

POLYESTIMATE: A LIBRARY FOR NEAR-INSTANTANEOUS SURFACE CODE ANALYSIS

AUSTIN G. FOWLER^{1,2}

¹*CNSI, Department of Physics, University of California, Santa Barbara, California 93106, USA*

²*CQC2T, School of Physics, The University of Melbourne, Victoria 3010, Australia*

Received September 10, 2013

Revised July 28, 2014

The surface code is highly practical, enabling arbitrarily reliable quantum computation given a 2-D nearest-neighbor coupled array of qubits with gate error rates below approximately 1%. We describe an open source library, Polyestimate, enabling a user with no knowledge of the surface code to specify realistic physical quantum gate error models and obtain logical error rate estimates. Functions allowing the user to specify simple depolarizing error rates for each gate have also been included. Every effort has been made to make this library user-friendly. Polyestimate provides data essentially instantaneously that previously required hundreds to thousands of hours of simulation, statements which we discuss and make precise. This advance has been made possible through careful analysis of the error structure of the surface code and extensive pre-simulation.

Keywords:

Communicated by: R Jozsa & B Terhal

1 Introduction

Quantum computing hardware is not expected to achieve the same level of reliability as classical computing hardware due to its complexity and reliance on the fragile phenomena of quantum mechanics. Arbitrarily reliable quantum computation can, however, be achieved through the use of quantum error correction [1, 2, 3, 4]. A bright hope in the field of quantum error correction is the surface code [5, 6, 7, 8, 9], which has the very experimentally reasonable requirements of a 2-D array of nearest-neighbor coupled qubits capable of implementing initialization, measurement, and one- and two-qubit unitary gates, all with error rates below approximately 1% [10, 11]. Trade-offs are also possible, such as a measurement error rate of 10% or more at the cost of somewhat lower two-qubit gate error rates. An open source analysis tool, Autotune [12], exists that is capable of taking into account such error model details.

Autotune is, however, rather computationally expensive to run and complex to use. Consider Fig. 1. This contains two examples of different size surface codes containing a single logical qubit. The operators in Fig. 1 are measured using the circuits of Fig. 2. The circuits of Fig. 2 are built out of eight types of gates, namely initialization, measurement, Hadamard, CNOT, and four potentially different identity gates of durations equal to each of these non-

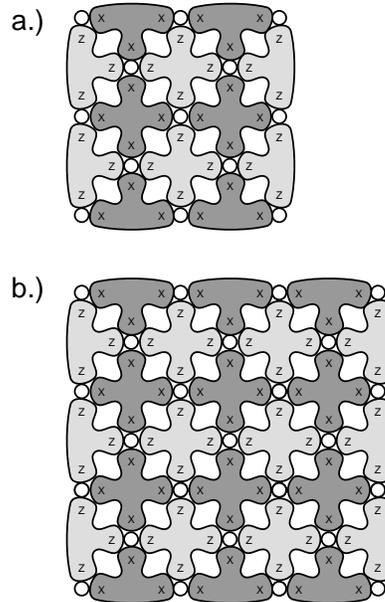


Fig. 1. a.) Distance 3 surface code. b.) Distance 4 surface code. Circles represent qubits. Bubbles represent operators (tensor product of Pauli X or Z operators) that are measured to detect errors. Note that all operators commute. Each bubble is associated with its own (syndrome) qubit, used to measure its operator (stabilizer [13]) via the circuits shown in Fig. 2. Each surface code contains a single logical qubit. A distance d surface code, properly implemented, can correct any combination of $\lfloor (d-1)/2 \rfloor$ errors.

trivial gates. It would be nice if a simple library existed taking error models or simple error rates for each gate and a code distance d as input and producing accurate estimates of the logical X and Z error rates of the protected logical qubit. In this work, we report the creation of Polyestimate, a library which returns such estimates essentially instantaneously. More precise comparative benchmarks can be found in Section 5.

Polyestimate is designed to work accurately on error models that are not too extremely asymmetric within a single error type on the CNOT gate. For example, Polyestimate will return inaccurate results (meaning excessively high logical error rate estimates) if, for some reason, the probability of X error on the control qubit of the CNOT is very much higher than the probability of an X error on the target qubit. Asymmetry between X and Z errors is, however, accurately handled on all gates. Polyestimate can also work quite accurately quite close to threshold — it does not rely on error rates being low as we did in [14], work which required error rates to be of order 10^{-4} or lower. These statements will be made more precise as the details of Polyestimate are explained.

