Unitary averaging with fault and loss tolerance

Ryan J. Marshman,[1] Deepesh Singh,[1] Timothy C. Ralph,[1] and Austin P. Lund[2,1]
[1]*Centre for Quantum Computation and Communication Technology, School of Mathematics and Physics, University of Queensland, Brisbane, Queensland 4072, Australia*
[2]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*

(Received 7 December 2023; revised 26 May 2024; accepted 29 May 2024; published 27 June 2024)

We consider the impact of the unitary-averaging framework on single- and two-mode linear optical gates. We demonstrate that this allows a trade-off between the probability of success and gate fidelity, with perfect fidelity gates being achievable for a finite decrease in the probability of success, at least in principle. Furthermore, we show that the encoding and decoding errors in the averaging scheme can also be suppressed up to the first order. We also look at how unitary averaging can work in conjunction with existing error-correction schemes. Specifically, we consider how parity encoding might be used to counter the extra loss due to the decreased probability of success, with the aim of achieving fault tolerance. We also consider how unitary averaging might be utilized to expand the parameter space in which fault tolerance may be achievable using standard fault-tolerant schemes.

DOI: 10.1103/PhysRevA.109.062436

I. INTRODUCTION

With noisy intermediate-scale quantum (NISQ) devices becoming more common, the attention is shifting to the creation of full-scale quantum computing devices [1]. These future devices will differ from current technology in their scale, which necessitates the development of practical methods of dealing with noise in quantum devices. Specifically, full-scale quantum computation will require the used of fault-tolerant schemes. However, these require enormous overhead in resources to be implemented. Quantum computing architectures and error-correction schemes that offer a reduction in this overhead will be necessary for at least near future quantum computing devices and are thus the focus of much interest. Indeed, improvements in this space which are easily implementable would be useful both in the NISQ era and in the longer term.

Linear optics provides a promising platform for achieving the transition from existing technologies such as boson sampling [2] to universal quantum computing realizations such as fusion-based quantum computing [3]. The implementation of deterministic two-qubit entangling gates with noninteracting photons, photon loss, and the buildup of errors throughout the quantum circuits represents a major challenge for optical quantum computers. The first two issues can be addressed by loss-tolerant encoding [4] to enable recovery from failed probabilistic components. As such, loss tolerance must be a fundamental component of any optical quantum computer, a fact that our scheme relies on to address the limitation of its probabilistic nature.

We explore the framework of unitary averaging (UA) [5–7], which allows one to alleviate the effect of imperfections within the applied transformations. Unitary averaging employs N copies of a noisy transformation circuit one wishes to implement to passively, but probabilistically, reduce the buildup of errors. The scheme has a simple implementation with linear optics, and so we will use this as the example system, although it is compatible with more than just linear optics. Unitary averaging implements a transformation equivalent to the average of a given set of transformations. In the limit when all the individual transformations are close approximations of some target, their average is a good approximation of the target in total variation distance. The resulting improvement in the gate fidelity, however, comes at the cost of success probability, as we will discuss throughout this paper.

As is the case with all error-correction schemes, encoding errors present a challenge to their practical usefulness. An exploration of the same is therefore considered for the UA framework here. It is shown that errors present in any encoding utilized in the UA scheme are naturally suppressed to the first order.

Moreover, the trade-off between gate fidelity and loss may be useful given the necessity of loss protection in any realistic optical quantum computer due to both photon absorption and the probabilistic nature of optical quantum computation [8,9]. We consider the loss-tolerant parity-encoding scheme and show its compatibility with UA. Thus, in the limit of large redundancy, UA could be employed to ensure all logical errors are converted into heralded losses which are recoverable by the parity encoding.

Finally, we consider how UA may be used to expand the parameter space for which fault tolerance can be achieved, performing a simplified analysis for a few example error codes. Specifically, we consider some older fault-tolerant codes (the seven-qubit Steane code and the 23-qubit Golay code [10]) for which the benefit of UA is seen to be significant, as well as more modern surface-code-based implementations [3,11] for which the effect of UA is more modest.

The structure of this paper is as follows. We begin by summarizing the known results on UA in Sec. II. Section III explores the suppression of encoding errors in single-mode unitary-averaged gates, while Sec. IV details the effect unitary

averaging has on an arbitrary single-mode unitary in terms of both the gate fidelity and probability of success. Section V discusses how two-mode gates are implemented and protected in a similar fashion. We then introduce the loss-tolerant scheme of parity encoding in Sec. VI before highlighting how it can be used together with UA in Sec. VII. Section VIII then discusses the use of UA with the threshold for fault tolerance. A brief summary is provided in Sec. IX.

II. UNITARY AVERAGING

Given access to multiple unitaries, the unitary-averaging framework allows one to apply an average of these unitaries on the intended modes, with the success probability of this transformation depending on the exact value of the unitaries. Unitary averaging acts to apply an averaged-unitary evolution [5,6], given by

$$\hat{\mathcal{U}} = \frac{1}{N}\sum_{j=1}^{N} \hat{U}_j, \tag{1}$$

on a target set of M modes using N noisy copies of a target unitary $\hat{U}_T$, each labeled $\hat{U}_j$. There is also an accompanying $(N-1)M$ set of error modes which must be heralded in the vacuum state. If each applied unitary $\hat{U}_j$ is approximately implementing a target unitary $\hat{U}_T$ with an independent, identically distributed, and unbiased noise, then UA will apply an averaged unitary $\hat{\mathcal{U}}$. This itself will be a stochastic operator which approximates the target unitary $\hat{U}_T$, with variance reduced by a factor of N when compared to each of the original transformations $\hat{U}_j$. As demonstrated later, we can further write the individual transformation in the form $\hat{U}_j = \hat{U}_T + \hat{E}_j$, where $\hat{E}_j$ is the stochastic operator containing all of the noise terms. Note that the use of $\mathcal{U}$ rather than U is to remind the reader that the resulting transformation is nonunitary but aims to be as close to unitary as possible. Specifically, the target unitary can be implemented arbitrarily accurately using a sufficiently large N, after renormalization due to heralding each error mode in the vacuum state. The cost of this reduction in variation of the applied transformation is that it is implemented probabilistically. The probability of success $P_s(N)$ depends on both the variance in the individual unitaries and number of copies used N. Detecting a photon at any of the error heralding modes applies the transformation

$$\hat{\mathcal{U}}_e = \frac{1}{N}\sum_{j=1}^{N} f_j \hat{U}_j, \tag{2}$$

where the weights f_j are phase factors such that $|f_j| = 1\,\forall\, j$. The values f_j depend on the encoding and decoding used in the averaging process. When the encoding unitary is $H^{\otimes i}$ for any i, as will be used throughout this paper, $f_j = \pm 1$ such that at least one $f_j = -1$. Furthermore, depending on the nature of the detectors used, it may destroy the state. As such, observing a photon in the heralding modes applies an unintended transformation which we treat as unrecoverable. As such, UA can be viewed as a mapping between logical errors and heralded loss. The probability of such a heralded loss occurring scales proportionally to the variance of the applied unitaries $\hat{U}_j$ and the amount of averaging N.

Throughout this paper we are concerned only with the success modes. As such, we consider the output state to only be the correct modes, dropping the error heralding modes. The output state is thus

$$\hat{\rho}(N) = \hat{\mathcal{U}}(N)|\psi\rangle\langle\psi|\hat{\mathcal{U}}^{\dagger}(N), \tag{3}$$

which is an unnormalized state, where $|\psi\rangle$ is the initial state. The normalization is returned though the postselection process, which is also the source of the probabilistic nature of the correction being applied. The probability of success $P_s(N)$ can be defined by the amplitude of the unnormalized state which can be calculated for a general input state using

$$P_s(N) = |\hat{\mathcal{U}}(N)|\psi\rangle|^2 = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle\psi|\hat{U}_j^{\dagger}\hat{U}_k|\psi\rangle. \tag{4}$$

Throughout this paper, we will reserve the subscripts j and k to indicate the noisy copy of the target unitary $\hat{U}_j$ or $\hat{U}_k$ or a parameter within said unitary. The (normalized) density operator after the postselection step is then given by

$$\hat{\rho}_{\rm ps} = [P_s(N)]^{-1}\hat{\rho}(N). \tag{5}$$

The other figure of merit to characterize the effect UA has on a gate is the fidelity. The fidelity encodes how likely it is that, upon measurement, the transformation returns the target state. The target state in this case is $|\Psi\rangle = \hat{U}_T|\psi\rangle$, where $|\psi\rangle$ is the initial state and $\hat{U}_T$ is the target unitary, which corresponds to $\hat{U}_j$ in the instance of no noise. The gate fidelity is then defined as

$$\mathcal{F}(N) = \langle\Psi|\rho_{\rm ps}(N)|\Psi\rangle. \tag{6}$$

Before we can start characterizing the effect UA has on single- and two-qubit gates, we must also consider the encoding and decoding steps and to what extent they impact on the output.

III. UNITARY-AVERAGING ENCODING ERRORS

The encoding and decoding can be achieved in multiple different manners, including using Hadamard encoding [5] or the more general W-state or quantum Fourier transform encodings [6,7]. In this instance we consider the Hadamard encoding, which is simple in its construction for scaling to higher levels of encoding (N). We can then take the gates to be encoded using only beam splitters, splitting the input evenly between each redundant physical gate. This can then be performed iteratively, giving $N = 2^n$ for $n \in \mathbb{N}$. This process is shown in Fig. 1 going from $N = 2$ to $N = 4$. This choice of encoding is also optimal in that it maintains constant optical depth of each interferometric path while also minimizing the optical depth. Using this method, the optical depth for each path increases logarithmically as $2\log_2(N)$ for N copies of the unitary.

