



CHORUS

This is the accepted manuscript made available via CHORUS. The article has been published as:

Optimizing the Frequency of Quantum Error Correction

Ali Abu-Nada, Ben Fortescue, and Mark Byrd

Phys. Rev. Lett. **119**, 190502 — Published 9 November 2017

DOI: [10.1103/PhysRevLett.119.190502](https://doi.org/10.1103/PhysRevLett.119.190502)

Optimizing the Frequency of Quantum Error Correction

Ali Abu-Nada¹, Ben Fortescue¹, Mark Byrd^{1,2}

¹Physics Department, Southern Illinois University, Carbondale, IL 62901, USA and

²Computer Science Department, Southern Illinois University, Carbondale, IL 62901, USA

(Dated: July 24, 2017)

A common assumption is that one applies fault-tolerant quantum error correction (FTQEC) after every gate during quantum computing. However, it is known that this is not always optimal since the FTQEC procedure itself can introduce errors. Here we vary the number of logical gates between FTQEC operations given that a failure of a postselection condition may cause FTQEC to be skipped. We derive an expression for the logical error rate as a function of error-correction frequency, and find the *optimal frequency* for the application of FTQEC. Furthermore, we show this is relatively insensitive to postselection failure probability for a large range of such probabilities. We provide an example of the application of the analytic expression to the $[[7, 1, 3]]$ Steane code and data derived from Monte Carlo simulation.

Quantum error correcting codes (QECCs) [1–5] will be necessary for reliable storage and processing of quantum information. The redundant information used in these codes protects quantum information by enabling the detection and correction of errors assuming they are below a certain threshold value [6] and assuming that they are operated in a fault tolerant manner [7]. Fault tolerance means that information can be stored (and/or manipulated) reliably even if individual components are faulty.

The largest known classes of QECCs use ancillary qubits (ancillas) to extract the measurement syndrome that identifies correctable errors. The ancillas may be prepared in several different ways: using the Shor technique (the ancilla is in a cat state) [8], Steane technique (the ancilla is in a logical state) [9], or Knill ancillas [10]. In all cases the errors must not propagate to the data in order to ensure fault tolerance. The qubits used for encoding together with the ancillas, lead to a large overhead since many physical qubits are utilized to encode and protect one logical (encoded) qubit.

Reducing these overheads is a particularly important theoretical task. Topological codes [4, 11] may be preferable, but recent work comparing resources in topological and concatenated codes suggests that the best code to use is highly dependent on the details of the physical constraints [12]. Brooks, et al. analyzed the Bacon-Shor code when X and Z errors have different probabilities thus reducing the size of the code for those errors which are less prevalent [13]. Weinstein has provided a relation between the fidelity and physical error rates for different numbers of gates in the $[[7, 1, 3]]$ Steane code [14]. He found that the fidelity is reduced slowly after skipping the QEC [14, 15]. This leads one to believe that QEC after every step is certainly not necessary and thus uses resources unnecessarily.

In addition to large overhead requirements of QEC, checking the ancillas for errors before use helps ensure fault tolerance. Most QECCs assume an error correction step after each step. (A step may be a WAIT gate, or a logical gate.) However, if an ancilla does not pass the

check, the QEC must be skipped as a practical matter.

Our work simultaneously addresses both of these problems. First, if one skips a QEC, how is the logical error probability affected? Second, if one skips QEC to save resources, how many can one skip without a large increase in the logical error probability given that an ancilla check may fail? We find the optimal number of steps before implementing a full QECC procedure. This saves both time and resources.

We emphasize that this optimization procedure identifies the number of steps a computation completes before QEC, i.e. the frequency of the QEC. Other researchers have optimized the encoding and decoding procedure to find the best code for a set of errors [16–18]. In other work, a message-passing algorithm was shown to aid in QEC for a concatenated code [19]. Here we assume a given code, without concatenation, and provide the optimal QEC frequency. Our methods may be combined with these other methods of optimization/resource reduction.

In the next section we find an analytic expression for the cumulative logical error probability P_L after a logical data qubit undergoes operations from N logical gates in total, with m gates in between each QEC. We consider the $[[7, 1, 3]]$ Steane code [2] and the Steane ancilla technique and determine P_L after (N/m) “blocks” of gates and QEC operations. We express P_L as a function of m and the physical gate error ϵ_g , and then minimize as a function of m . We then compare the predicted error rate to that obtained via Monte Carlo simulation.