The discussion is organized as follows. In Section 2, we describe how to reduce detailed error models for the eight types of gates described above to three simple error rates. In Section 3, we describe our simulation method and give selected examples of the data collected. In Section 4, we describe how Polyestimate interpolates and extrapolates from the simulation

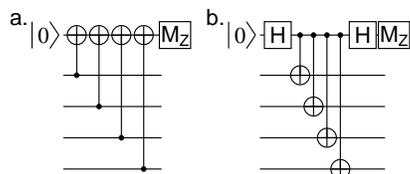


Fig. 2. Quantum circuits measuring a.) $ZZZZ$ and b.) $XXXX$ operators. The $|0\rangle$ represents initialization, the H represents Hadamard, the M_Z represents measurement of the operator Z , and the dot and target symbols connected by lines each represent a CNOT gate. The interaction sequence is with the North data qubit as shown in Fig. 1, then West, East, South. If no data qubit is present in a given direction, the CNOT is simply replaced with an identity gate of duration CNOT.

generated data to the user specified error models and code distance d . In Section 5, we compare the output of Polyestimate to that of detailed Autotune simulations for a range of situations, characterizing the accuracy of Polyestimate. Section 6 concludes and discusses further work.

2 Error classes

When using circuits of the form shown in Fig. 2, currently the only available algorithm to correct errors is based on minimum weight perfect matching [15, 16, 17]. While other algorithms for decoding the surface code certainly exist, including algorithms based on renormalization group methods [18, 19], Markov chain Monte Carlo methods [20, 21], expanding diamonds [22], and iterative methods [23], none of these algorithms have been tested on the output of fault-tolerant circuits making use of only nearest-neighbor two-qubit gates.

In this paper, we do not take into account correlations between X and Z errors. For example, if after a given quantum gate there is a probability p of X , Y , or Z error, each with probability $p/3$, then if we believe with high confidence that we have detected an X error after this gate we should also believe there to be a Z error with 50% probability since $Y = iXZ$. This information can be used to improve error correction [21, 24], leading to at most a factor of $2^{\lceil d/2 \rceil}$ reduction in the logical error rate, achieved for a standard depolarizing channel. If X and Z errors are uncorrelated, however, no benefit is obtained. Even if maximally correlated, if Z errors are much more likely than X errors, correction of Z errors will benefit little from the rare X error information, and the presence of Z errors will not strongly imply the presence of X errors. Building a fast library with the capacity to accurately take all of this complexity into account is unlikely to be feasible. Our goal is reasonable accuracy near instantaneously. The most accurate and optimized results will always come from CPU intensive simulations.

Treating X and Z errors independently is enormously helpful. An arbitrary single-qubit error model with probabilities p_X, p_Y, p_Z of errors X, Y, Z can simply be treated as two independent error models implying an X error with probability $p'_X = p_X + p_Y$ and a Z error with probability $p'_Z = p_Y + p_Z$. Missing out on the potential factor of $2^{\lceil d/2 \rceil}$ lower logical error rates may seem unacceptable, however the threshold error rate is not significantly improved by accounting for correlations [24], so the feasibility of a given technology can still be well tested using Polyestimate. Furthermore, the maximum factor of improvement is only achieved for quite low balanced depolarizing error rates $p \lesssim 2 \times 10^{-4}$ [24], with higher error rates, less

correlation, and more X - Z asymmetry all substantially limiting the benefit. Even when the maximum benefit is achieved, at a couple of orders of magnitude below the threshold error rate, the difference in the estimated code distance required to achieve a given logical error rate is very small, meaning Polyestimate is still viable as an algorithm overhead calculator, the primary purpose for its construction.