Starting with $N = 2$, the output of a single-qubit-averaged gate after postselection is determined by

$$|\psi\rangle_{\rm out} = \langle\mathbf{0}|_{\epsilon}\hat{B}^{(2)}_{1\leftrightarrow 2}(\hat{U}_2 \otimes \hat{U}_1)\hat{B}^{(1)}_{1\leftrightarrow 2}|\psi\rangle_{\rm in}. \tag{7}$$

where $\hat{B}^{(i)}_{a\leftrightarrow b}$ acts to evenly mix the pair of modes a and b, each of which is then separately acted on by independent unitary $\hat{U}_{a,b}$, and $\langle\mathbf{0}|_{\epsilon}$ represents the projection onto the vacuum for the error modes. The superscript i serves to remind the reader

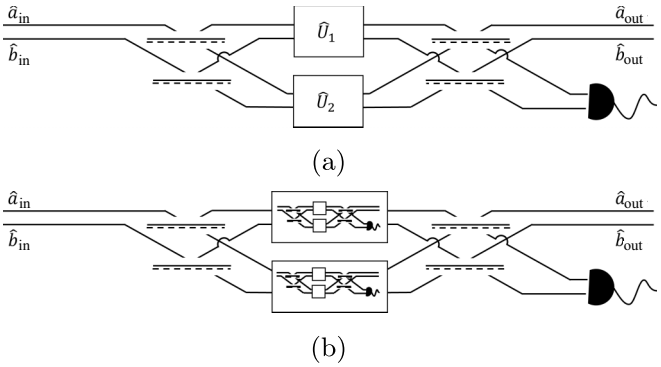


FIG. 1. Arbitrary dual-rail single-qubit unitary with unitary averaging for (a) $N = 2$ levels of redundancy and (b) similarly $N = 4$ but where each physical unitary $\hat{U}_j$ is itself an $N = 2$ encoded unitary. All beam splitters outside of $\hat{U}_j$ aim to be 50:50, enacting the operation $\hat{B}^{(i)}_{1\leftrightarrow 2}$ in Eq. (8).

that each beam splitter is independent and thus has its own unique noise associated with it. In a dual rail, this can be written as

$$\hat{B}^{(i)}_{1\leftrightarrow 2} = \begin{bmatrix} \sin(\theta_i) & 0 & \cos(\theta_i) & 0 \\ 0 & \sin(\theta_i') & 0 & \cos(\theta_i') \\ \cos(\theta_i) & 0 & -\sin(\theta_i) & 0 \\ 0 & \cos(\theta_i') & 0 & -\sin(\theta_i') \end{bmatrix}, \quad (8)$$

where $\theta_i^{(\prime)} = \frac{\pi}{2} + \delta\theta_i^{(\prime)}$ for $\delta\theta_i^{(\prime)} \ll \frac{\pi}{2}$, allowing us to write $\sin(\theta_i) \approx \frac{1}{\sqrt{2}} + \delta\theta_i^{(\prime)}$ and $\cos(\theta_i) \approx \frac{1}{\sqrt{2}} - \delta\theta_i^{(\prime)}$. For $N = 4$, given the concatenated nature of the encoding choice, each single-qubit gate $\hat{U}_1$ and $\hat{U}_2$ can be replaced by an $N = 2$ circuit $\hat{B}^{(2)}_{1\leftrightarrow 2}(\hat{U}_2 \otimes \hat{U}_1)\hat{B}^{(1)}_{1\leftrightarrow 2}$ and relabeling the elements. This process can then be further repeated for higher $N = 2^n$, where $n \in \mathbb{N}$. However, care is needed when relabeling each single-qubit unitary and beam splitter to ensure each act on the appropriate modes.

The transformation implemented in Eq. (7) after heralding, but before renormalization, will evolve the annihilation operators according to

$$\begin{aligned} &\hat{a}_{\text{out}}(N=2) \\ &\quad = [\sin(\theta_1)\sin(\theta_2)\hat{U}_1 + \cos(\theta_1)\cos(\theta_2)\hat{U}_2]\hat{a}_{\text{in}} \\ &\qquad + [\sin(\theta_1')\sin(\theta_2')\hat{U}_1 + \cos(\theta_1')\cos(\theta_2')\hat{U}_2]\hat{b}_{\text{in}} \\ &\quad = [\hat{U}_T + \tfrac{1}{2}(\hat{E}_1 + \hat{E}_2)](\hat{a}_{\text{in}} + \hat{b}_{\text{in}}) + O(\hat{E}^2,\ \delta\theta\hat{E},\ \delta\theta^2), \end{aligned} \quad (9)$$

where we have Taylor expanded the applied single-qubit gates, setting $\hat{U}_j = \hat{U}_T + \hat{E}_j$ (as discussed in more detail later). All terms linear in $\delta\theta_i$ naturally cancel. An equivalent expression can be written for $\hat{b}_{\text{out}}$. For the details of this calculation see Appendix A.

The $N = 4$ result can be calculated by replacing each single-qubit unitary $\hat{U}_i$ with the entire $N = 2$ result, with the appropriate relabeling of the parameters giving

$$\hat{a}_{\text{out}}(N=4) \approx \left(\hat{U}_T + \tfrac{1}{4}\sum_{j=1}^{4}\hat{E}_j\right)(\hat{a}_{\text{in}} + \hat{b}_{\text{in}}), \quad (10)$$

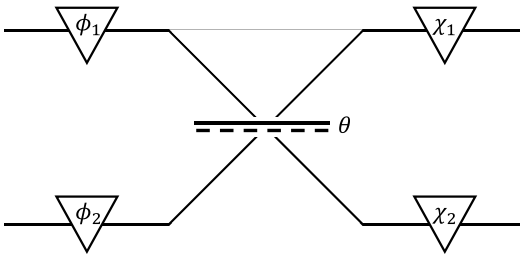


FIG. 2. Arbitrary dual-rail single-qubit unitary.

where again the linear encoding error terms cancel one another (see Appendix A for details). One can see by induction that this pattern continues for all $N = 2^n$, where $n \in \mathbb{Z}$. Thus encoding errors are naturally suppressed to the first order. As such, it is sufficient to account for only errors within the averaged unitaries themselves, so encoding errors will not be considered for the remainder of this paper.

One might be tempted to view this suppression of the encoding errors as a result of the concatenation; however, given the linear encoding errors for the $N = 2$ case [as shown in Eq. (9)] this cannot be the full story. While the concatenation will help further suppress noise, it is the postselection that removes the last two terms from Eq. (9).

IV. UNITARY-AVERAGED SINGLE-QUBIT GATES

Here we consider how an arbitrary single-qubit physical gate might be protected using UA. Specifically, we detail how phase- and bit-flip errors are converted into heralded loss allowing for a trade-off between fidelity and known loss. Given the results of the preceding section, it is not necessary to include any encoding errors and so we will consider noise only in the linear optical unitary.

The first thing to do is build an error model for an unencoded gate and then consider the impact of UA on these errors. To do this we use an overcomplete gate description so as to allow errors to arise anywhere within the optical circuit which implements the unitary. The circuit depth is also then the same regardless of the path taken. The chosen optical layout allows an arbitrary single-qubit gate to be implemented, with specific parameters tuned such that it implements the intended transformation. This is shown in Fig. 2. This gate $\hat{U}_j$ will implement the transformation

$$\hat{U}_j = \begin{bmatrix} e^{i\phi_{1,j}}e^{i\chi_{1,j}}\sin(\theta_j) & e^{i\phi_{2,j}}e^{i\chi_{1,j}}\cos(\theta_j) \\ e^{i\phi_{1,j}}e^{i\chi_{2,j}}\cos(\theta_j) & -e^{i\phi_{2,j}}e^{i\chi_{2,j}}\sin(\theta_j) \end{bmatrix}. \quad (11)$$

We can then take each parameter $\mathcal{O}_j$ to be given by the intended target value $\mathcal{O}$ and an additional noise term δO_j, where $\mathcal{O}_j = \mathcal{O} + \delta\mathcal{O}_j$. The appropriate parameters for a number of gates of interest are shown in Table I. We can then Taylor expand each parameter around its target value, as shown in Appendix B. Doing so allows us to write the applied unitary as $\hat{U}_j = \hat{U}_T + \hat{E}_j$, where $\hat{U}_T$ is the target unitary. Thus, after employing UA, we will have implemented the transformation

TABLE I. Arbitrary single-qubit gate parameters.

Gate	θ	ϕ_1	ϕ_2	χ_1	χ_2
$\hat{I}$	$\frac{\pi}{2}$	0	0	0	π
$\hat{X}$	0	0	0	0	0
$\hat{Y}$	0	$\frac{\pi}{2}$	0	$-\frac{\pi}{2}$	0
$\hat{Z}_\alpha$	$\frac{\pi}{2}$	0	0	0	α
$\hat{H}$	$\frac{\pi}{4}$	0	0	0	0

on the output mode

$$\hat{\mathcal{U}}(N) = \frac{1}{N}\sum_{j=1}^{N}(\hat{U}_T + \hat{E}_j) \equiv \hat{U}_T + \frac{\hat{\epsilon}(N)}{N}, \tag{12}$$

where $\hat{\epsilon}(N)$ is some stochastically varying transformation. Note that again we are not tracking what happens to the error modes as a success requires these to be heralded in the vacuum state. Thus, so long as each error is unbiased and independent, we might expect any desired gate fidelity to be achievable with sufficiently large N. Specifically, with $\lim_{N\to\infty}\hat{\mathcal{U}}(N) = \hat{U}_T$, arbitrarily high gate fidelity may be possible with sufficient redundancy. This will be at the cost of a decreasing probability of success $P_s(N)$. We can write the state at the output modes in general as

$$\begin{aligned}\hat{\rho}(N) &= \hat{\mathcal{U}}(N)|\psi\rangle\langle\psi|\hat{\mathcal{U}}^\dagger(N)\\ &= \left(\hat{U}_T + \frac{\hat{\epsilon}(N)}{N}\right)|\psi\rangle\langle\psi|\left(\hat{U}_T + \frac{\hat{\epsilon}^\dagger(N)}{N}\right)\\ &= |\Psi\rangle\langle\Psi| + \frac{1}{N}[\hat{\epsilon}(N)|\psi\rangle\langle\psi|\hat{U}_T + \hat{U}_T|\psi\rangle\langle\psi|\hat{\epsilon}^\dagger(N)]\\ &\quad + \frac{1}{N^2}\hat{\epsilon}(N)|\psi\rangle\langle\psi|\hat{\epsilon}^\dagger(N)\end{aligned} \tag{13}$$

and the density operator after postselection as

$$\hat{\rho}_{\rm ps} = [P_s(N)]^{-1}\hat{\rho}(N), \tag{14}$$

with the associated probability of success as defined in Eq. (4).

