}}, \mathcal{Z}\}$ decomposition, climbing the Clifford hierarchy in terms of $\mathcal{T}_n^{\frac{1}{n}}$ channels results in a larger degree of saving compared to the $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$ Clifford decomposition.

We now consider the case where the non-Clifford channels in our basis set are subject to noise, as originally motivated in the quasiprobability method. Non-Clifford operations (in our case $\mathcal{T}_n^{\frac{1}{n}}$) can be implemented via encoded gate teleportation circuits using noisy encoded magic states and error-corrected Clifford gates as shown in figure 2. The most relevant noise channel for these diagonal rotation gates in our basis set is the dephasing noise channel (see appendix C), which is defined as

$$\varepsilon(\cdot) = (1 - p)(\cdot) + pZ(\cdot)Z, \quad (20)$$

where p quantifies the amount of dephasing noise. The action of the dephasing noise channel on a unitary channel in our basis set of unitary operations is given as

$$\mathcal{U}^{\text{deph}}(\cdot) = \varepsilon(\mathcal{U}^{\text{ideal}}(\cdot)) = (1 - p)\mathcal{U}^{\text{ideal}}(\cdot) + p\mathcal{Z} \circ \mathcal{U}^{\text{ideal}}(\cdot). \quad (21)$$

As the number of iterations of the teleportation circuit increases for the implementation of gates from higher levels of the Clifford hierarchy (figure 2), we find that the dephasing noise changes correspondingly. We assume that each non-stabiliser state $|T_n^{\frac{1}{n}}\rangle$ is prepared with uniform (independent of θ) fidelity. However, in section 5, we discuss how θ -dependent fidelities could lead to further performance improvements.

If the dephasing noise for a T gate implementation (requiring a $|T\rangle$ magic state) is p , the effective dephasing noise for a $T_n^{\frac{1}{n}}$ gate can be found by considering the errors on the non-stabiliser states that lead to an error on the $T_n^{\frac{1}{n}}$ gate. For example, the probability of a Z error acting on a $T^{\frac{1}{2}}$ gate is given by the probability of an error occurring on either the $|T\rangle$ or $|T^{\frac{1}{2}}\rangle$ state [50], that is

$$p_{\text{eff}} = p\left(1 - \frac{p}{2}\right) + \frac{p}{2}(1 - p) = \frac{3}{2}p - p^2 < \frac{3}{2}p, \quad (22)$$

where we take into account that there is a 50% probability that a T gate correction needs to be applied (recall figure 2). Therefore, the effective dephasing noise on a $T_n^{\frac{1}{n}}$ gate is bounded by $p_{\text{eff}} = (2 - n^{-1})p$.

As a result of this, the l_1 -norm for the $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ decomposition transforms from the ideal case in equation (17) to the following:

$$\begin{aligned} \Lambda_{\mathcal{G}}(\mathcal{R}_z^{\theta}(\cdot)) &= \left| \cos^2\left(\frac{\theta}{2}\right) - \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi) \right] \frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} \right| \\ &\quad + \left| \frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} \right| \\ &\quad + \left| \sin^2\left(\frac{\theta}{2}\right) - \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi) \right] \frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} \right|. \end{aligned} \quad (23)$$

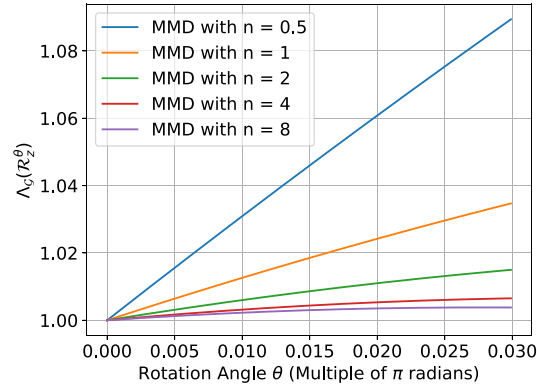


Figure 3. $\Lambda_G(\mathcal{R}_z^\theta)$ as a function of target rotation angle θ for different optimal decompositions of the form $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ including the optimal Clifford decomposition of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$. A dephasing error of 0.1% is assumed for non-Clifford state preparation.

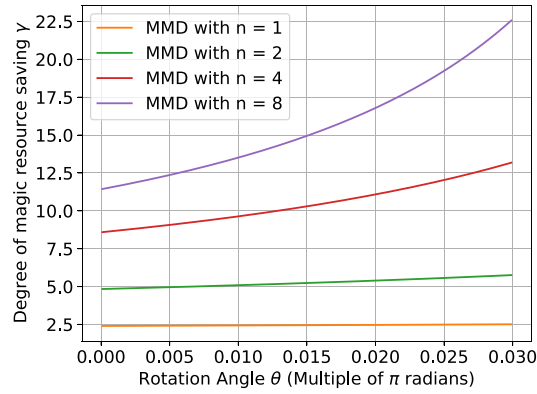


Figure 4. Degree of magic resource saving γ as a function of target rotation angle θ for different optimal decompositions of the form $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ relative to the optimal Clifford decomposition of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$. A dephasing error of 0.1% is assumed for non-Clifford state preparation.

In figure 3 we present $\Lambda_G(\mathcal{R}_z^\theta)$ as function of θ for $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ decompositions with $n \in \{0.5, 1, 2, 4, 8\}$ for $p = 0.1\%$ dephasing noise, where $n = 0.5$ corresponds to the $\mathcal{S}$ channel.

In particular, we are interested in the degree of magic resource saving γ to benchmark against the classical Clifford decomposition. As $\theta \rightarrow 0$, γ with respect to the $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$ Clifford decomposition is approximated by (see appendix B.2)

$$\frac{1}{\gamma} \approx \frac{\csc(\phi)}{(1 - 2p_{\text{eff}})} - \cot(\phi). \quad (24)$$

From figure 4, we find that for $p = 0.1\%$ and in the limit of small θ , we can achieve at least a ~ 2.4 degree of saving in the simplest case of a $\{\mathcal{I}, \varepsilon(\mathcal{T}), \mathcal{Z}\}$ decomposition, and greater than ~ 11.4 degree of saving relative to the Clifford decomposition by replacing $\mathcal{T}$ with a $\mathcal{T}^{\frac{1}{8}}$ channel.

We further note the dependence of $\Lambda_G(\mathcal{R}_z^\theta)$ on the dephasing noise p . For $p = 0.1\%$, increasing n in our protocol results in a greater advantage up to $n = 8$, as shown in figure 4. However, for higher p , this is not always the case. In fact, for 1% dephasing noise, going beyond $n = 2$ presents no further improvement in resource saving as shown in table 1 (see appendix D for corresponding values of $\ln(\Lambda_G(\mathcal{R}_z^\theta))$). In other words, the $\{\mathcal{I}, \varepsilon(\mathcal{T}^{\frac{1}{2}}), \mathcal{Z}\}$ decomposition is optimal and provides a lower bound of ~ 3.6 degree of saving for small θ . In general, the smallest value of n that provides the optimal saving can be found by maximising γ with respect to n (indicated in bold in table 1).

