Optimal number of stabilizer measurement rounds in an idling surface code patch

Áron Márton[1] and János K. Asbóth[1,2,3]

[1]Department of Theoretical Physics, Institute of Physics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary

[2]HUN-REN Wigner Research Centre for Physics, H-1525 Budapest, P.O. Box 49., Hungary

[3]HUN-REN-BME-BCE Quantum Technology Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary

arXiv:2408.07529v4 [quant-ph] 30 May 2025

Logical qubits can be protected against environmental noise by encoding them into a highly entangled state of many physical qubits and actively intervening in the dynamics with stabilizer measurements. In this work, we numerically optimize the rate of these interventions: the number of stabilizer measurement rounds for a logical qubit encoded in a surface code patch and idling for a given time. We model the environmental noise on the circuit level, including gate errors, readout errors, amplitude and phase damping. We find, qualitatively, that the optimal number of stabilizer measurement rounds is getting smaller for better qubits and getting larger for better gates or larger code sizes. We discuss the implications of our results to some of the leading architectures, superconducting qubits, and neutral atoms.

1 Introduction

Quantum error-correcting codes are necessary for useful quantum algorithms to be run on noisy hardware. One of the most promising candidates among such codes is the surface code [1, 2], due to its planar connectivity and high threshold [3, 4, 2]. Quantum memory experiments with surface codes have been performed on various architectures [5, 6, 7, 8, 9, 10, 11], and logical operations have also been demonstrated recently [12, 13, 14].

The performance of the surface code is routinely investigated under static noise models, i.e., errors occurring with fixed per-cycle probability. This class of errors includes random Pauli and readout errors, which are modeled both on phenomenological [4] and circuit level [2], coherent errors (small unitary rotations) were also investigated both theoretically [15, 16] and numerically [17, 18, 19].

In a real quantum computer, errors are dynamic: the per-cycle error rate depends on the time between quantum error correction rounds. This raises an optimization problem for a logical qubit based on the surface code: What is the optimal number of stabilizer measurement rounds to get the best performance from a logical qubit over a fixed amount of time?

In this work, we answer the question of optimizing the number of stabilizer measurement rounds by numerically simulating a circuit-level noise model including static errors (gate and readout errors) and dynamic errors (amplitude and phase damping) as well. We investigate how the optimal number depends on different noise parameters, and on the code distance. We find that, as could be expected, for higher quality physical qubits the optimal number of rounds is lower, while for higher quality gates the optimal number of rounds is higher. Interestingly, the optimal number of rounds is also higher for larger code sizes. We also discuss the practical implications of our results to quantum error correction experiments with superconducting qubits and neutral atoms. Broadly speaking, for superconducting qubits performing as many stabilizer measurement rounds as possible is beneficial; however, for neutral atoms, much fewer measurement rounds may offer optimal performance.

The rest of this paper is structured as follows. In Sec. 2 we introduce the basic concepts of the surface code and discuss the logical noise channel corresponding to the logical qubit encoded into a surface code patch. In Sec. 3 we describe our noise model in detail. Finally, in Sec. 4 we discuss our results on the optimal number of stabilizer measurement rounds and explain how it depends on the noise parameters and the code distance. We also discuss the implications of our work for superconducting and neutral atom devices. In Sec. 5 we summarize our results and discuss possible extensions of our work.

2 The logical noise channel of the surface code

In a rotated surface code patch[20] data qubits are on the vertices of a square grid, while ancilla qubits in the middle of the faces, as shown in Fig. 1. A logical qubit is encoded in the collective quantum state of the data qubits, more precisely into the so-called logical subspace: the $+1$ eigensubspace of the X- and Z-

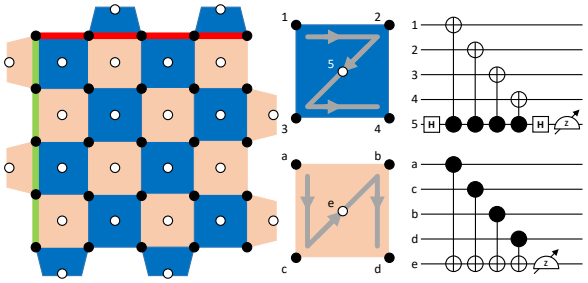

Figure 1: A rotated surface code patch with distance $d = 5$. Data (ancilla) qubits are black (white) circles, on the vertices (faces). The X- (Z-) stabilizers are dark (light) faces; Quantum circuits for their measurements are shown. Logical Z (X) operator is a red (green) line at the top (left side).

stabilizers, called A_f and B_f,

$$A_f = \prod_{j\in\partial f} X_j; \qquad B_f = \prod_{j\in\partial f} Z_j. \tag{1}$$

Stabilizers are X and Z parity checks on the data qubits at the corners of the corresponding faces; their measurement can detect local Pauli errors. These measurements are realized via a syndrome extraction circuit which entangles data qubits with ancillas, as in Fig. 1.