The probability of success is determined by the moments of the noise distributions, at least under the assumptions that each noise term is independent. Throughout this paper we set each noise parameter to be independent and identically distributed, with a Gaussian noise profile. Thus, each parameter's noise $\delta\mathcal{O}_j$ has the following properties: $\langle\delta\mathcal{O}_j\rangle = 0$, $\langle\delta\mathcal{O}_j\delta\mathcal{O}_k\rangle = \nu\delta_{j,k}$, $\langle\delta\mathcal{O}_j^3\rangle = 0$, $\langle\delta\mathcal{O}_j^4\rangle = 3\nu^2$, and $\langle\delta\mathcal{O}_j\delta\mathcal{O}_k\rangle = \langle\delta\mathcal{O}_j\rangle\langle\delta\mathcal{O}_k\rangle\ \forall\, j\neq k$. The probability of success for a transformation characterized by Eq. (11) will be

$$P_s(N) \approx 1 - 3\nu + \frac{3\nu}{N} + \frac{9\nu^2}{2} - \frac{9\nu^2}{2N}, \tag{15}$$

as shown in Fig. 3. For the details of this calculation up to first order in ν, see Appendix B 2; the higher-order calculation was performed with the aid of *Mathematica*. While we have taken the noise to be Gaussian to simplify the result, this could be replaced by other unbiased probability distributions without significantly impacting the results. The postselected gate fidelity can be calculated similarly

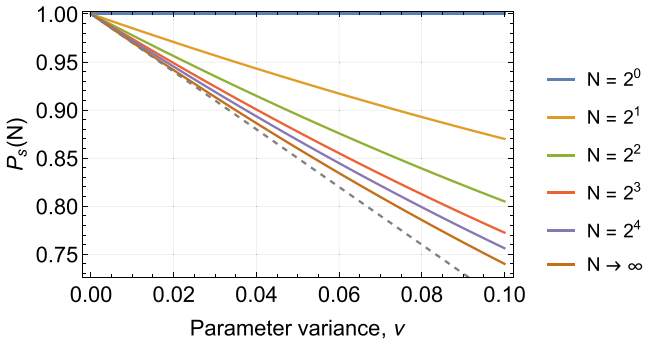


FIG. 3. Probability of success scaling for a unitary-averaged single-qubit gate as given by Eq. (15). The gray dashed line shows the first-order scaling of $\lim_{N\to\infty} P_s(N) = 1 - 3\nu$.

using

$$\begin{aligned}\mathcal{F}(N) &= \langle\Psi|\rho_{\rm ps}(N)|\Psi\rangle\\ &= [P_s(N)]^{-1}\langle\psi|\hat{U}_T^\dagger\hat{\mathcal{U}}(N)|\psi\rangle\langle\psi|\hat{\mathcal{U}}^\dagger(N)\hat{U}_T|\psi\rangle\\ &= [P_s(N)]^{-1}\left(1 + \frac{1}{N}\sum_{j=1}^{N}\langle\psi|\hat{U}_T^\dagger\hat{E}_j|\psi\rangle\right)^2.\end{aligned} \tag{16}$$

Again, the fidelity is characterized by the moments of the noise probability distribution with

$$\mathcal{F}(N) \approx \frac{1 - 3\nu + \frac{9}{4\nu^2}}{1 - 3\nu + \frac{\nu}{N} + \frac{9\nu^2}{2} - \frac{9\nu^2}{2N}} \tag{17}$$

to the second order in the parameter variance ν. The details of the calculation are provided in Appendix B 3. The fidelity scaling is shown in Fig. 4.

V. TWO-QUBIT GATES

To be sufficient for arbitrary quantum computation, it is necessary to have a corrected universal gate set. While we have just shown it is possible to protect any single-qubit gate, it is still necessary to provide for two-qubit gates such as the controlled-NOT gate. Two-qubit gates under parity encoding (see Sec. VII) can be enacted via single-qubit gates and reencoding using type-I and type-II fusion gates [12]. As such, to protect against errors in two-qubit gates using unitary

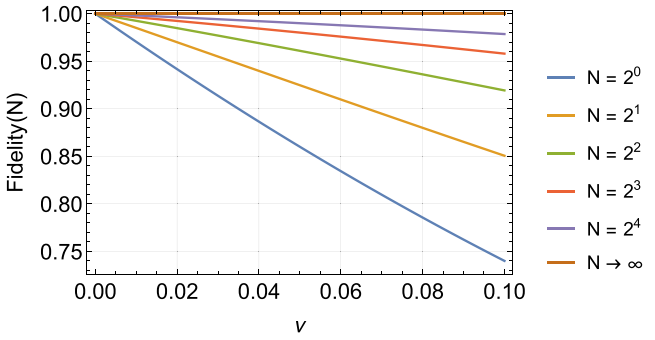


FIG. 4. Postselected fidelity for a unitary-averaged single-qubit gate as given by Eq. (17).

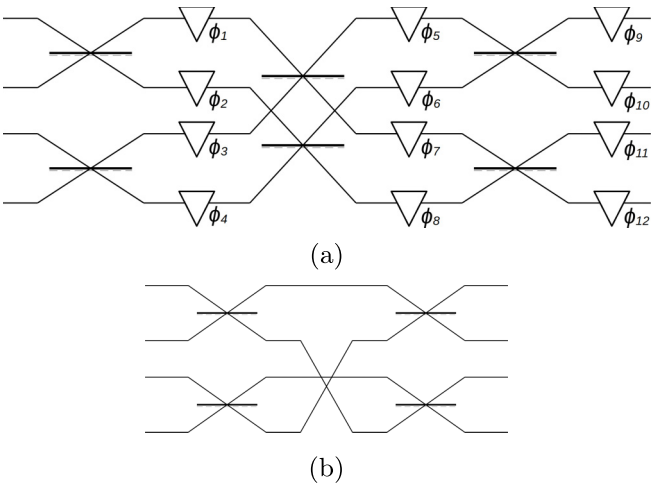


FIG. 5. (a) General four-mode linear optical gate for two qubits consisting of beam splitters between pairs of modes and phase shifters marked by their phase ϕ_i. The type-II fusion can be implemented with a reduced version of this as shown in (b) with all present beam splitters set to 50:50.

averaging, we need to consider how to protect fusion gates. Recent work has shown that UA can be employed to correct fusion gates [7]; however, this study was considering the low-N limit. A general dual-rail two-qubit linear optical gate can be implemented by the interferometer diagram shown in Fig. 5. The fusion gates can be implemented by a simplified version of this system, with additional measurements on the outputs which are taken to occur separately from the UA protocol.

The same procedure for determining the probability of success for the one-photon gate can be employed, only, due to the increased problem size, *Mathematica* was used to simplify the expressions. It was found that the four-mode interferometer shown in Fig. 5, when employing UA, will implement the target gate with a postselected fidelity of

$$\mathcal{F}_{\text{4-mode}}(N) = P_{s,\text{4-mode}}(N)^{-1}(1-6\nu) \tag{18}$$

and a success probability

$$P_{s,\text{4-mode}}(N) = 1 - 6\nu + \frac{6\nu}{N} + 18\nu^2 - \frac{18\nu^2}{N^2}. \tag{19}$$

These can be further improved in systems that do not require the full generality enabled by such an interferometer. Specifically, when implementing a type-II fusion gate, no phase manipulation is required and the central column of beam splitters can be replaced with a single physical swapping of the second and fourth modes. With this reduced system, the postselected fidelity becomes

$$\mathcal{F}_{\text{II}}(N) = P_{s,\text{II}}(N)^{-1}(1-2\nu+2\nu^2), \tag{20}$$

with a success probability of

$$P_{s,\text{II}}(N) = 1 - 2\nu + \frac{2\nu}{N} + 2\nu^2 - \frac{2\nu^2}{N}. \tag{21}$$

It appears that the gate fidelity and probability of success scale linearly with the optical depth of the circuit, as seen by comparing Eqs. (17), (18), and (20) and Eqs. (15), (19), and (21). We can in general write the probability of success and fidelity to the first order in the variance as

$$P_{s,g}(N) = 1 - \mathcal{V} + \frac{\mathcal{V}}{N}, \tag{22}$$

$$\begin{aligned}\mathcal{F}_g(N) &= P_{s,g}(N)^{-1}(1-\mathcal{V}) \\ &= 1 - \frac{\mathcal{V}}{N+\mathcal{V}-N\mathcal{V}},\end{aligned} \tag{23}$$

where $\mathcal{V} = d \times \nu$ is the characteristic noise parameter for a circuit with an optical depth d.