A further useful metric to consider is the expected number of magic states required per sample for an optimal decomposition of the form equation (16) which we derive from the proportion of $|x_n|$ relative to $\Lambda_G(\mathcal{R}_z^\theta)$, and the total number of magic states (per sample) required to implement a $T^{\frac{1}{n}}$ gate using figure 2:

$$E = \left(2 - \frac{1}{n}\right) \frac{|x_n|}{\Lambda_G(\mathcal{R}_z^\theta)}. \quad (25)$$

Table 1. Degree of magic resource saving γ for a basis set of channels $\mathcal{G}$ with optimal channel decomposition $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ relative to the $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$ channel combination for increasing n up to the sixth level of the Clifford hierarchy, with non-Clifford gates subject to dephasing noise with probability p . Values in bold indicate the decomposition that offers optimal γ for a given dephasing noise.

n	ϕ	$\mathcal{R}_z^\phi$	γ (Small θ)			
			$p = 0.01\%$	$p = 0.1\%$	$p = 0.5\%$	$p = 1.0\%$
1	$\frac{\pi}{4}$	$\mathcal{T}$	2.41	2.40	2.33	2.26
2	$\frac{\pi}{8}$	$\sqrt{\mathcal{T}}$	5.01	4.84	4.19	3.58
4	$\frac{\pi}{16}$	$\sqrt[4]{\mathcal{T}}$	9.97	8.58	5.27	3.52
8	$\frac{\pi}{32}$	$\sqrt[8]{\mathcal{T}}$	18.88	11.43	4.10	2.24

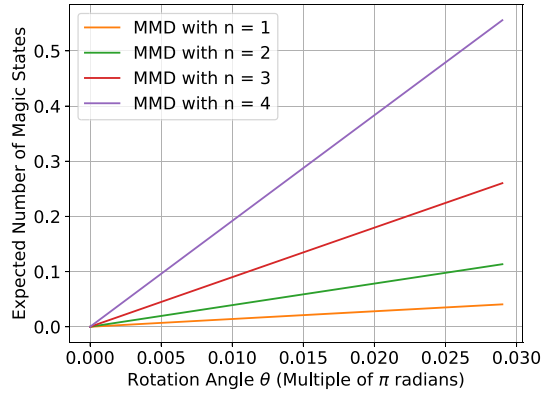


Figure 5. Expected number of magic states E per sample as a function of target rotation angle θ for different optimal decompositions of the form $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ relative to the optimal Clifford decomposition of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$. A dephasing error of 0.1% is assumed for non-Clifford state preparation.

In this expression, x_n is the coefficient of the $\mathcal{T}_n^{\frac{1}{n}}$ term in the LCC decomposition of the form $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$, and $\Lambda_{\mathcal{G}}$ is the l_1 -norm of the coefficients in this decomposition. As the target rotation angle θ decreases, $|x_n|$ also decreases, resulting in a smaller number of expected magic states. Thus, this method ‘dilutes’ magic resource in simulating smaller rotations from $\mathcal{T}_n^{\frac{1}{n}}$ channels of higher magic resource. This dilution phenomenon is evident in figure 5, which shows the expected number of magic states (per sample) decreasing with the rotation angle θ .

We also present a comparison of our quantum protocol to the classical sum-over-Cliffords simulation method [34]. The total runtime of this method scales as $\mathcal{O}(\xi/\epsilon^4)$ where ξ is the stabiliser extent and ϵ is the precision of the simulation. Therefore, the sum-over-Cliffords method has a worse scaling w.r.t. the desired precision than MMD where the scaling is proportional to $1/\epsilon^2$ as in equation (4). In the small θ limit, the stabiliser ext. for a Z -rotation is given by

$$\xi(R(\theta)) = \exp(\tan(\pi/8)\theta). \quad (26)$$

Therefore, we can define a resource saving degree γ_{SE} such that $(\Lambda_{\mathcal{G}}^2)^{\gamma_{\text{SE}}} = \xi$ analogously to equation (24). This results in

$$\gamma_{\text{SE}}(\theta) = \frac{\ln(\xi(R(\theta)))}{2\ln(\Lambda_{\mathcal{G}}(\mathcal{R}_z^\theta))}. \quad (27)$$

We present this with the caveat that MMD has an additional advantage with respect to precision ϵ scaling, which is not captured by γ_{SE} .

As shown in table 2, the MMD method is able to achieve a better-than-quadratic advantage for $p = 0.1\%$ dephasing noise with respect to the stabiliser extent. Meanwhile for higher dephasing noise, our method provides comparable performance with a slight increase in magic resource for $p = 1.0\%$.

4. Fermi–Hubbard model simulation

The Fermi–Hubbard model describes the behaviour of interacting electrons in 2D materials. The Hubbard Hamiltonian is composed of hopping terms H_h which represent the kinetic energy of electrons that can tunnel between neighbouring lattice sites, and interaction terms H_i which represent the potential energy due

Table 2. Degree of magic resource saving γ_{SE} for a basis set of channels $\mathcal{G}$ with optimal channel decomposition $\{\mathcal{I}, \varepsilon(\mathcal{T}_n^{\frac{1}{n}}), \mathcal{Z}\}$ relative to the stabiliser extent method for increasing n up to the sixth level of the Clifford hierarchy, with non-Clifford gates subject to dephasing noise with probability p . Values in bold indicate the decomposition that offers optimal γ_{SE} for a given dephasing noise.

n	ϕ	$\mathcal{R}_z^\phi$	γ_{SE} (Small θ)			
			$p = 0.01\%$	$p = 0.1\%$	$p = 0.5\%$	$p = 1.0\%$
1	$\frac{\pi}{4}$	$\mathcal{T}$	0.50	0.50	0.48	0.47
2	$\frac{\pi}{8}$	$\sqrt{\mathcal{T}}$	1.04	1.00	0.87	0.74
4	$\frac{\pi}{16}$	$\sqrt[4]{\mathcal{T}}$	2.07	1.78	1.09	0.73
8	$\frac{\pi}{32}$	$\sqrt[8]{\mathcal{T}}$	3.91	2.37	0.85	0.46

to the on-site repulsion of electrons. For this analysis, we consider H_i subject to a chemical potential shift as per [6]. The resulting Hamiltonian takes the form

$$H = H_h + H_i = \sum_{\sigma \in \uparrow, \downarrow} \sum_{i \neq j} R_{i,j} a_{i,\sigma}^\dagger a_{j,\sigma} + \frac{u}{4} \sum_i \hat{z}_{i,\uparrow} \hat{z}_{i,\downarrow} \quad (28)$$

where $a, a^\dagger$ are the creation and annihilation operators respectively, u is the repulsive interaction strength between spin-up and spin-down electrons at each site, and $\hat{z}$ is related to the number operator $\hat{n} = a^\dagger a$ by $\hat{z} = (2\hat{n} - \mathbb{1})$. The hopping strength is defined as $R_{i,j} = \tau$ if i, j are nearest-neighbour lattice sites, allowing electrons to tunnel between adjacent sites, and $R_{i,j} = 0$ otherwise. We choose parameters that lie within the regime widely considered to be classically challenging for simulation [6, 42]. As such, we set $u = 8$, $\tau = 1$ and consider 2D square lattices of size $L \times L$ with $L \in \{4, 6, 8\}$. The total number of spin orbitals is given by $N = 2L^2$.

To estimate the resource requirements for simulating the Fermi–Hubbard model using MMD, we employ the Fermionic swap network Trotter step algorithm to implement a second-order Trotterisation of the system, following [51]. We determine the number of rotations and corresponding angle of rotations required for each second-order Trotter step. Specifically, for r Trotter steps, $N_h = 8Nr$ arbitrary rotations of angle $\theta_h = \tau t/4r$ are needed to simulate the hopping terms, while $N_i = Nr/2$ arbitrary rotations of angle $\theta_i = ut/4r$ are required to simulate the interaction terms. This makes use of a slight improvement upon [51] due to the shifted form of the interaction Hamiltonian [6].