Derivation of P_L : Logical gate and QEC model—In general, the dependence of the logical error rate on the physical gate error rate depends on the circuits used to implement gates and QEC. While our analysis is based on the Steane code, we use a model that could be readily adapted to other codes. We thus produced a semi-abstracted model based on the Steane code with Steane ancillas. In this QEC approach X and Z errors are corrected separately, and our model thus considers the propagation and correction of only one kind of error (with the same analysis applicable to both X and Z).

Logical errors may be introduced in two ways, from the logical gates and the QEC itself (in this analysis we do not treat errors from movement or hold operations as a separate category, they are incorporated into the above categories). We model a noisy physical gate as performing the desired operation followed by, with probability ϵ_g , an error. We treat logical gates as transverse, that is, consisting of a single physical gate applied to each qubit.

Our model of the QEC is more approximate. We divide errors induced by QEC into four separate parts (as shown in Figure 1) with the following probabilities:

1. “Correction errors”, with probability ϵ_c per qubit, are defined as errors in the QEC affecting data qubits only (not the ancilla measurement). Since the correction operations can be implicitly performed using the “Pauli frame” [20] rather than physical gates, such errors are limited to those failures in two-qubit gates used to interact the data with ancillas where only the data is affected.

The next three types of errors affect the data by causing an incorrect syndrome measurement, and thus an incorrect correction operation.

2. We will define “syndrome errors”, with probability ϵ_s per qubit, as those where an error on the QEC ancilla (or its measurement) *only* (i.e., with no errors directly applied to the data) cause a data qubit to be wrongly “corrected”, in addition to any errors already present.

3. “Omission errors”, with probability ϵ_o per qubit, represent erroneous syndromes which combine with existing data errors (where present) to (wrongly) return a syndrome indicating no errors. Thus if the data is initially without error, an omission error in the QEC will lead to an error on the data, but if a single qubit error is present on the data prior to the QEC, an omission error will lead to this error remaining uncorrected.

4. Finally, “double errors”, with probability ϵ_d per qubit, are where a single failure in the gate joining the data to the ancilla leads to errors on both outputs. (We only consider this class of errors when source and target errors on two-qubit gates are correlated). This is different than a standard syndrome error since, for example, a double error in isolation is equivalent to a data error correctly propagated to the ancilla and thus corrected. Conversely, a double error along with a syndrome error in the same QEC can lead to two data errors (while 2 syndrome errors will only produce 1 data error).

When the ancilla does not pass the check, which happens with probability ϵ_a (“ancilla” error, not per qubit) the QEC operation is not performed. Such events do not produce data errors, but result in existing errors not being corrected when they should be. Note that all of the above are functions of the physical gate error, ϵ_g , but the exact relationship depends on the QECC used, thus we treat them as separate variables.

Derivation of P_L : Logical error rate—The sequence of (N/m) blocks may, in a distance-3 code, produce a logical error if two or more data qubits end up with un-

corrected physical errors. To estimate the logical error rate P_L , we enumerate the ways in which this can occur, given that successful QEC operations will remove errors, limiting the possible ways to create a logical error. The fault-tolerant design means a logical error will only occur with probability second-order or higher in the various error probabilities. We are particularly interested in the regime where the ancilla errors are significantly larger than the other errors, (as can occur with complex ancilla creation circuits). While ancilla errors do not affect the data directly, they can result in multiple QEC operations being skipped with a consequent increase in the logical error probability.

We suppose a set of m gates (referred to “block” of gates) is implemented followed by a QEC operation. In the absence of skipped QECs the leading-order contributions to the logical error rate are limited: two errors can occur within a block or across two adjacent blocks. *Additional* errors that may arise by one or more skipped QECs follow a simple pattern: the only second-order errors arising from f skipped QECs, but wouldn’t arise from $f - 1$, is where the skipped QECs all occur sequentially, and the two errors in question are on the block containing the first skipped QEC and the block following the final skipped QEC.