With this we can achieve arbitrarily high gate fidelities with sufficiently large N and heralded loss. This process is considered heralded loss rather than postselection as it does not require any postprocessing and any corrective action can be taken straight away, potentially correcting for such loss. An obvious hypothetical use case is then employing effectively infinite averaging ($N \to \infty$) along with loss tolerance to achieve a regime of effective fault tolerance, by which we mean that, within the bounds of the error model considered here, each gate implements a logical transformation with perfect fidelity and so errors do not accumulate. However, perfect fidelity only truly occurs for infinite averaging ($N = \infty$) and so may not reasonably constitute true fault tolerance. Nonetheless, the next section presents this process, with the behavior of UA within standard fault-tolerant protocols presented in Sec. VIII.

VI. PARITY ENCODING

Parity encoding (PE) provides a loss-tolerant encoding in which extra physical qubits can be added throughout the computation in response to the occurrence of loss [4,13]. Given that UA turns logical errors into heralded loss and PE corrects for photon loss, here we seek to combine the two to produce a general error-correction scheme.

Full parity encoding employs two separate encoding steps: parity type and redundancy type. Parity type encodes logical states as

$$\begin{aligned}|0\rangle_L &= |0\rangle^{(n)} = (|+\rangle^{\otimes n} + |-\rangle^{\otimes n})/\sqrt{2}, \\ |1\rangle_L &= |1\rangle^{(n)} = (|+\rangle^{\otimes n} - |-\rangle^{\otimes n})/\sqrt{2},\end{aligned} \tag{24}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|H\rangle \pm |V\rangle)$ as we consider polarization-encoded qubits in these sections. The redundancy-type encoding uses q copies of each of these parity-encoded states. With this, an arbitrary single-qubit state is encoded as

$$|\Psi\rangle_L = \alpha \bigotimes^{q} |0\rangle^{(n)} + \beta \bigotimes^{q} |1\rangle^{(n)}. \tag{25}$$

Parity encoding in optical quantum computing has been shown to contain a universal gate set [13] with $\hat{X}$ and $\hat{Z}$ rotations implemented simply by applying the gate to a single physical qubit or all physical qubits, respectively, while other gates require reencoding.

VII. PARITY-ENCODED AND UNITARY-AVERAGED SINGLE-QUBIT GATES

In this section we explore how parity encoding can be utilized in conjunction with unitary averaging to implement high-fidelity logical operations with a high probability of success using noisy unitary operations. To begin, we present the case in which J errors occur, but only within a single redundant encoded state. Consider the initial state

$$|\psi\rangle_L = \alpha|0\rangle^{(n)} \bigotimes^{q-1} |0\rangle^{(n)} + \beta|1\rangle^{(n)} \bigotimes^{q-1} |1\rangle^{(n)} \tag{26}$$

on which we act a single logical qubit unitary $\hat{U} = \bigotimes_{i=1}^{qn} \hat{u}_i$ which has some effect on each individual physical qubit. Note that here we single out a single redundant copy of the parity encoding which we consider to be the location of the $J < n$ errors. We take the individual unitaries to be unitary averaged with $N \to \infty$ and the error ports monitored such that the resulting transformation is projected onto either the correct target result or a heralded loss. When a heralded loss occurs a unique stochastic phase factor is then written on to the physical qubit. Thus, unitary averaging acts to transform each physical qubit according to

$$|H\rangle_i \to \hat{u}_T|H\rangle_i + \sum_{k=2}^{N} \delta H_{i,k}|\epsilon_k\rangle_i, \tag{27}$$

$$|V\rangle_i \to \hat{u}_T|V\rangle_i - \sum_{k=2}^{N} \delta V_{i,k}|\epsilon_k\rangle_i, \tag{28}$$

where we have neglected normalization, as after projecting on either the error or no-error result, it will not be important. Here each $\delta H_{i,k}$ and $\delta V_{i,k}$ is a unique stochastic phase factor and $|\epsilon_k\rangle_i$ represents the photon in the kth error mode which is to be measured, heralding a photon loss.

Any parity-encoded state can be expanded, using the identity

$$|0\rangle^{(l+m)} = \frac{1}{\sqrt{2}}(|0\rangle^{(l)}|0\rangle^{(m)} + |1\rangle^{(l)}|1\rangle^{(m)}), \tag{29}$$

$$|1\rangle^{(l+m)} = \frac{1}{\sqrt{2}}(|0\rangle^{(l)}|1\rangle^{(m)} + |1\rangle^{(l)}|0\rangle^{(m)}), \tag{30}$$

which, in the instance of J errors occurring in only the first redundant encoding, allows us to write

$$\begin{aligned}
\langle\boldsymbol{\epsilon}|_\epsilon^{1\to J}\langle\mathbf{0}|_\epsilon^{J+1\to n}\hat{U}|\psi\rangle_L =& \frac{\alpha}{\sqrt{2}}\left(\bigotimes^{k}\hat{u}_T\right)(\delta\Theta|0\rangle^{(k)} + \delta\Phi|1\rangle^{(k)}) \bigotimes^{q-1}\left[\left(\bigotimes^{n}\hat{u}_T\right)|0\rangle^{(n)}\right] \\
&+ \frac{\beta}{\sqrt{2}}\left(\bigotimes^{k}\hat{u}_T\right)(\delta\Theta|1\rangle^{(k)} + \delta\Phi|0\rangle^{(k)}) \bigotimes^{q-1}\left[\left(\bigotimes^{n}\hat{u}_T\right)|1\rangle^{(n)}\right] \\
=& \frac{\alpha}{2}\left(\bigotimes^{k}\hat{u}_T\right)[(\delta\Theta + \delta\Phi)|+\rangle^{\otimes k} + (\delta\Theta - \delta\Phi)|-\rangle^{\otimes k}] \bigotimes^{q-1}\left[\left(\bigotimes^{n}\hat{u}_T\right)|0\rangle^{(n)}\right] \\
&+ \frac{\beta}{2}\left(\bigotimes^{k}\hat{u}_T\right)[(\delta\Theta + \delta\Phi)|+\rangle^{\otimes k} - (\delta\Theta - \delta\Phi)|-\rangle^{\otimes k}] \bigotimes^{q-1}\left[\left(\bigotimes^{n}\hat{u}_T\right)|1\rangle^{(n)}\right],
\end{aligned} \tag{31}$$

where $J + k = n$, $|\mathbf{0}\rangle_\epsilon^{a\to b}$ corresponds to error modes a through to b in the vacuum (no errors), and $|\epsilon\rangle_\epsilon^{a\to b}$ corresponds to an error occurring in modes a through to b. Also,

$$\delta\Theta = \langle\boldsymbol{\epsilon}|_\epsilon^{1\to J}\left(\bigotimes_{i=1}^{J}\hat{\mathcal{U}}_{i,e}\right)|0\rangle^{(J)}, \tag{32}$$

$$\delta\Phi = \langle\boldsymbol{\epsilon}|_\epsilon^{1\to J}\left(\bigotimes_{i=1}^{J}\hat{\mathcal{U}}_{i,e}\right)|1\rangle^{(J)} \tag{33}$$

are stochastic $\mathbb{C}$ numbers whose value depends on the specific location of the errors and the individual operations implemented on the qubits. After renormalization, these factors randomly take the values ± 1.

If any one of the remaining k physical qubits of the first redundant encoding is measured in the $\hat{u}_T|\pm\rangle^{(1)} = \frac{1}{\sqrt{2}}\hat{u}_T(|0\rangle^{(1)} \pm |1\rangle^{(1)})$ basis, the state gains an undetectable global phase factor which depends on the measurement result, with $(\delta\Theta + \delta\Phi)$ for $|+\rangle^{(1)}$ and $(\delta\Theta - \delta\Phi)$ for $|-\rangle^{(1)}$. Also, if $|-\rangle^{(1)}$ is returned there will also be a known sign error on the logical $|1\rangle_L$ state. Given only relative phases are important, the global phase can be ignored. Dropping the global phase and conducting projective measurements on all error channels produces a final state after renormalization of

$$\begin{aligned}
|\psi_{\text{out}}(\pm)\rangle_L =& \left(\bigotimes^{k-1}\hat{u}_T|\pm\rangle\right) \\
&\otimes \hat{U}_T\left(\alpha\bigotimes^{q-1}|0\rangle^{(n)} \pm \beta\bigotimes^{q-1}|1\rangle^{(n)}\right).
\end{aligned} \tag{34}$$

The component of the state with an error present is unentangled from the remainder of the state which successfully has the target unitary applied, although potentially with a known phase error. This can be repeated for any of the remaining $q-1$ redundant copies in the parity encoding in which an error is detected. Provided at least one entire redundant copy is heralded as error free, the target unitary will be successfully applied. The phase error occurs if an odd number of $|-\rangle$ states are returned during the projective measurements. We therefore have the following set of success criteria.

(i) At most $q-1$ redundant copies of the encoding are heralded to have encountered an error. This ensures that there

remains a logical state on which the gate is successfully applied.

(ii) For all redundant copies of the encoding which are heralded to have encountered an error, at least one physical qubit must not have been heralded to be in error so that the projected measurement can be performed.