From this, we calculate the number of magic states required per sample as

$$N_m = N_h E_h + N_i E_i, \quad (29)$$

where the number of magic states $E_{h(i)}$ (per sample) for each rotation is given by

$$E_{h(i)} = \left(2 - \frac{1}{n}\right) \frac{|x_n|}{\Lambda_{\mathcal{G}}(\mathcal{R}_z^{\theta_{h(i)}})} \quad (30)$$

in analogy to equation (25).

The total number of samples (N_{sample}) is calculated from Hoeffding's inequality to be

$$N_{\text{sample}} = \frac{2 \ln(2/\delta)}{\epsilon_{\text{sample,EM}}^2} \left[(\Lambda_{\mathcal{G}}(\mathcal{R}_z^{\theta_h}))^{2N_h} (\Lambda_{\mathcal{G}}(\mathcal{R}_z^{\theta_i}))^{2N_i} \right]. \quad (31)$$

Here we take the probability of error mitigation failing, δ to be 0.01 and the error bound for failure mitigation $\epsilon_{\text{sample,EM}}$ to be 0.02. The total number of magic states over all samples then becomes

$$N_{\text{total}} = N_m \times N_{\text{sample}}. \quad (32)$$

Considering the exponential scaling of the sampling overhead, we first evaluate the total number of samples for which our framework is practically feasible as a function of the evolution time t . We assume a dephasing noise of $p = 0.1\%$ throughout the analysis in this section. In the limit of large Trotter steps for the MMD method, the number of samples for a lattice size of $L = 6$ scales with t as per figure 6.

We see that a purely classical implementation of our MMD method (for which $n = 0.5$) requires 10^{70} samples to simulate $t = 0.35$, whereas a quantum implementation of our method (so $n \geq 1$) requires from 10^{31} to 10^7 samples as n increases. We also overlay the upper-bounded runtime scaling of the stabiliser extent method with precision $\epsilon = 0.01$ in figure 6.

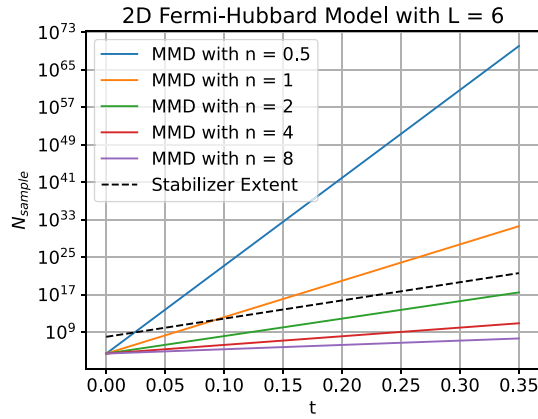


Figure 6. Number of samples (N_{sample}) required to perform MMD for the second-order Trotterised time evolution of the 2D Fermi–Hubbard model in the limit of large (10^6) Trotter steps.

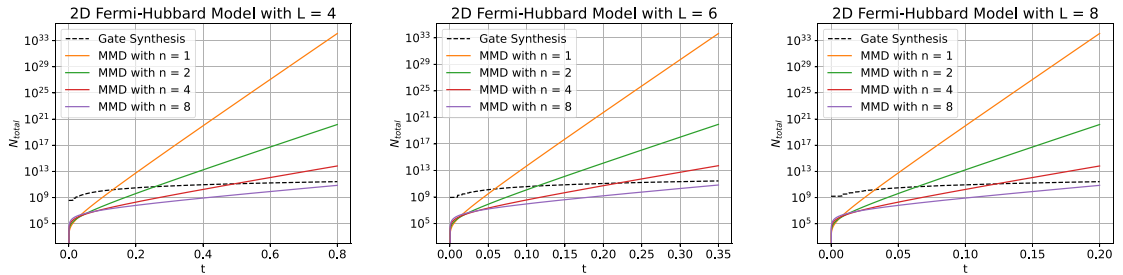


Figure 7. Number of magic states over all samples (N_{total}) to perform MMD for the second-order Trotterised time evolution of the 2D Fermi–Hubbard model compared to gate synthesis. Results are shown for an $L \times L$ lattice with $L \in \{4, 6, 8\}$.

We finally present a representative comparison of this framework with direct gate synthesis methods. Asymptotically optimal unitary synthesis is possible using the Ross–Selinger method [29] with improvements possible by employing random compiling as shown in [52]. Adding an ancillary qubit, mixed fallback techniques [53, 54] obtain the best T -count known to date:

$$T_{\text{synth}} = 0.53 \log_2 \left(\frac{N_R}{\epsilon_{\text{synth}}} \right) + 4.86, \quad (33)$$

where $N_R = N_h + N_i$ is the total number of rotations.

To ensure a fair comparison, we allow an error budget of $\epsilon_{\text{synth}} + \epsilon_{\text{trotter}} = 0.01$, a favourable sampling error $\epsilon_{\text{sample,RS}} = 0.01$, and N_{sample} is a constant. Here, $\epsilon_{\text{trotter}} = W_{\text{FS}} t^3 / s^2$ is the Trotter error for the Fermionic swap network algorithm, taking W_{FS} values from [51]. From this, we optimise the number of required Trotter steps s that minimise $N_R T_{\text{synth}}$ for gate synthesis.

In figure 7, we see that the MMD method provides a significant reduction in the magic state count for very small t , for all cases of n . Smaller lattice sizes see up to a few orders of magnitude saving and could provide an advantage for longer t compared to direct gate synthesis. To achieve the same advantage as the lattice size and evolution time increases, MMD requires gates from higher levels (larger n) of the Clifford hierarchy as shown. In particular, for $L = 4, 6$ and 8 , MMD presents at least an order of magnitude resource saving for time evolution up to $0.8, 0.35$ and 0.2 respectively.

5. Discussion and conclusions

In this paper, we have presented our MMD framework for implementing small-angle single-qubit Z -rotations, making use of error mitigation to provide an advantage in terms of the sampling overhead when benchmarked against a classical approach.

We have further identified a use-case of MMD for simulating time evolution of the Fermi–Hubbard Hamiltonian using Trotterisation. Through this, we have demonstrated a resource benefit of implementing MMD over conventional direct synthesis when taking into account the number of magic states and sampling overhead. Thus, our framework is suitable for early fault-tolerant quantum computers. Broadly speaking,

our approach is more advantageous for algorithms containing a large number of very small angle rotations, such as Trotterised simulations. Our work hints towards a wide variety of directions for future work when larger angle rotations are encountered, including hybrid approaches with MMD used for small angles and synthesis used for larger angles, and hybrid approaches where the gate set G used in MMD includes partially synthesised rotations.

A key assumption we have made in our MMD method is that all magic states are prepared with the same fidelity. Consequently, we observed optimal performance at some finite level of the Clifford hierarchy, fixed by the fidelity. However, recent literature indicates that this can be improved such that magic states of smaller rotations—and therefore gates higher in the Clifford hierarchy—can be prepared with a lower logical error rate [8, 55]. This is particularly relevant for this work as it suggests that the optimal performance will be at even higher Clifford hierarchy levels, leading to higher degrees of magic resource saving and hence speedup. Recent methods including magic state cultivation [56] to construct high fidelity $|T\rangle$ magic states also complement the methods we present in this paper, however further research is needed to determine the feasibility and effectiveness of applying similar techniques to other magic states such as $|T^{\frac{1}{2}}\rangle$ for higher levels of the Clifford hierarchy.