For a sequence of $B \equiv N/m$ blocks, there are $B - f$ ways to have f sequential skipped QECs. Thus if two errors within the same block can produce a logical error when no QECs are skipped (for an overall probability weighted by the number of blocks N/m), the corresponding combination of errors occurring in different blocks separated by f skipped QECs is weighted by a factor

$$\gamma \equiv \sum_{f=1}^{B-1} \epsilon_a^f (B - f) = \frac{B\epsilon_a(1 - \epsilon_a) - \epsilon_a + \epsilon_a^{B+1}}{(1 - \epsilon_a)^2} \quad (1)$$

Similarly if two errors across two adjacent blocks can produce a logical error when no QECs are skipped, the corresponding weighting factor for f skipped QECs is

$$\gamma_3 \equiv \sum_{f=1}^{B-1} \epsilon_a^f (B - f - 1) = (\gamma - (B - 1)\epsilon_a)/\epsilon_a. \quad (2)$$

Summing all possible second-order errors (a full breakdown is provided in the supplemental material), we obtain the second-order formula for P_L :

$$\begin{aligned} P_L = 42 \Big[& B \left(m\epsilon_g \left(\frac{m\epsilon_g}{2} + (1 - \epsilon_a)(\epsilon_s + \epsilon_d) \right) \right. \\ & + (1 - \epsilon_a) \left(\epsilon_c \left(\epsilon_s + \epsilon_o + \frac{\epsilon_c}{2} \right) + \epsilon_d \left(\epsilon_s + \frac{\epsilon_d}{2} \right) \right) \\ & + ((B - 1 + \gamma_3)(1 - \epsilon_a)(\epsilon_c + \epsilon_s + \epsilon_o) + \gamma m\epsilon_g) \\ & \left. \times (m\epsilon_g + (1 - \epsilon_a)(\epsilon_s + \epsilon_d)) \right] \quad (3) \end{aligned}$$

(See supplemental material for further detail.)

Minimizing $P_L(m)$ — P_L is a discrete function of m . In the limit of many blocks ($B \rightarrow \infty$) we can express P_L as

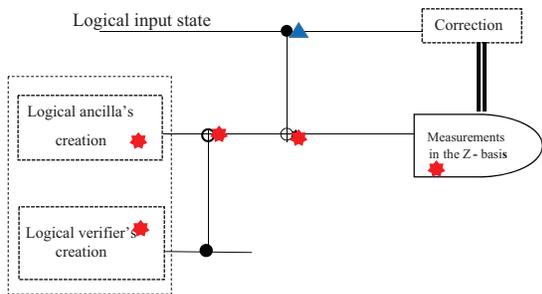


FIG. 1. (Color online) “Noisy” QEC for correction of X errors. The star represents the location of a syndrome error, and triangle a location of a correction error. The Steane ancilla state is $|+L\rangle$. Measurements are performed in the Z -basis.

a function $P_{L_p}(m) \simeq dm^{-1} + c_0 + c_1m$, where

$$d = 42B(1 - \epsilon_a) \left[\begin{aligned} &\epsilon_c \left(\epsilon_s + \epsilon_o + \frac{\epsilon_c}{2} \right) + \epsilon_d \left(\epsilon_s + \frac{\epsilon_d}{2} \right) \\ &+ (\epsilon_s + \epsilon_d)(\epsilon_c + \epsilon_s + \epsilon_o) \end{aligned} \right] \quad (4)$$

$$c_0 = 42B\epsilon_g(\epsilon_c + 2\epsilon_s + \epsilon_o + \epsilon_d), \quad c_1 = 42B\epsilon_g^2 \left(\frac{1}{1 - \epsilon_a} - \frac{1}{2} \right) \quad (5)$$

m_{min} satisfies $P_L(m_{min}) < P_L(m_{min} - 1), P_L(m_{min}) < P_L(m_{min} + 1)$, thus from we have $m_{min}(m_{min} + 1)c_1 - d > 0, m_{min}(m_{min} - 1)c_1 - d > 0$ and hence, since m_{min} is positive, it is the unique integer satisfying

$$\sqrt{\frac{1}{4} + \frac{d}{c_1}} - \frac{1}{2} < m_{min} < \sqrt{\frac{1}{4} + \frac{d}{c_1}} + \frac{1}{2}. \quad (6)$$