VIII. UNITARY AVERAGING AND STANDARD FAULT TOLERANCE

Now that we have established the action of UA on single- and two-qubit gates and considered how it can be integrated into existing loss-correction schemes, it is insightful to translate these results into fault-tolerant error-correction schemes. The purpose here is to explore the potential benefits of employing UA within existing error-correction schemes on the physical qubit level. The aim is to expand the parameter space for which fault tolerance might be achievable by exchanging loss and noise. The analysis presented here is intended as a guide to how UA might impact the fault-tolerant parameter space and not a full analysis, which is beyond the scope of the present study. However, given that certain implementations maybe be capable of achieving low loss rates but higher than tolerable error rates, we seek to show that these schemes may be made compatible with certain fault-tolerant schemes through the use of UA. Furthermore, when the gate-error rate is at or very close to the fault-tolerant boundary, a significant overhead is required to achieve reasonable error rates. To this end, we consider how the fault-tolerant parameter space changes if the operations performed on each physical qubit are done utilizing UA. We concern ourselves with the per gate depolarization probability ε and per qubit, per gate loss rate γ, which with UA become

$$\Gamma = \frac{\gamma}{3}[3 + 2\log_2(N)] + \varepsilon\left(1 - \frac{1}{N}\right), \tag{35}$$

$$\mathcal{E} = \frac{\varepsilon}{N + \varepsilon - N\varepsilon}. \tag{36}$$

See Appendix D for more details.

In the same manner as discussed above, we can consider UA to sit within existing algorithms and serve to reduce the effective gate error at the cost of increasing loss. We use the above results in conjunction with existing fault-tolerant thresholds for some example error-correction schemes. If we take the earlier results of Dawson *et al.* [10] and Fujii and Tokunaga [11], we can estimate the effect UA has on the fault-tolerant parameter space for both the more modern fault-tolerant architectures such as fusion-based quantum computation (FBQC) [3] and the more simple seven-qubit Steane code. The results are shown in Fig. 6.

This suggests a potential improvement in the parameter space at which fault tolerance can be achieved despite the increase in effective loss due to a lossy encoding circuit and the probabilistic nature of unitary averaging. This may enable fault tolerance in systems with low inherent loss but higher error rates. This improvement primarily arises due to the difference in scaling between the effective loss rate Γ and the gate-error rate $\mathcal{E}$, where for large N we see that loss scales logarithmically while the error scales proportionally to $1/N$. We also see in Fig. 6(a) that this benefit is only observed for

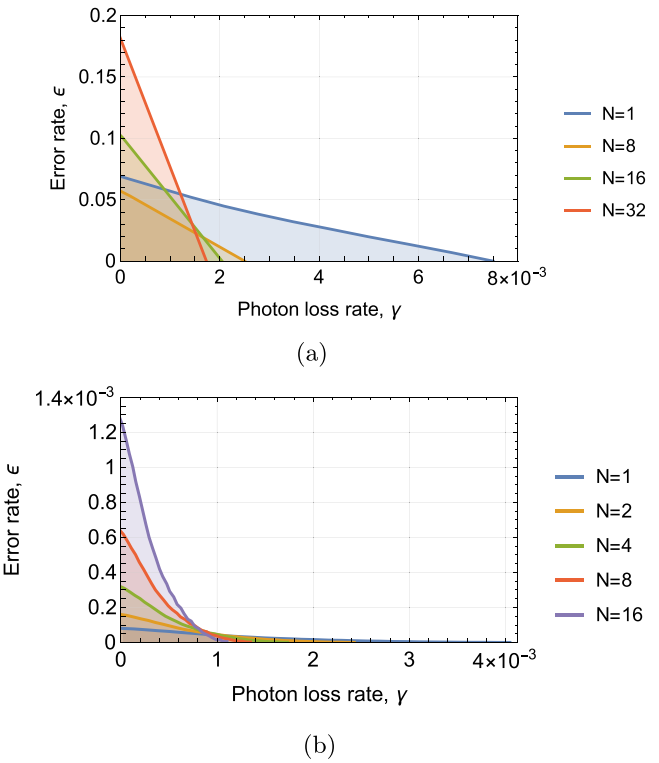


FIG. 6. Fault-tolerant parameter space improvement estimation when employing unitary averaging in (a) a FBQC utilizing a four-star architecture and (b) a seven-qubit Steane code. The $N = 1$ data were taken from [3,10] for the FBQC and the Steane code, respectively. Fault tolerance is achieved within the shaded region for each value of N. Note that, depending on the original ($N = 1$) parameter space, applying UA could be either beneficial or detrimental to the parameter space.

sufficiently large N; when $N \leqslant 8$, the additional induced loss eliminates any benefit of improved tolerable error rates. This suggests that some error-correction codes require a larger N to obtain any benefit from applying UA.

It is worth highlighting that the results in Fig. 6 are dependent on some significant assumptions, specifically that each component that goes into implementing these encodings and optical circuits can be implemented using UA and that UA has the same effect on their output as seen above. For example, implementing a measurement-based quantum computer will likely require many components not considered above such as quantum memories and detectors. It seems reasonable that memories will be compatible with UA. However, destructive measurements clearly cannot be treated as unitaries that can be averaged; thus these results will likely apply only to the preparation stages of such a device. This also captures only the first-order effect of UA, where encoding errors cancel. We are here treating all loss as due to the photon absorption situation and so are overestimating the additional loss induced with the increased optical depth of UA. This has the effect of underestimating the tolerable loss to maintain fault tolerance. We are also neglecting all effects due to the larger level of encoding, such as longer range gates. As such, these results need to be viewed as approximate scaling.

IX. CONCLUSION

This paper has explored the use of UA within the context of applying arbitrary one- and two-qubit gate transformations. We have shown that UA can be implemented in such a way as to avoid encoding errors while also introducing only a logarithmic number of optical elements to the path depth. Unitary averaging may enable arbitrary high gate fidelity at the cost of a reduced probability of success. This cost, however, is linear in the initial parameter variance within a single unitary. We have also demonstrated that this loss can be mitigated by employing existing techniques for loss protection which will already be necessary in any large-scale system. Furthermore, we have suggested how UA can be used within existing fault-tolerant schemes to modify and potentially improve the parameter space for which error tolerance is achieved. We have done so considering both early fault-tolerant schemes, which can greatly benefit from the use of UA, and more modern surface-code-based schemes for which the benefit of applying UA appears to be more modest.

ACKNOWLEDGMENTS

A.P.L. acknowledges support from BMBF (QPIC) and the Einstein Research Unit on Quantum Devices. This research was supported by the Australian Research Council through the Centre of Excellence for Quantum Computation and Communication Technology (Project No. CE170100012).

APPENDIX A: ENCODING-ERROR CALCULATION

The following gives a detailed analysis of encoded-error propagation through a UA system. Starting with $N = 2$, the output of a single-qubit-averaged gate after postselection is given by

$$|\psi\rangle_{\text{out}} = \langle \mathbf{0}|_\epsilon \hat{B}^{(2)}_{1\leftrightarrow 2}(\hat{U}_2 \otimes \hat{U}_1)\hat{B}^{(1)}_{1\leftrightarrow 2}|\psi\rangle_{\text{in}}, \tag{A1}$$

where $\hat{B}^{(j)}_{a\leftrightarrow b}$ acts to evenly mix the pair of modes a and b, each of which is then separately acted on by the independent unitaries $\hat{U}_a$ and $\hat{U}_b$ respectively, and $\langle \mathbf{0}|_\epsilon$ represents the projection onto the vacuum for the error modes. The superscript j serves as a reminder that each beam splitter is independent and thus has its own unique noise associated with it. In a dual rail, this can be written as

$$\hat{B}^{(j)}_{1\leftrightarrow 2} = \begin{bmatrix} \sin(\theta_j) & 0 & \cos(\theta_j) & 0 \\ 0 & \sin(\theta'_j) & 0 & \cos(\theta'_j) \\ \cos(\theta_j) & 0 & -\sin(\theta_j) & 0 \\ 0 & \cos(\theta'_j) & 0 & -\sin(\theta'_j) \end{bmatrix}, \tag{A2}$$

where $\theta^{(\prime)}_j = \frac{\pi}{2} + \delta\theta^{(\prime)}_j$ for $\delta\theta^{(\prime)}_j \ll \frac{\pi}{2}$, allowing us to write $\sin(\theta_j) \approx \frac{1}{\sqrt{2}} + \delta\theta^{(\prime)}_j$ and $\cos(\theta_j) \approx \frac{1}{\sqrt{2}} - \delta\theta^{(\prime)}_j$. For $N = 4$, given the concatenated nature of the encoding choice, each single-qubit gate $\hat{U}_1$ and $\hat{U}_2$ can be replaced by an $N = 2$ circuit $\hat{B}^{(2)}_{1\leftrightarrow 2}(\hat{U}_2 \otimes \hat{U}_1)\hat{B}^{(1)}_{1\leftrightarrow 2}$ and relabeling the elements. This process can then be further repeated for higher $N = 2^j$, where $j \in \mathbb{N}$. Care is needed when relabeling each single-qubit unitary and beam splitter to ensure each acts on the appropriate modes. We will also assume that the implemented transformations $\hat{U}_j$ are in some sense close to the target transformation $\hat{U}_T$ such that we can write $\hat{U}_j \approx \hat{U}_T + \hat{E}_j$, where $\hat{E}_j$ is the error in $\hat{U}_j$ and $\hat{E}^2_j$ is negligible.