It would be interesting to further analyse the resource requirements when compiling down to lattice surgery operations for particular algorithms. Since our scheme uses repeated teleportation circuits (recall figure 2), these can be realised with logical ZZ measurements without using Hadamard gates or patch rotations. In contrast, when using the gate synthesis approach, we require either Hadamard gates that take a long time to execute [57, 58], or a fast data block structure that comes with an additional qubit overhead compared to compact layouts [46].

The Trotterised time evolution studied in this paper lends itself well into statistical phase estimation for future work. Furthermore, we have studied the effect of single-qubit dephasing noise channels, however for specific quantum hardware, it may be worthwhile to explore other noise models. Finally, further advancements in better logical magic states as discussed earlier will improve the advantages presented in this paper.

We have compared our approach against two families of near-Clifford simulators. Clearly, there are a range of other simulators to consider, including tensor network simulators that are suited to shallow local circuits on 2D arrays of qubits [59]. However, our techniques are directly applicable to simulating systems with long-range interactions or in higher-dimensional geometries, and in such settings tensor-network methods perform poorly. Conversely, we envisage further improvements to MMD by extending to a dyadic decomposition approach similar to that used in some classical simulators [32].

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.5281/zenodo.15427988> [60].

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council [Grant Numbers EP/S021582/1, EP/Y004140/1, EP/Y004310/1]. We thank Nick Blunt and Alex Gramolin for comments on the manuscript.

Appendix A. Minimum requirement of three channels

In the general case, A and b in equation (10) can be written in matrix form as

$$A = \begin{bmatrix} \cos^2\left(\frac{\phi_1}{2}\right) & \cos^2\left(\frac{\phi_2}{2}\right) & \cos^2\left(\frac{\phi_3}{2}\right) & \cdots \\ \cos\left(\frac{\phi_1}{2}\right)\sin\left(\frac{\phi_1}{2}\right) & \cos\left(\frac{\phi_2}{2}\right)\sin\left(\frac{\phi_2}{2}\right) & \cos\left(\frac{\phi_3}{2}\right)\sin\left(\frac{\phi_3}{2}\right) & \cdots \\ \sin^2\left(\frac{\phi_1}{2}\right) & \sin^2\left(\frac{\phi_2}{2}\right) & \sin^2\left(\frac{\phi_3}{2}\right) & \cdots \end{bmatrix}, \quad (\text{A1})$$

and

$$b = \begin{bmatrix} \cos^2\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \sin^2\left(\frac{\theta}{2}\right) \end{bmatrix}. \quad (\text{A2})$$

Specifically for the two channel case, we have

$$\begin{bmatrix} \cos^2\left(\frac{\phi_1}{2}\right) & \cos^2\left(\frac{\phi_2}{2}\right) \\ \cos\left(\frac{\phi_1}{2}\right)\sin\left(\frac{\phi_1}{2}\right) & \cos\left(\frac{\phi_2}{2}\right)\sin\left(\frac{\phi_2}{2}\right) \\ \sin^2\left(\frac{\phi_1}{2}\right) & \sin^2\left(\frac{\phi_2}{2}\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \cos^2\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ \sin^2\left(\frac{\theta}{2}\right) \end{bmatrix} \quad (\text{A3})$$

from which we write three simultaneous equations:

$$x_1 \cos^2\left(\frac{\phi_1}{2}\right) + x_2 \cos^2\left(\frac{\phi_2}{2}\right) = \cos^2\left(\frac{\theta}{2}\right), \quad (\text{A4a})$$

$$x_1 \cos\left(\frac{\phi_1}{2}\right)\sin\left(\frac{\phi_1}{2}\right) + x_2 \cos\left(\frac{\phi_2}{2}\right)\sin\left(\frac{\phi_2}{2}\right) = \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right), \quad (\text{A4b})$$

$$x_1 \sin^2\left(\frac{\phi_1}{2}\right) + x_2 \sin^2\left(\frac{\phi_2}{2}\right) = \sin^2\left(\frac{\theta}{2}\right). \quad (\text{A4c})$$

Upon solving equations (A4a) and (A4c), we obtain

$$x_1 + x_2 = 1, \quad (\text{A5})$$

which arises due to the unitarity of the channels. From this, we can find two solutions for x_1 by substituting this into equations (A4a) and (A4b) respectively, resulting in

$$\begin{aligned} x_1 &= \frac{\cos((\theta + \phi_1)/2)\sin((\theta + \phi_2)/2)}{\cos((\phi_1 + \phi_2)/2)\sin((\phi_1 + \phi_2)/2)} \\ &= \frac{\sin((\theta + \phi_1)/2)\sin((\theta + \phi_2)/2)}{\sin((\phi_1 + \phi_2)/2)\sin((\phi_1 + \phi_2)/2)}. \end{aligned} \quad (\text{A6})$$

Therefore all three equations in this overdetermined system are only satisfied when

$$\phi_1 = \theta + 2k\pi \Rightarrow x_1 = 1, x_2 = 0 \quad (\text{A7})$$

or

$$\phi_2 = \theta + 2k\pi \Rightarrow x_1 = 0, x_2 = 1, \quad (\text{A8})$$

where $k \in \mathbb{Z}$, i.e. the trivial case of only a single channel in the decomposition. So, for multiple channels in the decomposition, at least three channels are required.

Appendix B. LCC decomposition for $\{\mathcal{I}, \mathcal{T}_n^{\frac{1}{n}}, \mathcal{Z}\}$ channels

In the main text, we showed that the optimal decompositions found using convex optimisation were of the form $\{\mathcal{I}, \mathcal{T}_n^{\frac{1}{n}}, \mathcal{Z}\}$ for both ideal and noisy $\mathcal{T}_n^{\frac{1}{n}}$ channels. In the following sections, we derive the l_1 -norm and γ corresponding to these optimal decompositions.

B.1. Ideal case

A $\mathcal{T}_n^{\frac{1}{n}}$ channel can be written as a $\mathcal{Z}$ -rotation channel with rotation angle ϕ ,

$$\begin{aligned} \mathcal{R}_Z^\phi(\rho) &= e^{-i\frac{\phi}{2}Z} \rho e^{i\frac{\phi}{2}Z} \\ &= \cos^2\left(\frac{\phi}{2}\right)\rho + \sin^2\left(\frac{\phi}{2}\right)Z\rho Z + i\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right)[\rho Z - Z\rho], \end{aligned} \quad (\text{B1})$$

where $\phi = \frac{\pi}{4n}$. Recall that $n = 2^{i-1}$ where $i \in \mathbb{Z}_{\geq 0}$ such that $i = 0$ is the $\mathcal{S}$ channel, $i = 1$ is the $\mathcal{T}$ channel, etc.