Splitting CNOT gate error models into pure X and Z terms requires care. Given 15 error probabilities $p_{IX}, p_{IY}, \dots, p_{ZZ}$ one can certainly make a similar identification $p'_{IX} = p_{IX} + p_{IY} + p_{ZX} + p_{ZY}$, $p'_{XX} = p_{XX} + p_{XY} + p_{YX} + p_{YY}$, etc, however if $p'_{XX} \neq p'_{IX}$ or $p'_{XX} \neq p'_{XI}$ we first artificially set all three parameters to be equal to the maximum to obtain a balanced $p'_{IX} = p'_{XI} = p'_{XX}$ error model then define $p_{2X} = 5(p'_{IX} + p'_{XI} + p'_{XX})/4$. The multiplicative factor of $5/4$ scales the 12 fundamental error rates making up $p'_{IX}, p'_{XI}, p'_{XX}$ to a 15 component depolarizing error rate with the correct probability of each of the derived X terms. If the CNOT error model is very asymmetric within a single type of error, the process of raising one or two of $p'_{IX}, p'_{XI}, p'_{XX}$ can make the logical error rate estimates excessively high. We shall give a concrete example of this in Section 4. Repeating the above process for Z errors gives a second value p_{2Z} . Note that p_{2X} and p_{2Z} as defined could be greater than 1, but only if one of their three constituent components had probability $4/15$ or greater, which would not be below the threshold error rate of the surface code, resulting in rejection of the error model.

Referring to Fig. 2, data qubit identity errors can be lumped into a single variable $p_{1X} = 3(p'_X(\text{IdInit}) + 2p'_X(\text{IdHad}) + p'_X(\text{IdMeas}))/8$, and similarly for Z errors, giving p_{1Z} . The abbreviations stand for identity of duration initialization, identity of duration Hadamard, and identity of duration measurement, respectively. Identity of duration CNOT does not appear in the equation as in even a small distance surface code only a few such identity gates are required around the boundary of the qubit lattice and their influence is negligible — the vast majority of data qubits are busy performing CNOTs when any data qubit is involved in a CNOT. While this may not be immediately apparent from Fig. 2, it should be remembered that all stabilizers are measured simultaneously and each data qubit not on the boundary of the surface code interacts with each of its four neighboring syndrome qubits. The factor of $3/8$ in the definition of p_{1X} consists of a factor of $3/2$ to scale the individual error type error rate to a depolarizing error rate, and a factor of $1/4$ to reduce the result to an error rate per gate rather than per stabilizer measurement cycle.

Syndrome qubits associated with Z stabilizer measurements, meaning X error detection, only suffer initialization and measurement errors in addition to CNOT errors. Initialization and measurement errors cannot propagate off the syndrome qubit. We can therefore define $p_{0X} = p_X(\text{Init}) + p_X(\text{Meas})$. Note that for initialization and measurement $p_Y = p_Z = 0$. For syndrome qubits associated with X stabilizer measurements, meaning Z error detection, we have two additional Hadamard gates, the first of which is only dangerous if it introduces Z errors, the second of which is only dangerous if it introduces X errors. We therefore get a second definition $p_{0Z} = p_X(\text{Init}) + p'_X(\text{Had}) + p'_Z(\text{Had}) + p_X(\text{Meas})$. We choose to leave p_{0A} as an error rate per error detection cycle as the number of gates for $A = X$ and $A = Z$ differs.

Given arbitrary error models, we have described how to obtain two sets of error rates p_{0A}, p_{1A}, p_{2A} focusing on either $A = X$ or $A = Z$ errors. These simple error rates well characterize the behavior of the detailed error models provided the CNOT error model does not deviate

very far from balanced depolarizing within a single type of error.

3 Simulations

It is straightforward to use Autotune to simulate the performance of a given surface code of distance d making use of gate error rates p_{0A} , p_{1A} , p_{2A} . Autotune will set up the appropriate array of qubits, execute the appropriate sequence of gates, insert stochastic errors and use minimum weight perfect matching to insert corrections. When too many errors occur in the wrong places, logical errors will occur. Autotune simulates many continuous cycles of error detection and calculates the probability of logical error per round.

Our goal is to obtain the logical X and Z error rates corresponding to a sufficiently broad range of d , p_{0A} , p_{1A} , p_{2A} to enable the logical X and Z error rates at any other combination of values of d , p_{0A} , p_{1A} , p_{2A} to be accurately determined.

Note that below threshold the logical error rate is provably exponentially suppressed with increasing d [4], or, to be more precise, exponentially suppressed with increasing $d_e = \lfloor (d + 1)/2 \rfloor$. This means that we only need to simulate distances $d = 3, 4, 5, 6$ as the logical error rate at all higher distances can be obtained by taking the ratio of odd or even distance data to the appropriate power with the appropriate pre-factor. Details are given in Section 4.