Thus the dependence of P_L on m is determined by the variable

$$\frac{d}{c_1} = [\epsilon_g^2(1 + \epsilon_a)]^{-1} 2(1 - \epsilon_a)^2 \left[\begin{aligned} &\epsilon_c \left(\epsilon_s + \epsilon_o + \frac{\epsilon_c}{2} \right) \\ &+ \epsilon_d \left(\epsilon_s + \frac{\epsilon_d}{2} \right) + (\epsilon_s + \epsilon_d)(\epsilon_c + \epsilon_s + \epsilon_o) \end{aligned} \right]. \quad (7)$$

Monte Carlo Simulation of P_L —As discussed above, our analytical formula for the logical error simplifies the description of QEC errors (in general a function of complex ancilla circuits) to the variables $\epsilon_{s,o,c,d}$, which we assume are the same for every qubit. In order to check the accuracy of this approximation, we performed Monte Carlo simulations of the QEC for the $[[7,1,3]]$ Steane code and Steane ancilla technique using QASM-P, simulation software based on QASM [21], in order to compare the logical error rates obtained with those predicted.

Initially, all gates were simulated using the stochastic error model for depolarizing noise [22]. In this case, we considered bit-flip (X) errors only on the data qubits (Z errors may be dealt with independently in the $[[7,1,3]]$ code, and we assume at equal rates). We used $N = 1000$

with varying block sizes $m \in \{1, 2, 4, 5, 8, 10, 20, 25, 100\}$, thus B varied between 1000 and 10. In order to accurately simulate the errors obtained in the verification process but still vary ϵ_a independently of ϵ_g , QECs are skipped with probability ϵ_a , but if the QEC is not skipped, ancilla verification is repeated until verification is passed (at which point that ancilla is used in the QEC).

To determine P_L , the data is prepared without error, in a logical $|0\rangle$ state. Then a series of blocks of m transversal logical gates (to simulate errors, these are simply wait operations which do not change the qubits' state in the absence of errors), followed by one QEC operation per block, are applied, for a total of N logical gates and $B = N/m$ blocks and (attempted) QECs. Finally, the data is checked for logical X errors. Each simulation (for a given choice of variable values ϵ_g and ϵ_a has 10^6 runs). From the Monte Carlo simulation we therefore obtain P_L as a function of the physical error rates.

The graphs below show detailed agreement between the numerical monte carlo simulations of the Steane code with the Steane ancilla and the analytic expression.

Numerical estimation of ϵ_s and ϵ_o —By our definition, the only source of correction errors is the CNOT gate interacting the data with the ancilla. Such errors occur only when the gate failure leads to an X error on the CNOT source (the data) but not the CNOT target. Similarly, double errors occur when a CNOT failure leads to X errors on both outputs. In our depolarizing error model single qubit gates undergo X, Y or Z errors with equal probability $\epsilon/3$, and two-qubit gates undergo the 15 possible two-qubit errors ($X \otimes I, Y \otimes I \dots Z \otimes Z$) with equal probability $\epsilon/15$. Since our analysis only considers bit errors (introduced by either X or Y operators), we have a single-qubit gate bit error probability of $\epsilon_g = 2\epsilon/3$. Thus the probability of a CNOT source-only error, coming from $X \otimes I, X \otimes Z, Y \otimes I, Y \otimes Z$ is $\epsilon_c = 4\epsilon/15 = 2\epsilon_g/5$. Likewise the probability of a double error comes from $X \otimes X, X \otimes Y, Y \otimes X, Y \otimes Y$ and hence $\epsilon_d = 2\epsilon_g/5$

ϵ_s and ϵ_o were determined directly from simulation. By our definition, a syndrome error or double error in a QEC will, for an input logical qubit containing one error, add a second error, leading to an overall logical error, and are the only first-order QEC errors which do this. Thus to estimate error rate $\epsilon_s + \epsilon_d$, the data was first prepared in a logical eigenstate, with an error on one of the seven qubits. The QEC procedure for the $[[7,1,3]]$ code was then performed using the stochastic error model for depolarizing noise [22]. Finally the logical qubit was checked for logical errors to determine the error rate.