The transformation implemented in Eq. (A1) after heralding but before renormalization on no error detection gives the following mode operator transformation:

$$\begin{aligned}
\hat{a}_{\text{out}}(N=2) =& [\sin(\theta_1)\sin(\theta_2)\hat{U}_1 + \cos(\theta_1)\cos(\theta_2)\hat{U}_2]\hat{a}_{\text{in}} + [\sin(\theta'_1)\sin(\theta'_2)\hat{U}_1 + \cos(\theta'_1)\cos(\theta'_2)\hat{U}_2]\hat{b}_{\text{in}} \\
=& \tfrac{1}{2}[(1+\delta\theta_1)(1+\delta\theta_2)(\hat{U}_T + \hat{E}_1) + (1-\delta\theta_1)(1-\delta\theta_2)(\hat{U}_T + \hat{E}_2)]\hat{a}_{\text{in}} \\
&+ \tfrac{1}{2}[(1+\delta\theta'_1)(1+\delta\theta'_2)(\hat{U}_T + \hat{E}_1) + (1-\delta\theta'_1)(1-\delta\theta'_2)(\hat{U}_T + \hat{E}_2)]\hat{b}_{\text{in}} + O(\delta\theta^2, \hat{E}^2) \\
=& [\hat{U}_T + \tfrac{1}{2}(\hat{E}_1 + \hat{E}_2)](\hat{a}_{\text{in}} + \hat{b}_{\text{in}}) + \tfrac{1}{2}(\hat{E}_1 - \hat{E}_2)[(\delta\theta_1 + \delta\theta_2)\hat{a}_{\text{in}} + (\delta\theta'_1 + \delta\theta'_2)\hat{b}_{\text{in}}] \\
=& [\hat{U}_T + \tfrac{1}{2}(\hat{E}_1 + \hat{E}_2)](\hat{a}_{\text{in}} + \hat{b}_{\text{in}}) + O(\hat{E}^2, \delta\theta\hat{E}, \delta\theta^2).
\end{aligned} \tag{A3}$$

Note that all terms linear in $\delta\theta_j$ cancel. An equivalent expression can be written for $\hat{b}_{\text{out}}$. Similarly, the $N = 4$ result can be calculated by replacing each single-qubit unitary $\hat{U}_i$ with the entire $N = 2$ result, with the appropriate relabeling of the parameters. For compactness, only the $\hat{a}_{\text{in}}$ terms are expanded to improve legibility, giving

$$\begin{aligned}
\hat{a}_{\text{out}}(N=4) =& (\sin(\theta_1)\sin(\theta_2)\{[\sin(\theta_3)\sin(\theta_4)\hat{U}_1 + \cos(\theta_3)\cos(\theta_4)\hat{U}_2]\} + \cos(\theta_1)\cos(\theta_2)\{[\sin(\theta_5)\sin(\theta_6)\hat{U}_3 \\
&+ \cos(\theta_5)\cos(\theta_6)\hat{U}_4]\})\hat{a}_{\text{in}}(\sin(\theta'_1)\sin(\theta'_2)\{[\sin(\theta'_3)\sin(\theta'_4)\hat{U}_1 + \cos(\theta'_3)\cos(\theta'_4)\hat{U}_2]\} \\
&+ \cos(\theta'_1)\cos(\theta'_2)\{[\sin(\theta'_5)\sin(\theta'_6)\hat{U}_3 + \cos(\theta'_5)\cos(\theta'_6)\hat{U}_4]\})\hat{b}_{\text{in}} \\
\approx& \tfrac{1}{4}((1+\delta\theta_1)(1+\delta\theta_2)\{[(1+\delta\theta_3)(1+\delta\theta_4)(\hat{U}_T + \hat{E}_1) + (1-\delta\theta_3)(1-\delta\theta_4)(\hat{U}_T + \hat{E}_2)]\} \\
&+ (1-\delta\theta_1)(1-\delta\theta_2)\{[(1+\delta\theta_5)(1+\delta\theta_6)(\hat{U}_T + \hat{E}_3) + (1-\delta\theta_5)(1-\delta\theta_6)(\hat{U}_T + \hat{E}_4)]\})\hat{a}_{\text{in}} \\
&+ (\cdots)\hat{b}_{\text{in}} + O(\delta\theta^2, \hat{E}^2) \\
=& \tfrac{1}{4}[(1+\delta\theta_1+\delta\theta_2+\delta\theta_3+\delta\theta_4)(\hat{U}_T + \hat{E}_1) + (1+\delta\theta_1+\delta\theta_2-\delta\theta_3-\delta\theta_4)(\hat{U}_T + \hat{E}_2)
\end{aligned}$$

$$
\begin{aligned}
&+ (1 - \delta\theta_1 - \delta\theta_2 + \delta\theta_5 + \delta\theta_6)(\hat{U}_T + \hat{E}_3) + (1 - \delta\theta_1 - \delta\theta_2 - \delta\theta_5 - \delta\theta_6)(\hat{U}_T + \hat{E}_4)]\hat{a}_{\text{in}} \\
&+ (\cdots)\hat{b}_{\text{in}} + O(\delta\theta^2, \hat{E}^2) \\
&= \left(\hat{U}_T + \tfrac{1}{4}\sum_{j=1}^{4}\hat{E}_j\right)(\hat{a}_{\text{in}} + \hat{b}_{\text{in}}) + O(\delta\theta^2, \delta\theta\hat{E}, \hat{E}^2),
\end{aligned} \tag{A4}
$$

where again the linear encoding-error terms cancel one another. Proceeding similarly, we can write

$$
\hat{a}_{\text{out}}(N = 2^n) \approx \left(\hat{U}_T + \frac{1}{N}\sum_{j=1}^{N}\hat{E}_j\right)(\hat{a}_{\text{in}} + \hat{b}_{\text{in}}). \tag{A5}
$$

APPENDIX B: SINGLE-QUBIT GATE-ERROR CALCULATION

Here we present the details of so solving for the expectation values used to calculate the probability of success when error averaging a single-qubit gate. Throughout we will assume that each noisy parameter is independent and identically distributed, with a Gaussian noise profile. Thus we can take $\langle\delta\mathcal{O}\rangle = 0$, $\langle\delta\mathcal{O}_j\delta\mathcal{O}_k\rangle = \nu\delta_{j,k}$, $\langle\delta\mathcal{O}^3\rangle = 0$, $\langle\delta\mathcal{O}^4\rangle = 3\nu^2$ for all noise terms $\delta\mathcal{O}$.

1. Single-qubit gate error

Here we consider each physical gate to be averaged $\hat{U}_j$. Taylor expanding each parameter and keeping only the terms that are bilinear or quadratic in the error terms gives

$$
\hat{U}_j = \begin{bmatrix} e^{i\phi_{1,j}}e^{i\chi_{1,j}}\sin(\theta_j) & e^{i\phi_{2,j}}e^{i\chi_{1,j}}\cos(\theta_j) \\ e^{i\phi_{1,j}}e^{i\chi_{2,j}}\cos(\theta_j) & -e^{i\phi_{2,j}}e^{i\chi_{2,j}}\sin(\theta_j) \end{bmatrix} = \begin{bmatrix} a_j & b_j \\ c_j & d_j \end{bmatrix}, \tag{B1}
$$

where

$$
\begin{aligned}
a_j &= e^{i\phi_1}e^{i\chi_1}\sin(\theta) + \left[e^{i\phi_1}e^{i\chi_1}\sin(\theta)\left(i\delta\phi_{1,j} + i\delta\chi_{1,j} - \delta\phi_{1,j}\delta\chi_1 - \frac{\delta\phi_{1,j}^2}{2} - \frac{\delta\chi_{1,j}^2}{2} - \frac{\delta\theta_j^2}{2}\right)\right. \\
&\quad \left. + e^{i\phi_1}e^{i\chi_1}\cos(\theta)\delta\theta_j(1 + i\delta\phi_{1,j} + i\delta\chi_{1,j})\right] \equiv a + [\delta a_j],
\end{aligned} \tag{B2}
$$

$$
\begin{aligned}
b_j &= e^{i\phi_2}e^{i\chi_1}\cos(\theta) + \left[e^{i\phi_2}e^{i\chi_1}\cos(\theta)\left(i\delta\phi_{2,j} + i\delta\chi_{1,j} - \delta\phi_{2,j}\delta\chi_{1,j} - \frac{\delta\phi_{2,j}^2}{2} - \frac{\delta\chi_{1,j}^2}{2} - \frac{\delta\theta_j^2}{2}\right)\right. \\
&\quad \left. - e^{i\phi_2}e^{i\chi_1}\sin(\theta)\delta\theta_j(1 + i\delta\phi_{2,j} + i\delta\chi_{1,j})\right] \equiv b + [\delta b_j],
\end{aligned} \tag{B3}
$$

$$
\begin{aligned}
c_j &= e^{i\phi_1}e^{i\chi_2}\cos(\theta) + \left[e^{i\phi_1}e^{i\chi_2}\cos(\theta)\left(i\delta\phi_{1,j} + i\delta\chi_{2,j} - \delta\phi_{1,j}\delta\chi_{2,j} - \frac{\delta\phi_{1,j}^2}{2} - \frac{\delta\chi_{2,j}^2}{2} - \frac{\delta\theta_j^2}{2}\right)\right. \\
&\quad \left. - e^{i\phi_1}e^{i\chi_2}\sin(\theta)\delta\theta_j(1 + i\delta\phi_{1,j} + i\delta\chi_{2,j})\right] \equiv c + [\delta c_j],
\end{aligned} \tag{B4}
$$

$$
\begin{aligned}
d_j &= -e^{i\phi_2}e^{i\chi_2}\sin(\theta) + \left[-e^{i\phi_2}e^{i\chi_2}\sin(\theta)\left(i\delta\phi_{2,j} + i\delta\chi_{2,j} - \delta\phi_{2,j}\delta\chi_{2,j} - \frac{\delta\phi_{2,j}^2}{2} - \frac{\delta\chi_{2,j}^2}{2} - \frac{\delta\theta_j^2}{2}\right)\right. \\
&\quad \left. - e^{i\phi_2}e^{i\chi_2}\cos(\theta)\delta\theta_j(1 + i\delta\phi_{2,j} + i\delta\chi_{2,j})\right] \equiv d + [\delta d_j].
\end{aligned} \tag{B5}
$$