Note that the threshold error rate when the surface code is subjected to only CNOT gate errors is approximately 1.25% [9]. We therefore do not simulate values of p_{2A} above 2%, as this is sufficiently high to be above threshold for all values of p_{0A} and p_{1A} .

Let $r_{0A} = p_{0A}/p_{2A}$ and $r_{1A} = p_{1A}/p_{2A}$. We have initially considered only syndrome errors in the range $r_{0A} \in [0.01, 200]$ and data qubit errors in the range $r_{1A} \in [0.01, 1]$. This is essentially a statement that we consider it possible that initialization and measurement error rates could be significantly less, comparable, or significantly greater than the CNOT error rate, whereas we expect the single-qubit identity error rates to be less than or equal to the CNOT error rate. It would be straightforward to simulate a broader range if a quantum technology called for it. We have restricted ourselves to $p_{2A} \geq 10^{-4}$ as lower error rates are well out of reach of current universal and scalable quantum technology. The range $p_{2A} \in [10^{-4}, 0.02]$ covers all existing technology, and technology likely to be available in the short to medium term. Note that p_{2A} is not expected to be 0 in any physical technology. Input of an error model with p_{2A} lower than the lowest simulated value of p_{2A} , currently 10^{-4} , will result rejection of the model. To keep the size of the simulation task under control, we have initially restricted ourselves to three points per decade, so for example $r_{1A} \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1\}$.

Samples of the data collected can be found in Figs. 3–4. These two graphs contain data for the highest and lowest error ratios, respectively. Error bars corresponding to 2σ confidence have been included. At the very lowest error rates p_{2A} it is computationally challenging to observe a sufficiently large number of logical errors to obtain an accurate logical error rate. Nevertheless, the data presented is sufficiently accurate for practical purposes. This data is available as part of the Polyestimate library, available online at [12].

4 Interpolation and extrapolation

Given detailed error models for each gate or simple error rates for each gate, in Section 2 we described how to define two sets of three values $r_{0A} = p_{0A}/p_{2A}$, $r_{1A} = p_{1A}/p_{2A}$, p_{2A} , where $A = X, Z$, that can be used to estimate the failure behavior of the surface code. In Section 3,

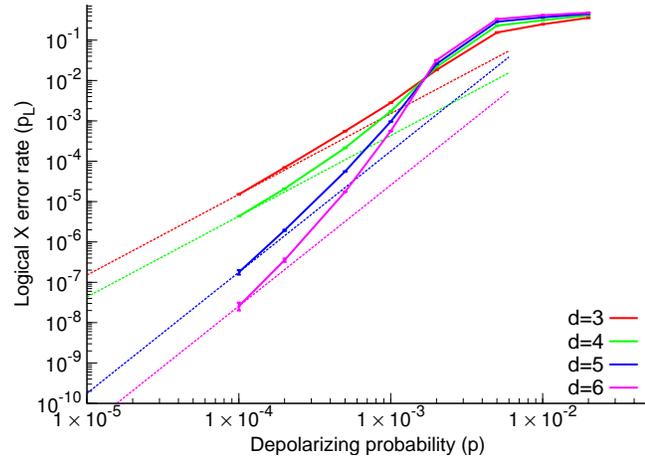


Fig. 3. (Color online) probability of logical X error as a function of the depolarizing error rate $p = p_{2A}$ for distances $d = 3, 4, 5, 6$ and $r_{0A} = 100$, $r_{1A} = 1$. Referring to the left of the figure, the distance $d = 3, 4, 5, 6$ curves are ordered top to bottom. Quadratic curves have been fit through the lowest $d = 3, 4$ data points, and cubic curves through the lowest $d = 5, 6$ data points.