For an LCC decomposition in terms of $\{\mathcal{I}, \mathcal{R}_z^\phi, \mathcal{Z}\}$ channels, we obtain

$$\begin{aligned}
 \mathcal{R}_z^\theta(\rho) &= x_0 \mathcal{I}(\rho) + x_1 \mathcal{R}_z^\phi(\rho) + x_2 \mathcal{Z}(\rho) \\
 &= x_0 \rho + x_1 \left(\cos^2\left(\frac{\phi}{2}\right) \rho + i \cos\left(\frac{\phi}{2}\right) \sin\left(\frac{\phi}{2}\right) [\rho Z - Z\rho] + \sin^2\left(\frac{\phi}{2}\right) Z\rho Z \right) + x_2 Z\rho Z \\
 &= \left(x_0 + \cos^2\left(\frac{\phi}{2}\right) x_1 \right) \rho + i \cos\left(\frac{\phi}{2}\right) \sin\left(\frac{\phi}{2}\right) x_1 [\rho Z - Z\rho] + \left(\sin^2\left(\frac{\phi}{2}\right) x_1 + x_2 \right) Z\rho Z \\
 &= \cos^2\left(\frac{\theta}{2}\right) \rho + i \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right) [\rho Z - Z\rho] + \sin^2\left(\frac{\theta}{2}\right) Z\rho Z.
 \end{aligned}$$

The coefficients x_i can be determined by solving three simultaneous equations:

$$x_0 + \cos^2\left(\frac{\phi}{2}\right) x_1 = \cos^2\left(\frac{\theta}{2}\right), \quad (\text{B2a})$$

$$\cos\left(\frac{\phi}{2}\right) \sin\left(\frac{\phi}{2}\right) x_1 = \cos\left(\frac{\theta}{2}\right) \sin\left(\frac{\theta}{2}\right), \quad (\text{B2b})$$

$$\sin^2\left(\frac{\phi}{2}\right) x_1 + x_2 = \sin^2\left(\frac{\theta}{2}\right), \quad (\text{B2c})$$

with solutions given by

$$x_0 = \cos^2\left(\frac{\theta}{2}\right) - \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \cos^2\left(\frac{\phi}{2}\right), \quad (\text{B3a})$$

$$x_1 = \frac{\sin(\theta)}{\sin(\phi)}, \quad (\text{B3b})$$

$$x_2 = \sin^2\left(\frac{\theta}{2}\right) - \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \sin^2\left(\frac{\phi}{2}\right). \quad (\text{B3c})$$

The l_1 -norm of this decomposition for target rotation angles $0 \leq \theta \leq \phi$ is

$$\begin{aligned}
 \Lambda_G(\mathcal{R}_z^\theta) &= |x_0| + |x_1| + |x_2| \\
 &= \left| \cos^2\left(\frac{\theta}{2}\right) - \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \cos^2\left(\frac{\phi}{2}\right) \right| + \left| \frac{\sin(\theta)}{\sin(\phi)} \right| + \left| \sin^2\left(\frac{\theta}{2}\right) - \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \sin^2\left(\frac{\phi}{2}\right) \right| \\
 &= \cos^2\left(\frac{\theta}{2}\right) - \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \cos^2\left(\frac{\phi}{2}\right) + \frac{\sin(\theta)}{\sin(\phi)} - \sin^2\left(\frac{\theta}{2}\right) + \left(\frac{\sin(\theta)}{\sin(\phi)} \right) \sin^2\left(\frac{\phi}{2}\right) \\
 &= \cos(\theta) + \frac{\sin(\theta)}{\sin(\phi)} (1 - \cos(\phi)).
 \end{aligned} \quad (\text{B4})$$

noting that $x_2 \leq 0$ while $x_0, x_1 \geq 0$ for this range of θ .

For small rotation angles θ , we approximate

$$\begin{aligned}
 \Lambda_G(\mathcal{R}_z^\theta) &= \cos(\theta) + \sin(\theta) \left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right) \\
 &\approx 1 - \frac{\theta^2}{2} + \theta \left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right) \\
 &\approx 1 + \theta \left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right).
 \end{aligned} \quad (\text{B5})$$

Thus γ in equation (14) is obtained as follows

$$\gamma(\theta) = \frac{\ln(\Lambda_G(\mathcal{R}_z^\theta))}{\ln(\Lambda_G(\mathcal{R}_z^\phi))} = \frac{\ln(1 + \theta)}{\ln(1 + \theta \left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right))} \approx \frac{\theta}{\left(\frac{1 - \cos(\phi)}{\sin(\phi)} \right) \theta} = \frac{\sin(\phi)}{1 - \cos(\phi)} = \cot\left(\frac{\phi}{2}\right). \quad (\text{B6})$$

B.2. Dephased $\mathcal{T}_n^{\frac{1}{n}}$ case

We consider a dephasing noise channel as given by equation (21). In the main text, we define an effective dephasing noise p_{eff} , which we use in the following derivation.

The dephasing noise channel can applied to $\mathcal{R}_z^\phi$ as follows:

$$\begin{aligned}\varepsilon(\mathcal{R}_z^\phi(\rho)) &= (1 - p_{\text{eff}})(\mathcal{R}_z^\phi(\rho)) + p_{\text{eff}}Z(\mathcal{R}_z^\phi(\rho))Z \\ &= \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]\rho + (1 - 2p_{\text{eff}})\left[i\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right)\right][\rho Z - Z\rho] \\ &\quad + \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]Z\rho Z.\end{aligned}\quad (\text{B7})$$

The LCC decomposition of the $\mathcal{R}_z^\theta(\rho)$ channel in terms of $\{\mathcal{I}, \varepsilon(\mathcal{R}_z^\phi), \mathcal{Z}\}$ is given by

$$\begin{aligned}\mathcal{R}_z^\theta(\rho) &= x_0\mathcal{I}(\rho) + x_1\varepsilon(\mathcal{R}_z^\phi(\rho)) + x_2\mathcal{Z}(\rho) \\ &= x_0\rho + x_1\left[\left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]\rho + (1 - 2p_{\text{eff}})\left[i\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right)\right][\rho Z - Z\rho] \right. \\ &\quad \left. + \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]Z\rho Z\right] + x_2Z\rho Z\end{aligned}\quad (\text{B8a})$$

$$\begin{aligned}&= (x_0 + \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]x_1)\rho + i(1 - 2p_{\text{eff}})\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right)x_1[\rho Z - Z\rho] \\ &\quad + \left(\left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]x_1 + x_2\right)Z\rho Z \\ &= \cos^2\left(\frac{\theta}{2}\right)\rho + i\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)[\rho Z - Z\rho] + \sin^2\left(\frac{\theta}{2}\right)Z\rho Z\end{aligned}\quad (\text{B8b})$$

The coefficients x_i can be determined by solving three simultaneous equations:

$$x_0 + \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]x_1 = \cos^2\left(\frac{\theta}{2}\right), \quad (\text{B9a})$$

$$(1 - 2p_{\text{eff}})\cos\left(\frac{\phi}{2}\right)\sin\left(\frac{\phi}{2}\right)x_1 = \cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right), \quad (\text{B9b})$$

$$\left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]x_1 + x_2 = \sin^2\left(\frac{\theta}{2}\right). \quad (\text{B9c})$$

with solutions given by

$$x_0 = \cos^2\left(\frac{\theta}{2}\right) - \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} \quad (\text{B10a})$$

$$x_1 = \frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} \quad (\text{B10b})$$

$$x_2 = \sin^2\left(\frac{\theta}{2}\right) - \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)}. \quad (\text{B10c})$$