Each term within the square brackets represents the noise term unique to each physical copy of the unitary, while the preceding term is the intended value of each matrix element. Thus we can write

$$
\hat{U}_j = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} \delta a_j & \delta b_j \\ \delta c_j & \delta d_j \end{bmatrix} = \hat{U}_T + \hat{E}_j. \tag{B6}
$$

2. Solving for probability of success

To solve for the probability of success $P_s(N)$ we assume each noise term is independent, with $\langle\delta\mathcal{O}\rangle = 0$ and $\langle\delta\mathcal{O}^2\rangle = \nu$ for all terms $\delta\mathcal{O}$, and use the Taylor expanded form of the applied transformations as given above. We also consider the general input

state

$$|\psi\rangle = \begin{bmatrix}\alpha\\ \beta\end{bmatrix}. \tag{B7}$$

The probability of success is then given by

$$\begin{aligned} P_s(N) =& \langle\psi|\hat{\mathcal{U}}^\dagger(N)\hat{\mathcal{U}}(N)|\psi\rangle = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle\psi|\hat{U}_j^\dagger\hat{U}_k|\psi\rangle \\ =& \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}(|\alpha|^2\langle a_j^* a_k + c_j^* c_k\rangle + |\beta|^2\langle b_j^* b_k + d_j^* d_k\rangle + \alpha^*\beta\langle a_j^* b_k + c_j^* d_k\rangle + \alpha\beta^*\langle b_j^* a_k + d_j^* c_k\rangle). \end{aligned} \tag{B8}$$

Going term by term through this gives

$$\begin{aligned} \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle a_j^* a_k\rangle =& \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\Bigg\{\sin^2(\theta)\Bigg[\langle 1 - \frac{1}{2}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j} + \delta\theta_{1,j}^2\big) - \frac{1}{2}\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k} + \delta\theta_{1,k}^2\big)\rangle \\ &+ \Big\langle\frac{1}{4}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\theta_{1,j}^2\big)\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k}^2 + \delta\theta_{1,k}^2\big)\Big\rangle + \delta_{j,k}\big\langle\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\phi_j^2\delta\chi_{1,j}^2\big\rangle\Bigg] \\ &+ \cos^2(\theta)\delta_{j,k}\big\langle\theta_j^2\big\rangle\big\langle 1 + \delta\phi_j^2 + \delta\chi_j^2\big\rangle\Bigg\} \\ =& \sin^2(\theta)\left(1 - 3\nu + \frac{9}{4}\nu^2 + \frac{2\nu+\nu^2}{N}\right) + \cos^2(\theta)\frac{\nu+2\nu^2}{N} = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle d_j^* d_k\rangle, \end{aligned} \tag{B9}$$

$$\frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle b_j^* b_k\rangle = \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle c_j^* c_k\rangle = \cos^2(\theta)\left(1 - 3\nu + \frac{9}{4}\nu^2 + \frac{2\nu+\nu^2}{N}\right) + \sin^2(\theta)\frac{\nu+2\nu^2}{N}, \tag{B10}$$

and

$$\begin{aligned} \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle a_j^* b_k\rangle =& \frac{e^{i(\phi_2-\phi_1)}}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\Bigg\{\sin(\theta)\cos(\theta)\Bigg[\Big\langle 1 - \frac{1}{2}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j} + \delta\theta_{1,j}^2\big) - \frac{1}{2}\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k} + \delta\theta_{1,k}^2\big)\Big\rangle \\ &+ \Big\langle\frac{1}{4}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\theta_{1,j}^2\big)\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k}^2 + \delta\theta_{1,k}^2\big)\Big\rangle + \delta_{j,k}\big\langle\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\phi_j^2\delta\chi_{1,j}^2\big\rangle\Bigg] \\ &- \sin(\theta)\cos(\theta)\delta_{j,k}\big\langle\theta_j^2\big\rangle\big\langle 1 + \delta\phi_j^2 + \delta\chi_j^2\big\rangle\Bigg\} \\ \Rightarrow& \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\alpha^*\beta\langle a_j^* b_k\rangle + \alpha\beta^*\langle a_j b_k^*\rangle = 2|\alpha||\beta|\cos(\gamma)\sin(\theta)\cos(\theta)\left(1 - 3\nu + \frac{9}{4}\nu^2 + \frac{\nu-\nu^2}{N}\right), \end{aligned} \tag{B11}$$

where $\gamma = \phi_1 - \phi_2 - \theta_\alpha + \theta_\beta$ for $\alpha = |\alpha|e^{i\theta_\alpha}$ and $\beta = |\beta|e^{i\theta_\beta}$. Similarly,

$$\begin{aligned} \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\langle c_j^* d_k\rangle =& -\frac{e^{i(\phi_2-\phi_1)}}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\Bigg\{\sin(\theta)\cos(\theta)\Bigg[\Big\langle 1 - \frac{1}{2}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j} + \delta\theta_{1,j}^2\big) - \frac{1}{2}\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k} + \delta\theta_{1,k}^2\big)\Big\rangle \\ &+ \Big\langle\frac{1}{4}\big(\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\theta_{1,j}^2\big)\big(\delta\phi_{1,k}^2 + \delta\chi_{1,k}^2 + \delta\theta_{1,k}^2\big)\Big\rangle + \delta_{j,k}\big\langle\delta\phi_{1,j}^2 + \delta\chi_{1,j}^2 + \delta\phi_j^2\delta\chi_{1,j}^2\big\rangle\Bigg] \\ &- \sin(\theta)\cos(\theta)\delta_{j,k}\big\langle\theta_j^2\big\rangle\big\langle 1 + \delta\phi_j^2 + \delta\chi_j^2\big\rangle\Bigg\} \\ \Rightarrow& \frac{1}{N^2}\sum_{j=1}^{N}\sum_{k=1}^{N}\alpha^*\beta\langle c_j^* d_k\rangle + \alpha\beta^*\langle d_j c_k^*\rangle = -2|\alpha||\beta|\cos(\gamma)\sin(\theta)\cos(\theta)\left(1 - 3\nu + \frac{9}{4}\nu^2 + \frac{\nu-\nu^2}{N}\right). \end{aligned} \tag{B12}$$

Putting this all together yields

$$P_s(N) = 1 - 3\nu + \frac{3\nu}{N} + \frac{9\nu^2}{4} + \frac{3\nu^2}{N}. \tag{B13}$$

Note that this function is only accurate up to $O(\nu)$ as ν^2 terms can arise due to $\langle O(\delta x^4)O(\delta x^0)\rangle$ which have not been included. For this reason a calculation including fourth order in δx was completed as above to yield

$$P_s(N) \approx 1 - 3\nu + \frac{3\nu}{N} + 4\nu^2 - \frac{4\nu^2}{N} - \frac{21\nu^3}{8} + \frac{\nu^3}{12N} + \frac{49\nu^4}{64} + \frac{13\nu^4}{6N}, \tag{B14}$$

which is similarly only complete up to $O(\nu^2)$. This is the cause of the erroneous behavior where $P_s(N=1) < 0$ for $\nu > 0$.

3. Solving for fidelity

Solving for the fidelity can be done in much the same manner, particularly as after renormalization the corrected state will be pure. As such we can write the corrected state fidelity as

$$\begin{aligned}\mathcal{F}(N) =& \langle\Psi|\hat{\rho}_{\rm ps}(N)|\Psi\rangle = [P_s(N)]^{-1}\left(1 + \frac{1}{N}\sum_{j=1}^{N}\langle\psi|\hat{U}_T^\dagger\hat{E}_j|\psi\rangle\right)^2 \\ =& [P_s(N)]^{-1}\left(1 + \frac{\mathcal{N}}{N}\sum_{j=1}^{N}(|\alpha|^2\langle a^*\delta a_j + c^*\delta c_j\rangle + |\beta|^2\langle b^*\delta b_j + d^*\delta d_j\rangle + \alpha^*\beta\langle a^*\delta b_j + c^*\delta d_j\rangle + \alpha\beta^*\langle b^*\delta a_j + d^*\delta c_j\rangle)\right).\end{aligned} \tag{B15}$$

So once again going term by term and here expanding δa_j, δb_j, δc_j, and δd_j up to fourth order in the error terms gives

$$\frac{1}{N}\sum_{j=1}^{N}\langle a^*\delta a_j\rangle = \frac{1}{N}\sum_{j=1}^{N}\langle d^*\delta d_j\rangle = \sin^2(\theta)\left(-\frac{3\nu}{2} + \frac{7\nu^2}{8}\right), \tag{B16}$$

$$\frac{1}{N}\sum_{j=1}^{N}\langle b^*\delta b_j\rangle = \frac{1}{N}\sum_{j=1}^{N}\langle c^*\delta c_j\rangle = \cos^2(\theta)\left(-\frac{3\nu}{2} + \frac{7\nu^2}{8}\right), \tag{B17}$$

and

$$\frac{1}{N}\sum_{j=1}^{N}\alpha^*\beta\langle a^*\delta b_j\rangle + \alpha\beta^*\langle b^*\delta a_j\rangle = 2|\alpha||\beta|\cos(\gamma)\sin(\theta)\cos(\theta)\left(-\frac{3\nu}{2} + \frac{7\nu^2}{8}\right), \tag{B18}$$

where $\gamma = \phi_1 - \phi_2 - \theta_\alpha + \theta_\beta$ for $\alpha = |\alpha|e^{i\theta_\alpha}$ and $\beta = |\beta|e^{i\theta_\beta}$. Similarly,

$$\frac{1}{N}\sum_{j=1}^{N}\alpha^*\beta\langle c^*\delta d_j\rangle + \alpha\beta^*\langle d^*\delta c_j\rangle = -2|\alpha||\beta|\cos(\gamma)\sin(\theta)\cos(\theta)\left(-\frac{3\nu}{2} + \frac{7\nu^2}{8}\right). \tag{B19}$$