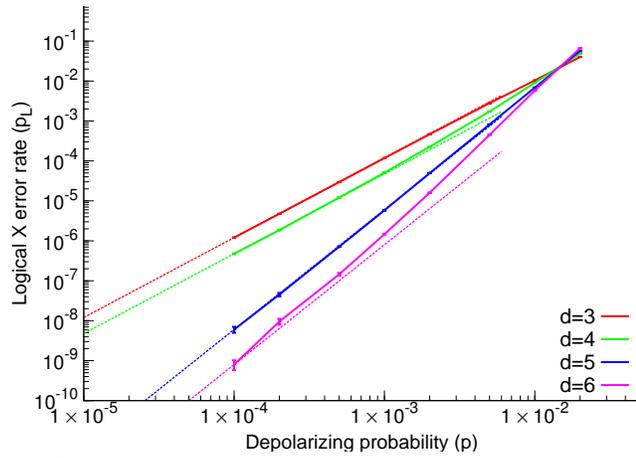


Fig. 4. (Color online) probability of logical X error as a function of the depolarizing error rate $p = p_{2A}$ for distances $d = 3, 4, 5, 6$ and $r_{0A} = 0.01$, $r_{1A} = 0.01$. Referring to the left of the figure, the distance $d = 3, 4, 5, 6$ curves are ordered top to bottom. Quadratic curves have been fit through the lowest $d = 3, 4$ data points, and cubic curves through the lowest $d = 5, 6$ data points.

we described the construction of a database of logical error rates for a range of code distances and values of r_{0A} , r_{1A} , p_{2A} . Note that a single database suffices for both $A = X$ and $A = Z$ as the database is for balanced depolarizing noise. We now describe how to obtain functions of d giving estimates of the logical X and Z error rates.

If $r_{0A}, r_{1A}, p_{2A} \neq 1, 2, 5 \times 10^i$ for some integer $i \in \mathbb{Z}$, interpolation will be required. Define superscripts H and L to mean the first variable value in the database higher or lower than the given value. For example, if $r_{0A} = 0.3$, then $r_{0A}^L = 0.2$ and $r_{0A}^H = 0.5$. Note that the logical error rates are strictly increasing smooth functions of r_{0A} , r_{1A} , p_{2A} . For a sufficiently fine-grained database, linear interpolation is therefore appropriate.

For each distance $d = 3, 4, 5, 6$ in the database and each type of logical error A , find the up to eight different values of logical A error for the up to eight different combinations of r_{0A}^B , r_{1A}^B , p_{2A}^B , where $B = H, L$. If there are eight different combinations of r_{0A}^B , r_{1A}^B , p_{2A}^B , these triples can be viewed as the coordinates of the corners of a cuboid. Along any edge of the cuboid, we define the logical error rate to be a linearly increasing or decreasing function based on the corner values. Across the faces and through the volume, we can similarly uniquely define a linearly interpolated logical error rate. Denote this linearly interpolated logical error rate by $p_{AL}(d, r_{0A}, r_{1A}, p_{2A})$. For brevity, we shall refer to this as simply $p_{AL}(d)$.

For $d > 6$, we extrapolate. At gate error rates below the surface code threshold error rate, suppression of logical error with increasing code distance is exponential (see eq. 12 of [4]). Define $x = p_{AL}(5)/p_{AL}(3)$, $C = p_{AL}(3)/x^2$, $y = p_{AL}(6)/p_{AL}(4)$, $D = p_{AL}(4)/y^2$. This gives us four expressions

$$p_{AL}(d) = Cx^{\lfloor \frac{d+1}{2} \rfloor}, \quad (1)$$

$$p_{AL}(d) = Dy^{\lfloor \frac{d+1}{2} \rfloor}, \quad (2)$$

valid for d odd and even, respectively, and $A = X, Z$.

5 Verification

The simplicity of derivation and form of Eqs. 1–2 is appealing, however verification of the accuracy of logical error rates obtained by Polyestimate is required. In this Section, we shall compare the output of Polyestimate with data generated in a number of situations by direct simulation with Autotune.