The l_1 -norm of this decomposition is

$$\begin{aligned}\Lambda_G(\mathcal{R}_z^\theta) &= |x_0| + |x_1| + |x_2| \\ &= \left|\cos^2\left(\frac{\theta}{2}\right) - \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)}\right| + \left|\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)}\right| \\ &\quad + \left|\sin^2\left(\frac{\theta}{2}\right) - \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right]\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)}\right| \\ &= \cos^2\left(\frac{\theta}{2}\right) - \left[\cos^2\left(\frac{\phi}{2}\right) - p_{\text{eff}}\cos(\phi)\right]\frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)} + \frac{\sin(\theta)}{(1 - 2p_{\text{eff}})\sin(\phi)}\end{aligned}$$

$$\begin{aligned}
& -\sin^2\left(\frac{\theta}{2}\right) + \left[\sin^2\left(\frac{\phi}{2}\right) + p_{\text{eff}}\cos(\phi)\right] \frac{\sin(\theta)}{(1-2p_{\text{eff}})\sin(\phi)} \\
& = \cos(\theta) - \frac{\sin(\theta)}{(1-2p_{\text{eff}})\sin(\phi)} \left((1-2p_{\text{eff}})\cos(\phi) - 1\right) \\
& = \cos(\theta) + \sin(\theta) \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right).
\end{aligned} \tag{B11}$$

for target rotation angles $0 \leq \theta \ll \phi$. For small θ , we can approximate

$$\begin{aligned}
\Lambda_G(\mathcal{R}_z^\theta) &= \cos(\theta) + \sin(\theta) \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right) \\
&\approx 1 - \frac{\theta^2}{2} + \theta \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right) \\
&\approx 1 + \theta \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right).
\end{aligned} \tag{B12}$$

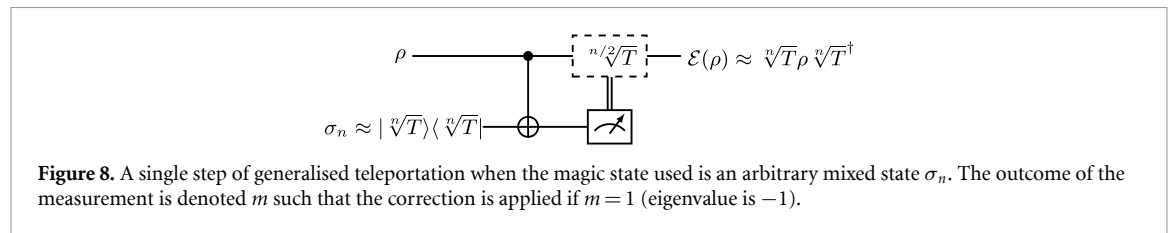
Therefore, γ becomes

$$\begin{aligned}
\gamma(\theta) &= \frac{\ln(\Lambda_G(\mathcal{R}_z^\theta))}{\ln(\Lambda_G(\mathcal{R}_z^\theta))} \\
&= \frac{\ln(1+\theta)}{\ln\left(1 + \theta \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right)\right)} \\
&\approx \frac{\theta}{\theta \left(\frac{1}{(1-2p_{\text{eff}})} \csc(\phi) - \cot(\phi)\right)} \\
&= \frac{1}{\frac{\csc(\phi)}{(1-2p_{\text{eff}})} - \cot(\phi)}.
\end{aligned} \tag{B13}$$

Appendix C. Noise in generalised gate teleportation circuit

We can show that the generalised gate teleportation circuit in figure 2 results in a rotation that differs from the target rotation θ by either dephasing noise, or coherent errors. We note that coherent errors can be handled through calibration, and accurate calibration is an underlying assumption required for the estimation of the level of dephasing noise present.

We proceed by considering a single step of generalised teleportation with noisy states as shown in figure 8.



The state σ_n will have an eigenvalue decomposition

$$\sigma_n = |\phi_0\rangle\langle\phi_0| + |\phi_1\rangle\langle\phi_1| \tag{C1}$$

where $|\phi_i\rangle$ are not individually normalised, but $\langle\phi_0|\phi_0\rangle + \langle\phi_1|\phi_1\rangle = 1$. It is a common assumption to take $|\phi_0\rangle = \sqrt{1-p}|\sqrt{T}\rangle$ and $|\phi_1\rangle$ as its orthogonal partner. However, the purpose of this appendix is to allow for a fully general σ_n .

After measurement outcome m , ρ gets mapped to

$$\mathcal{E}_m(\rho) = \kappa_{0,m}\rho\kappa_{0,m}^\dagger + \kappa_{1,m}\rho\kappa_{1,m}^\dagger, \tag{C2}$$

where $\kappa_{j,m}$ is the Kraus operator corresponding to teleportation using a pure state $|\phi_j\rangle$ after measurement outcome m . Denoting $|\phi_j\rangle = A_j|0\rangle + B_j|1\rangle$, calculating the action of the teleportation circuit one finds

$$\kappa_{0,0} = A_0|0\rangle\langle 0| + B_0|1\rangle\langle 1| = \left(\frac{A_0+B_0}{2}\right)I + \left(\frac{A_0-B_0}{2}\right)Z, \quad (\text{C3a})$$

$$\kappa_{0,1} = A_0|1\rangle\langle 1| + B_0|0\rangle\langle 0| = \left(\frac{A_0+B_0}{2}\right)I + \left(\frac{-A_0+B_0}{2}\right)Z, \quad (\text{C3b})$$

$$\kappa_{1,0} = A_1|0\rangle\langle 0| + B_1|1\rangle\langle 1| = \left(\frac{A_1+B_1}{2}\right)I + \left(\frac{A_1-B_1}{2}\right)Z, \quad (\text{C3c})$$

$$\kappa_{1,1} = A_1|1\rangle\langle 1| + B_1|0\rangle\langle 0| = \left(\frac{A_1+B_1}{2}\right)I + \left(\frac{-A_1+B_1}{2}\right)Z \quad (\text{C3d})$$

and so every relevant Kraus operator is diagonal in the Z -basis. Therefore, expanding out the channel we find that

$$\mathcal{E}_m(\rho) = E_{0,0}^m \rho + E_{0,1}^m \rho Z + E_{1,0}^m Z \rho + E_{1,1}^m Z \rho Z. \quad (\text{C4})$$

where $E_{k,\ell}^m$ are numbers depending on A_0, A_1, B_0, B_1 that in turn depend on σ_n .

We are not quite done because in the case where $m = 1$, we need to apply a correction (i.e. using a magic state from one level lower in the Clifford hierarchy $\sigma_{n/2}$). Nonetheless, this correction process will also perform diagonal Kraus operators, and since the group of diagonal operators is closed in multiplication, we conclude that generalised teleportation will result in some channel $\mathcal{E}$ that is a sum of diagonal Kraus operators. Expanding these out in the Pauli basis, we will again have an expression of the form

$$\mathcal{E}(\rho) = E_{0,0} \rho + E_{0,1} \rho Z + E_{1,0} Z \rho + E_{1,1} Z \rho Z, \quad (\text{C5})$$

where $E_{i,j}$ depend on the density matrices of all magic states $\{\sigma_n, \sigma_{n/2}, \sigma_{n/4}, \dots, \sigma_1\}$ in the generalised teleportation circuit. This means a large number of parameters are involved, but fortunately we can proceed with our proof using only the form of equation (C5) and the following observation: once all corrections are completed (we mix over all measurement outcomes) the full channel $\mathcal{E}$ must be CPTP. Note that, in contrast, each component $\mathcal{E}_m$ will be completely positive (CP) but not necessarily trace preserving (TP). Therefore, under the Choi–Jamiołkowski isomorphism, the Choi state for this channel must be a physical density matrix (Hermitian, positive semi-definite and trace normalised to unity). Furthermore, due to the form of equation (C5), the Choi state has the form