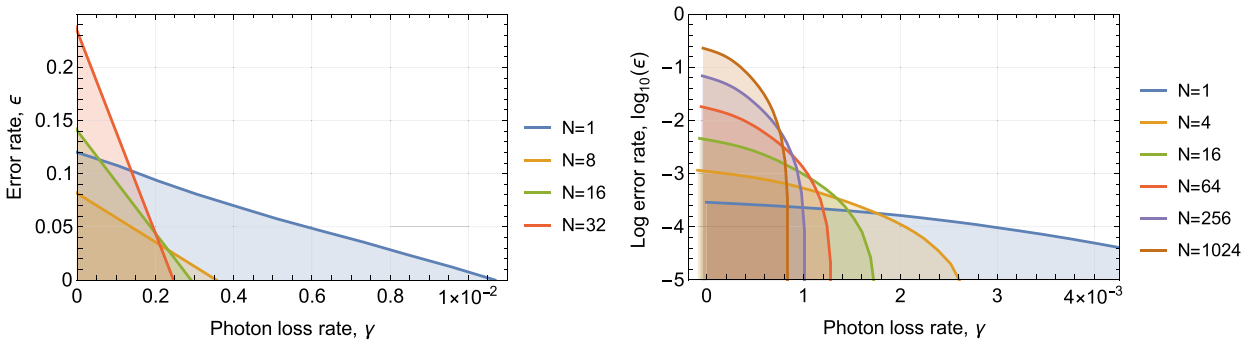
Putting this all together yields

$$\mathcal{F}(N) \approx [P_s(N)]^{-1}\left(1 - \frac{3\nu}{2} + \frac{7\nu^2}{8}\right)^2 \approx \left(1 - 3\nu + \frac{3\nu}{N} + 4\nu^2 - \frac{4\nu^2}{N}\right)^{-1}(1 - 3\nu + 4\nu^2). \tag{B20}$$

APPENDIX C: FUSION GATES

The type-II gate is implemented by the transformation

$$\begin{aligned}F_{\rm II} =& \begin{bmatrix}\sin(\theta_3) & \cos(\theta_3) & 0 & 0\\ \cos(\theta_3) & -\sin(\theta_3) & 0 & 0\\ 0 & 0 & \sin(\theta_4) & \cos(\theta_4)\\ 0 & 0 & \cos(\theta_4) & -\sin(\theta_4)\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\sin(\theta_1) & \cos(\theta_1) & 0 & 0\\ \cos(\theta_1) & -\sin(\theta_1) & 0 & 0\\ 0 & 0 & \sin(\theta_2) & \cos(\theta_2)\\ 0 & 0 & \cos(\theta_2) & -\sin(\theta_2)\end{bmatrix} \\ =& \begin{bmatrix}\sin(\theta_1)\sin(\theta_3) & \cos(\theta_1)\sin(\theta_3) & \cos(\theta_2)\cos(\theta_3) & -\sin(\theta_2)\cos(\theta_3)\\ \sin(\theta_1)\cos(\theta_3) & \cos(\theta_1)\cos(\theta_3) & -\cos(\theta_2)\sin(\theta_3) & \sin(\theta_2)\sin(\theta_3)\\ \cos(\theta_1)\cos(\theta_4) & -\sin(\theta_1)\cos(\theta_4) & \sin(\theta_2)\sin(\theta_4) & \cos(\theta_2)\sin(\theta_4)\\ -\cos(\theta_1)\sin(\theta_4) & \sin(\theta_1)\sin(\theta_4) & \sin(\theta_2)\cos(\theta_4) & \cos(\theta_2)\cos(\theta_4)\end{bmatrix}.\end{aligned} \tag{C1}$$



(a) (b)

FIG. 7. Fault-tolerant parameter space improvement estimation when employing unitary averaging in (a) a FBQC utilizing a six-ring architecture and (b) the 23-qubit Golay code. The $N = 1$ data were taken from [3] for the FBQC code and from [10] for the Golay code. Fault tolerance is achieved within the shaded region for each value of N.

Proceeding as before but with the aid of *Mathematica*, we can see, for a general two-photon input state, that the probability of success is given by

$$P_{s,\mathrm{II}}(N) = 1 - 2\nu + \frac{2\nu}{N} + \frac{5\nu^2}{3} - \frac{5\nu^2}{3N} \tag{C2}$$

and the postselected fidelity

$$\mathcal{F}_{\mathrm{II}}(N) = P_{s,\mathrm{II}}(N)^{-1}\left(1 - 2\nu + \frac{5\nu^2}{3}\right), \tag{C3}$$

where once again each parameter is taken to have an equal but independent noise spectrum with the parameters $\mathcal{O} = \mathcal{O}_T + \delta\mathcal{O}$, $\langle\mathcal{O}\rangle = \mathcal{O}_T$, $\langle\delta\mathcal{O}\rangle = 0$, $\langle\delta\mathcal{O}^2\rangle = \nu$, $\langle\delta\mathcal{O}^3\rangle = 0$, and $\langle\delta\mathcal{O}^4\rangle = 3\nu^2$.

APPENDIX D: UNITARY AVERAGING AND STANDARD FAULT TOLERANCE

Here we consider how the fault-tolerant parameter space changes if the operations performed on each physical qubit utilize UA. We concern ourselves with the per gate depolarization probability ε and per qubit, per gate loss rate γ.

The easiest consideration is the loss rate. This is because the UA scheme does nothing to correct for losses but does increase the optical depth due to the encoding and decoding circuits. As discussed earlier, it will increase the optical depth per gate by $2\log_2(N)$. Furthermore, the heralded error can be treated as loss. Here we will not seek to use the located nature of this loss to our advantage and treat it similarly to all other optical losses. As such, the UA scheme acts to increase the effective loss rate depending on the depth of encoding and decoding steps and the probability of success. If the original loss rate was γ per qubit per gate, we can estimate the new effective loss rate for a gate as $\Gamma = \frac{\gamma}{3}[3 + 2\log_2(N)] + [1 - P_s(N)]$ per qubit per gate. We have assumed all gates to have an optical depth of a single-qubit gate (3) and that encoding and decoding steps are as lossy as each individual component within the gate. Given some gates are likely to be much deeper than this, we are over estimating the relative loss introduced by the additional encoding.

We next turn our attention to the error rates. Consider that the gate fidelity relates to the probability that the qubit(s) it acts on would be measured in the incorrect state. After the gate is applied, with probability $\mathcal{F}(N) = 1 - \varepsilon$, the correct transformation is applied, and thus with probability ε any syndrome measurement will herald an error. This suggests an equivalence between our characteristic noise parameter $\mathcal{V}$ and the typical noise parameter considered in fault tolerance ε (the per qubit, per gate depolarization rate). This typical noise parameter corresponds to the more general depolarization errors occurring, which are not present in this model. However, if we take the depolarization effects as sourced by these same stochastic noises, we can indeed treat these two noise terms as equivalent. Comparing to the probability of success and fidelity shown in Eqs. (22) and (23), we have effective gate error $\mathcal{E}$ and loss Γ rates of

$$\mathcal{E} = \frac{\varepsilon}{N + \varepsilon - N\varepsilon}, \tag{D1}$$

$$\Gamma = \frac{\gamma}{3}[3 + 2\log_2(N)] + \varepsilon\left(1 - \frac{1}{N}\right). \tag{D2}$$

We can use these equations in conjunction with known fault-tolerant thresholds to gain insight into how the fault-tolerant parameter space may change when utilizing UA. If we take the earlier results of Dawson *et al.* [10] and Fujii and Tokunaga [11], we can estimate the effect UA has on the fault-tolerant parameter space for a 23-qubit Golay code. We can also consider its effect on a more modern fault-tolerant architecture such as fusion-based quantum computation [3]. The results are shown in Fig. 7.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[2] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Photonic boson sampling in a tunable circuit, Science **339**, 794 (2013).

[3] S. Bartolucci, P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant *et al.*, Fusion-based quantum computation, Nat. Commun. **14**, 912 (2023).

[4] T. C. Ralph, A. J. F. Hayes, and A. Gilchrist, Loss-tolerant optical qubits, Phys. Rev. Lett. **95**, 100501 (2005).

[5] R. J. Marshman, A. P. Lund, P. P. Rohde, and T. C. Ralph, Passive quantum error correction of linear optics networks through error averaging, Phys. Rev. A **97**, 022324 (2018).

[6] M. K. Vijayan, A. P. Lund, and P. P. Rohde, A robust W-state encoding for linear quantum optics, Quantum **4**, 303 (2020).

[7] D. Singh, A. P. Lund, and P. P. Rohde, Optical cluster-state generation with unitary averaging, arXiv:2209.15282.

[8] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, Nature (London) **409**, 46 (2001).

[9] T. Rudolph, Why I am optimistic about the silicon-photonic route to quantum computing, APL Photonics **2**, 030901 (2017).

[10] C. M. Dawson, H. L. Haselgrove, and M. A. Nielsen, Noise thresholds for optical quantum computers, Phys. Rev. Lett. **96**, 020501 (2006).

[11] K. Fujii and Y. Tokunaga, Error and loss tolerances of surface codes with general lattice structures, Phys. Rev. A **86**, 020303(R) (2012).

[12] A. J. F. Hayes, H. L. Haselgrove, A. Gilchrist, and T. C. Ralph, Fault tolerance in parity-state linear optical quantum computing, Phys. Rev. A **82**, 022323 (2010).

[13] A. J. F. Hayes, A. Gilchrist, and T. C. Ralph, Loss-tolerant operations in parity-code linear optics quantum computing, Phys. Rev. A **77**, 012310 (2008).