The first situation we shall consider is balanced depolarizing noise of equal strength on all gates. When all gates have probability $p = 10^{-3}$ of failure, Table 1 shows in columns 3 and 4 the Autotune determined probability of logical X and Z error for distances $d = 3, \dots, 8$. This should be compared with the Polyestimate data in columns 5 and 6. As the distance increases, the computational cost of using Autotune rapidly becomes prohibitive, whereas Polyestimate can be used for arbitrary distances, and generated the data in Table 1 in approximately one hundredth of one second, including loading the database from disk and calculating the appropriate values in Eqs. 1–2. This comparative timing information was obtained using a single core of a 16-core PowerEdge C6145 1.6GHz AMD processor with 2 GB of memory per core. The computational cost of evaluating Eqs. 1–2 for additional values of d is negligible after loading the database. Over 6.5 million evaluations can be completed in under one second using our chosen hardware.

d	T	p_{X_L}	p_{Z_L}	$p_{X_L}(d)$	$p_{Z_L}(d)$
3	10min	1.1×10^{-3}	1.4×10^{-3}	1.1×10^{-3}	1.4×10^{-3}
4	90min	4.5×10^{-4}	5.8×10^{-4}	4.5×10^{-4}	5.8×10^{-4}
5	7.7hr	1.0×10^{-4}	1.5×10^{-4}	9.9×10^{-5}	1.4×10^{-4}
6	48hr	3.2×10^{-5}	4.7×10^{-5}	3.2×10^{-5}	4.7×10^{-5}
7	250hr	8.5×10^{-6}	1.4×10^{-5}	9.0×10^{-6}	1.5×10^{-6}
8	500hr	2.5×10^{-6}	4.2×10^{-6}	2.2×10^{-6}	3.8×10^{-6}
9	-	-	-	8.2×10^{-7}	1.6×10^{-6}
10	-	-	-	1.6×10^{-7}	3.0×10^{-7}
36	-	-	-	1.5×10^{-22}	1.7×10^{-21}

Table 1. Comparison of Autotune and Polyestimate for depolarizing noise of rate $p = 10^{-3}$. Column 1 contains the code distance d . Column 2 contains the total CPU time required by Autotune to observe 10^4 logical errors enabling accurate estimation of the logical X and Z error rates shown in columns 3 and 4. Columns 5 and 6 contain Polyestimate generated logical X and Z error rates. Agreement is in all cases within 10%, sufficient for practical purposes.

The agreement between Autotune and Polyestimate is within 10% for the data shown in Table 1. It should be noted, however, that since Polyestimate uses a simple exponential expression (Eqs. 1–2), any small inaccuracy will grow with code distance, so absolute values of logical error rates for large distances will only be approximate. A user can still, however, quite accurately determine the code distance d required to suppress logical errors below some desired rate. For example, if a logical error rate below 10^{-20} is desired, Polyestimate can be used to predict that a code distance of approximately 36 is sufficient.

The next case we shall consider is depolarizing noise of rate $p = 10^{-3}$ for every gate except the measurement gate which shall have an error rate of 10%. It is possible that superconducting qubits may possess such a distribution of error rates as fast, low error measurement is especially challenging [25]. Agreement between Autotune and Polyestimate is again very good (Table 2), within 15% in all cases, frequently significantly better, and this agreement is exhibited over logical error rates varying by over two orders of magnitude. Similar levels of agreement have been observed for other ratios of simple depolarizing error rates on each gate.

To provide a little more insight into the computational cost of the simulations, consider the $d = 8$ data in Tables 1–2. All simulations in these tables were run until 10^4 of the least likely type of logical error were observed. In both these tables, logical X errors are least likely. This means that the total number of rounds of $d = 8$ error detection simulated can be calculated directly from the logical X error rates, so $10^4/2.5 \times 10^{-6} = 4.0 \times 10^9$ in Table 1 and $10^4/2 \times 10^{-4} = 5 \times 10^7$ in Table 2. This places the computational cost per round of error detection at $45\mu\text{s}$ in both cases, as expected since the computational cost is dominated by simulating the action of thousands of gates on hundreds of qubits each round of error detection.

Finally, we shall consider the case of all gates having total probability of error $p = 10^{-3}$, however the CNOT gate having an asymmetric error model satisfying $p'_{IX} = 10p'_{XI} = 100p'_{XX}$. We shall focus only on logical X errors. In this instance, since Polyestimate must raise the probability of XI and XX errors significantly in order to achieve an upper bound on the logical error rate using its pre-generated database, the logical X error estimates are excessively high (Table 3).