$$\Phi_{\mathcal{E}} = (\mathcal{E} \otimes \mathcal{I})(|\Phi\rangle\langle\Phi|) = \begin{pmatrix} a_{0,0} & 0 & 0 & a_{0,1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ a_{1,0} & 0 & 0 & a_{1,1} \end{pmatrix}. \quad (\text{C6})$$

where for instance $a_{1,1} = (E_{0,0} - E_{0,1} - E_{1,0} + E_{1,1})/2$. In particular, the Choi state has at most 2 non-zero eigenvalues, and by the CPTP property we can denote these as p and $1-p$ (with $1 \leq p \leq 0$), so they can be interpreted as probabilities. Therefore,

$$\Phi_{\mathcal{E}} = p|K_0\rangle\langle K_0| + (1-p)|K_1\rangle\langle K_1| \quad (\text{C7})$$

where $|K_j\rangle$ are a pair of orthogonal pure states supported on the non-trivial 2×2 submatrix of equation (C6). We may now reverse the Choi–Jamiołkowski isomorphism, so that

$$\mathcal{E}(\rho) = pK_0\rho K_0^\dagger + (1-p)K_1\rho K_1^\dagger \quad (\text{C8})$$

where K_j is the single Kraus operator isomorphic to the state $|K_j\rangle\langle K_j|$. Since the state representation is restricted to a specific submatrix, we can conclude that K_j are unitary operators diagonal in the Z basis and we are free to choose the global phase. Therefore, there exists angles φ_j such that $K_j = R_z(\varphi_j)$. Finally, since $|K_j\rangle$ are orthogonal to each other, we can conclude that $\text{Tr}[K_1^\dagger K_0] = 0$. This orthogonality entails that φ_1 is such that $K_1 = K_0 Z$ upto a global phase. This brings us to the final form

$$\mathcal{E}(\rho) = p\mathcal{R}_z^{\varphi_0}(\rho) + (1-p)Z\mathcal{R}_z^{\varphi_0}(\rho)Z \quad (\text{C9})$$

which has the claimed form of dephasing noise and potentially a coherent error which deviates from the ideal angle ϕ by a phase of $\delta = \varphi_0 - \phi$.

We have assumed throughout that physical noise is well characterised. The impact of changes in the magic states σ_n affects the form of the resulting channel $\mathcal{E}$. If there is an undesired coherent error, we can adjust the angle of the prepared magic state σ_n to eliminate this, thereby leaving only dephasing error as assumed throughout the main text.

Appendix D. Supplementary Data

Table 3. $\ln(\Lambda_G(\mathcal{R}_z^\theta))$ as a function of target rotation angle θ for different optimal decompositions of the form $\{\mathcal{I}, \varepsilon(\mathcal{T}^\frac{1}{2}), \mathcal{Z}\}$ including the optimal Clifford decomposition of $\{\mathcal{I}, \mathcal{S}, \mathcal{Z}\}$. The values presented here correspond to the relative values used to calculate γ in table 1.

n	ϕ	$\mathcal{R}_z^\phi$	$\ln(\Lambda_G(\mathcal{R}_z^\theta))$ (Small θ)			
			$p = 0.01\%$	$p = 0.1\%$	$p = 0.5\%$	$p = 1.0\%$
0	$\frac{\pi}{2}$	$\mathcal{S}$	1.00×10^{-07}	1.00×10^{-07}	1.00×10^{-07}	1.00×10^{-07}
1	$\frac{\pi}{4}$	$\mathcal{T}$	4.14×10^{-08}	4.17×10^{-08}	4.28×10^{-08}	4.43×10^{-08}
2	$\frac{\pi}{8}$	$\sqrt{\mathcal{T}}$	2.00×10^{-08}	2.07×10^{-08}	2.39×10^{-08}	2.40×10^{-08}
4	$\frac{\pi}{16}$	$\sqrt[4]{\mathcal{T}}$	1.00×10^{-08}	1.16×10^{-08}	1.90×10^{-08}	2.84×10^{-08}
8	$\frac{\pi}{32}$	$\sqrt[8]{\mathcal{T}}$	5.30×10^{-09}	8.75×10^{-09}	2.44×10^{-08}	4.47×10^{-09}

ORCID iDs

Surabhi Luthra  0009-0001-8029-5835
 Alexandra E Moylett  0000-0003-0163-5262
 Dan E Browne  0000-0003-3001-158X
 Earl T Campbell  0000-0002-5627-0021

References

- [1] Google Quantum AI and Collaborators 2024 Quantum error correction below the surface code threshold *Nature* **638** 920–6
- [2] Bluvstein D *et al* 2024 Logical quantum processor based on reconfigurable atom arrays *Nature* **626** 58
- [3] Rodriguez P S *et al* 2024 Experimental Demonstration of Logical Magic State Distillation (arXiv:2412.15165 [quant-ph])
- [4] Reichardt B W *et al* 2024 Demonstration of quantum computation and error correction with a tesseract code (arXiv:2409.04628 [quant-ph])
- [5] Campbell E 2024 A series of fast-paced advances in quantum error correction *Nat. Rev. Phys.* **6** 160
- [6] Campbell E T 2022 Early fault-tolerant simulations of the Hubbard model *Quantum Sci. Technol.* **7** 015007
- [7] Katabarwa A, Gratsea K, Caesura A and Johnson P D 2024 Early Fault-Tolerant Quantum Computing *PRX Quantum* **5** 020101
- [8] Toshio R, Akahoshi Y, Fujisaki J, Oshima H, Sato S and Fujii K 2024 Practical quantum advantage on partially fault-tolerant quantum computer (arXiv:2408.14848 [quant-ph])
- [9] Preskill J 2025 Beyond nisq: The megaquop machine *ACM Trans. Quantum Comput.* **6** 1–7
- [10] Gottesman D 1998 The Heisenberg Representation of Quantum Computers (arXiv:quant-ph/9807006)
- [11] Bravyi S and Kitaev A 2005 universal quantum computation with ideal Clifford gates and noisy ancillas *Phys. Rev. A* **71** 022316
- [12] Litinski D 2019 Magic State Distillation: Not as Costly as You Think *Quantum* **3** 205
- [13] Gidney C and Fowler A G 2019 Efficient magic state factories with a catalyzed $|\text{CCZ}\rangle$ to $2|T\rangle$ transformation *Quantum* **3** 135
- [14] Wills A, Hsieh M-H and Yamasaki H 2024 Constant-Overhead Magic State Distillation (arXiv:2408.07764 [quant-ph])
- [15] O’Gorman J and Campbell E T 2017 Quantum computation with realistic magic-state factories *Phys. Rev. A* **95** 032338
- [16] Lee J, Berry D W, Gidney C, Huggins W J, McClean J R, Wiebe N and Babbush R 2021 Even more efficient quantum computations of chemistry through tensor hypercontraction *PRX Quantum* **2** 030305
- [17] Gidney C 2025 How to factor 2048 bit rsa integers with less than a million noisy qubits (arXiv:2505.15917 [quant-ph])
- [18] Georges T N, Bothe M, Sünderhauf C, Berntson B K, Izsák R and Ivanov A V 2025 Quantum simulations of chemistry in first quantization with any basis set *npj Quantum Inform.* **11** 55
- [19] Peruzzo A, McClean J, Shadbolt P, Yung M-H, Zhou X-Q, Love P J, Aspuru-Guzik A and O’Brien J L 2014 A variational eigenvalue solver on a photonic quantum processor *Nat. Commun.* **5** 4213
- [20] Somma R D 2019 Quantum eigenvalue estimation via time series analysis *New J. Phys.* **21** 123025
- [21] Lin L and Tong Y 2022 Heisenberg-Limited Ground-State Energy Estimation for Early Fault-Tolerant Quantum Computers *PRX Quantum* **3** 010318
- [22] Wan K, Berta M and Campbell E T 2022 Randomized Quantum Algorithm for Statistical Phase Estimation *Phys. Rev. Lett.* **129** 030503
- [23] Blunt N S, Caune L, Izsák R, Campbell E T and Holzmann N 2023 Statistical phase estimation and error mitigation on a superconducting quantum processor *PRX Quantum* **4** 040341
- [24] Günther J, Witteveen F, Schmidhuber A, Miller M, Christandl M and Harrow A 2025 Phase estimation with partially randomized time evolution (arXiv:2503.05647)
- [25] Temme K, Bravyi S and Gambetta J M 2017 Error Mitigation for Short-Depth Quantum Circuits *Phys. Rev. Lett.* **119** 180509
- [26] Cai Z, Babbush R, Benjamin S C, Endo S, Huggins W J, Li Y, McClean J R and O’Brien T E 2023 Quantum Error Mitigation *Rev. Mod. Phys.* **95** 045005
- [27] Piveteau C, Sutter D, Bravyi S, Gambetta J M and Temme K 2021 Error Mitigation for universal Gates on Encoded Qubits *Phys. Rev. Lett.* **127** 200505