Similarly, if the input data has a single logical error entering and leaving the QEC, this will be due to either a correction or omission error. Hence to estimate the QEC physical error rate $\epsilon_c + \epsilon_o$, we prepare the input logical qubit with an error on one of the seven corresponding physical qubits, then we perform the same QEC simulation procedure as before, but determine the rate based

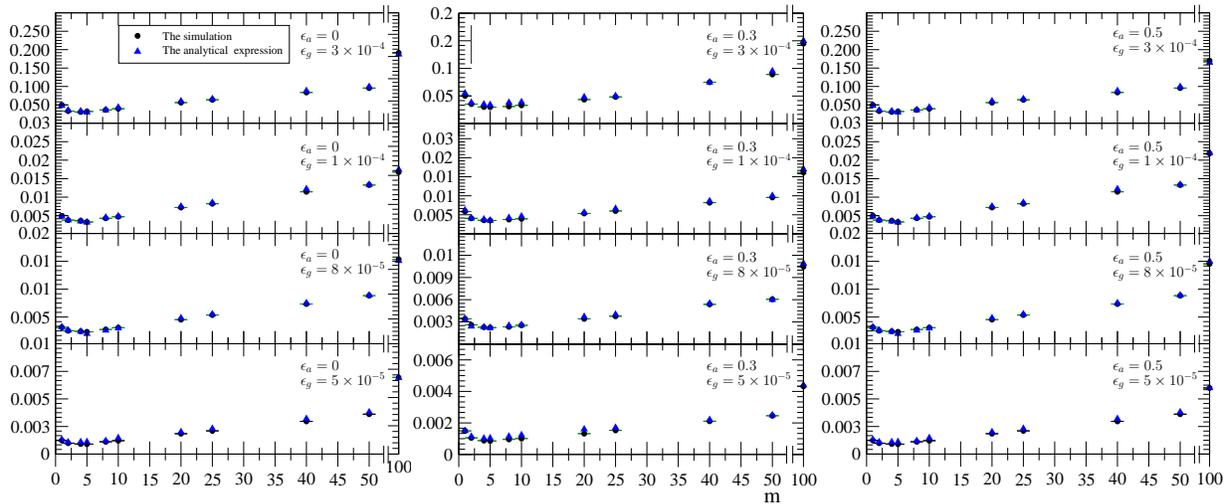


FIG. 2. (Color online) P_L vs m for $\epsilon_g = 5.0 \times 10^{-5}, 8.0 \times 10^{-5}, 1.0 \times 10^{-4}, 3.0 \times 10^{-4}$, where $\epsilon_a = 0.3$. The triangular blue points are the numerical values of P_L as given by the Equation (3) in the paper. The circular black points are the numerical simulation values of P_L .

on events when the output has a single error (rather than two). In both cases we varied the input error over all 7 qubits and took the mean resultant P_L .

We performed the simulation with a variety of numerical values for ϵ_g ($10^{-5} \sim 10^{-3}$). For each different value of ϵ_g , we determine the numerical values of ϵ_s and $\epsilon_c + \epsilon_o$. The relationship is fitted to a linear equation to determine the coefficients for ϵ_s vs. ϵ_g and $\epsilon_c + \epsilon_o$ vs. ϵ_g , which were: $\epsilon_s + \epsilon_d = 3.85\epsilon_g$, $\epsilon_c + \epsilon_o = 1.01\epsilon_g$, $\Rightarrow \epsilon_s = 3.45\epsilon_g$, $\epsilon_o = 0.61\epsilon_g$.

Results and discussions—In the supplemental material, we show the relationship between P_L and ϵ_a for a given ϵ_g and also m_{min} versus ϵ_a . These vary as expected.

More importantly, Figures 2 show the relationship between P_L and m for the case $\epsilon_a = 0.3$, for gate error values $\epsilon_g = 5.0 \times 10^{-5}, 8.0 \times 10^{-5}, 1.0 \times 10^{-4}$, and 3.0×10^{-4} . There is generally good agreement between the data generated by the formula and the simulation (and reasonably good agreement even given the assumption of large B , especially the location of m_{min}), and m_{min} is insensitive to variations in ϵ_a over the range of ϵ_a considered, with $m_{min} = 5$ in all cases. Within the region $m < m_{min}$, the error rate is reduced both by increasing m and by increasing ϵ_a , since both result in fewer QEC operations (the only difference being whether the skipped operations are regularly spaced or not), and QEC operations in this region produce more errors, on average, than they correct. Similarly the behavior is reversed for $m > m_{min}$. Again, this is shown explicitly in Figures 2.