d	T	p_{X_L}	p_{Z_L}	$p_{X_L}(d)$	$p_{Z_L}(d)$
3	11min	2.8×10^{-3}	3.4×10^{-3}	2.8×10^{-3}	3.4×10^{-3}
4	40min	1.8×10^{-3}	2.2×10^{-3}	1.7×10^{-3}	2.3×10^{-3}
5	2hr	9.6×10^{-4}	1.3×10^{-3}	9.9×10^{-4}	1.3×10^{-3}
6	6.3hr	5.7×10^{-4}	7.9×10^{-4}	5.9×10^{-4}	8.6×10^{-4}
7	16hr	3.4×10^{-4}	4.9×10^{-4}	3.5×10^{-4}	5.0×10^{-4}
8	47hr	2.0×10^{-4}	3.0×10^{-4}	2.0×10^{-4}	3.2×10^{-4}
9	86hr	1.2×10^{-4}	1.9×10^{-4}	1.2×10^{-4}	1.9×10^{-4}
10	199hr	7.6×10^{-5}	1.2×10^{-4}	6.6×10^{-5}	1.2×10^{-4}
11	422hr	4.6×10^{-5}	7.8×10^{-5}	4.3×10^{-5}	7.5×10^{-5}
12	960hr	2.8×10^{-5}	4.8×10^{-5}	2.2×10^{-5}	4.6×10^{-5}
13	1590hr	1.7×10^{-5}	3.2×10^{-5}	1.5×10^{-5}	2.9×10^{-5}
14	-	-	-	7.4×10^{-6}	1.7×10^{-5}
15	-	-	-	5.2×10^{-6}	1.1×10^{-5}

Table 2. Comparison of Autotune and Polyestimate for 10% measurement error and depolarizing noise of rate $p = 10^{-3}$ on all other gates. Column 1 contains the code distance d . Column 2 contains the total CPU time required by Autotune to observe 10^4 logical errors enabling accurate estimation of the logical X and Z error rates shown in columns 3 and 4. Columns 5 and 6 contain Polyestimate generated logical X and Z error rates. Agreement is in all cases within 15%, sufficient for practical purposes.

d	T	p_{X_L}	$p_{X_L}^{\max}(d)$	$p_{X_L}^{\text{av}}(d)$
3	10min	1.1×10^{-3}	4.0×10^{-3}	1.4×10^{-3}
4	47min	4.6×10^{-4}	2.5×10^{-3}	6.8×10^{-4}
5	6.8hr	9.7×10^{-5}	1.4×10^{-3}	2.1×10^{-4}
6	33hr	3.1×10^{-5}	8.7×10^{-4}	8.6×10^{-5}
7	199hr	7.9×10^{-6}	4.7×10^{-4}	3.0×10^{-5}
8	986hr	2.3×10^{-6}	3.0×10^{-4}	1.1×10^{-5}

Table 3. Comparison of Autotune and Polyestimate for the case of all gates having total probability of error $p = 10^{-3}$, however the CNOT gate having an asymmetric error model satisfying $p'_{IX} = 10p'_{XI} = 100p'_{XX}$. Column 1 contains the code distance d . Column 2 contains the total CPU time required by Autotune to observe 10^4 logical errors enabling accurate estimation of the logical X error rates shown in column 3. Column 4 contains Polyestimate generated logical X error rates when raising the low error rates to match the maximum. Agreement is poor, with Polyestimate significantly overestimating the logical error rate due significantly raising the probability of XI and XX errors. Column 5 instead takes the average error rate, which still leads to an overestimate, but one sufficiently close to simulation to be practically useful.

A simple alternative to raising all error probabilities to match the maximum when the underlying error model is highly asymmetric is to instead work with the average error probability. Polyestimate includes a switch to use averaging rather than the maximum. In many ways, it could be argued that averaging is always a more reasonable approach, however, when the asymmetry is low, using the maximum still gives good estimates with the additional desirable property that logical error rates are guaranteed to not be underestimated. Averaging intuitively should also lead to an upper bound on the logical error rate as the total error is still the same however the asymmetry in the model has not been exploited to achieve further logical error suppression the way Autotune does. This argument is not, however, rigorous.

The last column of Table 3 shows Polyestimate data with averaging. The obtained logical error rates are less than a factor of 4 off for odd $d < 9$, which is sufficient to obtain usefully accurate overhead estimates. Note that the level of asymmetry considered here is so extreme that it is highly unlikely to be found in a physical system, yet the logical error rate estimates are still acceptable for practical purposes.