- [28] Suzuki Y, Endo S, Fujii K and Tokunaga Y 2022 Quantum Error Mitigation as a universal Error Reduction Technique: Applications from the NISQ to the Fault-Tolerant Quantum Computing Eras *PRX Quantum* **3** 010345
- [29] Ross N J and Selinger P 2016 Optimal ancilla-free Clifford+T approximation of z-rotations (arXiv:1403.2975)
- [30] Veitch V, Hamed Mousavian S A, Gottesman D and Emerson J 2014 The resource theory of stabilizer quantum computation *New J. Phys.* **16** 013009
- [31] Howard M and Campbell E 2017 Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing *Phys. Rev. Lett.* **118** 090501
- [32] Seddon J R, Regula B, Pashayan H, Ouyang Y and Campbell E T 2021 Quantifying Quantum Speedups: Improved Classical Simulation From Tighter Magic Monotones *PRX Quantum* **2** 010345
- [33] Seddon J R and Campbell E T 2019 Quantifying magic for multi-qubit operations *Proc. R. Soc. A* **475** 20190251
- [34] Bravyi S, Browne D, Calpin P, Campbell E, Gosset D and Howard M 2019 Simulation of quantum circuits by low-rank stabilizer decompositions *Quantum* **3** 181
- [35] Koczor B, Morton J J L and Benjamin S C 2024 Probabilistic Interpolation of Quantum Rotation Angles *Phys. Rev. Lett.* **132** 130602
- [36] Pashayan H, Wallman J J and Bartlett S D 2015 Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities *Phys. Rev. Lett.* **115** 070501
- [37] Gottesman D and Chuang I L 1999 Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations *Nature* **402** 390
- [38] Campbell E T and O’Gorman J 2016 An efficient magic state approach to small angle rotations *Quantum Sci. Technol.* **1** 015007
- [39] Li Y 2015 A magic state’s fidelity can be superior to the operations that created it *New J. Phys.* **17** 023037
- [40] Hubbard J and Flowers B H 1963 Electron correlations in narrow energy bands *Proc. R. Soc. A* **276** 238
- [41] Editors of Nature Physics 2013 The Hubbard model at half a century *Nat. Phys.* **9** 523
- [42] Cade C, Mineh L, Montanaro A and Stanisic S 2020 Strategies for solving the Fermi-Hubbard model on near-term quantum computers *Phys. Rev. B* **102** 235122
- [43] Dagotto E 1994 Correlated electrons in high-temperature superconductors *Rev. Mod. Phys.* **66** 763
- [44] Imada M, Fujimori A and Tokura Y 1998 Metal-insulator transitions *Rev. Mod. Phys.* **70** 1039
- [45] Babbush R, McClean J R, Newman M, Gidney C, Boixo S and Neven H 2021 Focus beyond quadratic speedups for error-corrected quantum advantage *PRX Quantum* **2** 010103
- [46] Litinski D 2019 A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery *Quantum* **3** 128
- [47] Endo S, Benjamin S C and Li Y 2018 Practical Quantum Error Mitigation for Near-Future Applications *Phys. Rev. X* **8** 031027
- [48] Diamond S and Boyd S 2016 CVXPY: A Python-embedded modeling language for convex optimization *J. Mach. Learn. Res.* **17** 1
- [49] Agrawal A, Verschueren R, Diamond S and Boyd S 2018 A rewriting system for convex optimization problems *J. Control Decis.* **5** 42
- [50] Note that if Z errors occur on both the $|T\rangle$ and $|T^{\frac{1}{2}}\rangle$ states, then the overall effect on the $T^{\frac{1}{2}}$ gate is $Z^2 = I$, meaning this case is equivalent to no error occurring.
- [51] Kivlichan I D *et al* 2020 Improved Fault-Tolerant Quantum Simulation of Condensed-Phase Correlated Electrons via Trotterization *Quantum* **4** 296
- [52] Campbell E 2017 Shorter gate sequences for quantum computing by mixing unitaries *Phys. Rev. A* **95** 042306
- [53] Kliuchnikov V, Lauter K, Minko R, Paetznick A and Petit C 2023 Shorter quantum circuits via single-qubit gate approximation *Quantum* **7** 1208
- [54] Yoshioka N, Akibue S, Morisaki H, Morisaki H, Tsubouchi K and Suzuki Y 2025 Error crafting in mixed quantum gate synthesis *npj Quantum Inform.* **11** 95
- [55] Choi H *et al* 2023 Fault Tolerant Non-Clifford State Preparation for Arbitrary Rotations (arXiv:2303.17380 [quant-ph])
- [56] Gidney C, ShuTTY N and Jones C 2024 Magic state cultivation: Growing T states as cheap as CNOT gates (arXiv:2409.17595 [quant-ph])
- [57] Horsman D, Fowler A G, Devitt S and Meter R V 2012 Surface code quantum computing by lattice surgery *New J. Phys.* **14** 123011
- [58] Gehér G P, McLauchlan C, Campbell E T, Moylett A E and Crawford O 2024 Error-corrected Hadamard gate simulated at the circuit level *Quantum* **8** 1394
- [59] Boixo S, Isakov S V, Smelyanskiy V N, Babbush R, Ding N, Jiang Z, Bremner M J, Martinis J M and Neven H 2018 Characterizing quantum supremacy in near-term devices *Nat. Phys.* **14** 595
- [60] Luthra S, Moylett A E, Browne D E and Campbell E T 2025 Dataset for “Unlocking early fault-tolerant quantum computing with mitigated magic dilution” *Zenodo* (<https://doi.org/10.5281/zenodo.15427988>)