Note that P_L is the cumulative error for 1000 gates (plus QEC operations with some frequency), and hence may be larger than the underlying error for an individual physical gate, even when operating below threshold.

Conclusions—We have found the optimal number of quantum gates to perform before applying an error cor-

rection operation for a semi-abstract model that can be applied to a variety of codes although topological codes behave a bit differently and our method does not directly apply. The analytic expression provides explicit dependence on the error correction frequency as a function of the gate error rate, ancilla failure rate, and error rates for the correction operation. The various rates depend on the underlying physical gate error rate. The dependence is different for different circuits which are determined by the QECC used. To be explicit, we showed in detail how this works by example. Our example is the commonly used Steane $[[7,1,3]]$ code and Steane ancilla technique. Our model is applicable to either X or Z errors, although a full analysis of both would need to make the (reasonably straightforward) extension of considering errors of one kind produced in QECCs for the other, logical gates (e.g. the Hadamard) which convert between error types, and correlated X and Z errors from 2-qubit gate failures.

We have assumed a transversal gate model. Single-qubit logical operations may, depending on the computation and code, be dominated by non-transversal gates (e.g. the T gate). Such gates often require preparation of a post-selected ancilla and would require an error model similar to that used for QEC. Our treatment is expected to work very well for storage where one implements no operation, called a WAIT gate.

For the Steane code and Steane ancilla, we compared the results with the simulation using QASM-P and found excellent agreement showing that our coarse-graining (due to a rough classification of error types) provides enough precision to provide a very reliable estimate. Furthermore, we find that the optimum frequency to apply QEC operations is relatively insensitive to ancilla failure probability, (with the optimum varying from $m = 3$ to $m = 6$ but frequency changes within this range making

only small differences to the overall logical error), indicating that skipping QEC operations under ancilla failure will in many cases be a successful approach even in a design where QECs are performed infrequently. This will save resources while providing a better overall logical error rate for quantum error correcting codes.

Acknowledgments—We thank Brian Arnold, Ken Brown and Yaakov Weinstein for helpful discussions. Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D12PC00527. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

[1] P. W. Shor, Phys. Rev. A **52**, R2493(R) (1995).

[2] A. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[3] F. Gaitan, *Quantum error correction and fault tolerant quantum computing* (CRC Press, 2008).

[4] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, New York, 2013).

[5] B. Terhal, Rev. Mod. Phys. **87** (2015).

[6] D. Aharonov and M. Ben-Or, SIAM J. Comput. **38**, 1207 (2008).

[7] J. Preskill, *Introduction to Quantum Computation* (World Scientific, Singapore, H.-K. Lo, S. Popescu, T. P. Spiller, eds., 1998).

[8] Shor, P.W., in *Proceedings of the 37th Symposium on Foundations of Computing* (IEEE Computer Society Press, Los Alamitos, CA, 1996) p. 56, eprint quant-ph/9605011.

[9] A.M. Steane, Phys. Rev. Lett. **78**, 2252 (1997).

[10] E. Knill, Nature **434**, 39 (2005).

[11] A. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).

[12] M. Suchara, A. Faruque, C.-Y. Lai, G. Paz, F.T. Chong and J. Kubiatowicz, (2013), arXiv:1312.2316.

[13] P. Brooks and J. Preskill, Phys. Rev. A. **87**, 032310 (2013).

[14] Y. S. Weinstein, Phys. Rev. A **88**, 012325 (2013).

[15] Y. S. Weinstein, “Quantum error correction during 50 gates,” (2014).

[16] A.S. Fletcher, P.W. Shor and M.Z. Win, Phys. Rev. A **75**, 012338 (2007).

[17] M. Reimpell, R.F. Wrener and K Audenaert, (2006), arXiv:quant-ph/0606059.

[18] R. Kosut and D. Lidar, Quant. Inf. Proc. **8**, 443 (2009).

[19] D. Poulin, Phys. Rev. A **74**, 052333 (2006).

[20] A. M. Steane, Phys. Rev. A **68**, 042322 (2003).

[21] A. Cross, “QASM,” <https://www.media.mit.edu/quanta/quanta-web/projects/qasm-tools> (2006).

[22] A. Abu-Nada, B. Fortescue and M. Byrd, Phys. Rev. A **89**, 062304 (2014).