6 Conclusion

We have described an open source tool, Polyestimate, that is shipped with a database of Autotune simulation results sufficiently broad and detailed to provide, through interpolation and extrapolation, near-instantaneous surface code logical error rates for a very broad range of practical depolarizing error rates and arbitrary code distances d . Individual depolarizing error rates for initialization, measurement, Hadamard, CNOT, and identity gates of duration each of these gates, respectively, can be specified and handled accurately. The primary limitation of Polyestimate is its inability to cope with very asymmetric error models on the CNOT within a single type of error. In future work, we plan address this limitation with slower analytic techniques for such error models that will still be orders of magnitude faster than running Autotune.

Acknowledgements

This research was conducted by the Australian Research Council Centre of Excellence for Quantum Computation and Communication Technology (project number CE110001027), with support from the US National Security Agency and the US Army Research Office under contract number W911NF-13-1-0024. Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D11PC20166. 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.

References

1. E. Knill, R. Laflamme, and W. H. Zurek, *Accuracy Threshold for Quantum Computation*, Tech. Rep. LAUR-96-2199 (Los Alamos National Laboratory, 1996) quant-ph/9610011.
2. D. Aharonov and M. Ben-Or, Proc. ACM STOC **29**, 176 (1997), quant-ph/9611025.
3. A. Y. Kitaev, Russ. Math. Surv. **52**, 1191 (1997).
4. A. G. Fowler, Phys. Rev. Lett. **109**, 180502 (2012), arXiv:1206.0800.

5. S. B. Bravyi and A. Y. Kitaev, quant-ph/9811052 (1998).
6. E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. **43**, 4452 (2002), quant-ph/0110143.
7. R. Raussendorf and J. Harrington, Phys. Rev. Lett. **98**, 190504 (2007), quant-ph/0610082.
8. R. Raussendorf, J. Harrington, and K. Goyal, New J. Phys. **9**, 199 (2007), quant-ph/0703143.
9. A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012a), arXiv:1208.0928.
10. D. S. Wang, A. G. Fowler, and L. C. L. Hollenberg, Phys. Rev. A **83**, 020302(R) (2011), arXiv:1009.3686.
11. A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Phys. Rev. Lett. **108**, 180501 (2012b), arXiv:1110.5133.
12. A. G. Fowler, A. C. Whiteside, A. L. McInnes, and A. Rabbani, Phys. Rev. X **2**, 041003 (2012c), arXiv:1202.6111, <http://topqec.com.au/autotune.html>.
13. D. Gottesman, *Stabilizer Codes and Quantum Error Correction*, Ph.D. thesis, Caltech (1997), quant-ph/9705052.
14. A. G. Fowler, Phys. Rev. A **87**, 040301(R) (2013a), arXiv:1208.1334.
15. J. Edmonds, Canad. J. Math. **17**, 449 (1965a).
16. J. Edmonds, J. Res. Nat. Bur. Standards **69B**, 125 (1965b).
17. A. G. Fowler, A. C. Whiteside, and L. C. L. Hollenberg, Phys. Rev. A **86**, 042313 (2012d), arXiv:1202.5602.
18. G. Duclos-Cianci and D. Poulin, Quant. Inf. Comput. **14**, 0721 (2014), arXiv:1304.6100.
19. S. Bravyi and J. Haah, Phys. Rev. Lett. **111**, 200501 (2013), arXiv:1112.3252.
20. J. R. Wootton and D. Loss, Phys. Rev. Lett. **109**, 160503 (2012), arXiv:1202.4316.
21. A. Hutter, J. R. Wootton, and D. Loss, Phys. Rev. A **89**, 022326 (2014), arXiv:1302.2669.
22. J. R. Wootton, Phys. Rev. A **88**, 062312 (2013), arXiv:1310.2393.
23. P. Sarvepalli and R. Raussendorf, Phys. Rev. A **85**, 022317 (2012), arXiv:1111.0831.
24. A. G. Fowler, arXiv:1310.0863 (2013b).
25. Y. Chen, D. Sank, P. O'Malley, T. White, R. Barends, B. Chiaro, J. Kelly, E. Lucero, M. Mariantoni, A. Megrant, C. Neill, A. Vainsencher, J. Wenner, Y. Yin, A. N. Cleland, and J. M. Martinis, .