Logical operators of the surface code, also shown in Fig. 1, can be defined as

$$Z_L = \prod_{j\in\mathrm{TOP}} Z_j; \qquad X_L = \prod_{j\in\mathrm{LEFT}} X_j. \tag{2}$$

These commute with all the stabilizers and act as Pauli operators in the logical subspace. Note that the above definition is not unique: equivalently valid logical operators can be obtained by multiplying the Z_L and X_L defined above with any stabilizers, since this does not change how they act in the logical subspace. The code distance d is defined as the minimum weight of logical operators, which is just the linear size of the rotated surface code patch. For Fig. 1, the distance is $d = 5$.

2.1 Characterizing the noisy logical identity channel

The noisy logical identity channel describes what happens to an idling logical qubit, if errors affect the physical qubits of the error correcting code. The error model will be detailed in the next Section. In the case of incoherent physical errors the channel reads,

$$\varepsilon(\rho) = (1-p_L)\rho + p_x^L X_L\rho X_L + p_y^L Y_L\rho Y_L + p_z^L Z_L\rho Z_L, \tag{3}$$

with the parameters $p^L_{x,y,z}$ describing the probabilities of Pauli X, Y, Z errors. The total *logical failure rate* is p_L,

$$p_L = p_x^L + p_y^L + p_z^L. \tag{4}$$

Characterization of the noisy logical identity channel is the (numerical) computation of the parameters $p^L_{x,y,z}$, which depend on both the details of the quantum error correcting code and the errors affecting the physical qubits.

We fully characterize the noisy logical identity channel of a logical qubit encoded in a single surface code patch under circuit-level Pauli noise. We perform efficient Clifford simulations with the Python package STIM [21] of ideal memory experiments. The numerical data is available at [22]. Our approach, as shown in Fig. 2, is as follows. We first ideally (with a circuit without noise) initialize the patch in one of the logical Clifford states, $|0\rangle_L, |+\rangle_L$ or $|i\rangle_L$, then measure the stabilizers through noisy syndrome extraction circuits, N times. At the end we perform a perfect stabilizer measurement round to ensure the quantum state is left in the logical subspace up to a correction operator (see next paragraph). Finally, we measure all of the data qubits in the appropriate basis (in accordance with the initialized state) and extract the eigenvalue of the appropriate logical Pauli operator, Z_L, X_L or Y_L, which is sufficient to determine the logical failure rate (see Appendix D). By performing the noisy stabilizer measurement rounds with the quantum circuits shown in Fig. 1, the fault distance of the ideal memory experiments is uniformly d

For the decoding, i.e., to decide a correction operator based on the stabilizer measurement outcomes, we used the minimum weight perfect matching approach [3, 23] via PyMatching [24, 25]. We note that other decoders may offer better performance than minimum weight perfect matching [26, 27, 28, 29], but we expect qualitatively similar results with them. We used STIM to decompose the detector error model into weighted X and Z matching graphs, where the weights display the probabilities in the circuit-level error model.

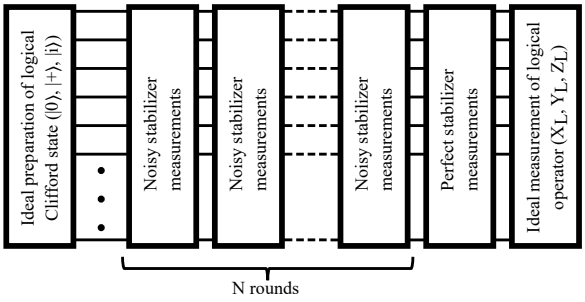

Figure 2: The sketch of quantum circuits realizing ideal X, Y and Z memory experiments with a logical qubit encoded into a surface code patch.

Note that we do not model in detail the preparation of the initial state – this is not relevant for our purpose. In principle, the preparation of logical Clifford states and the measurement of logical Pauli operators can be done fault tolerantly [30]. However, we would like to characterize the logical identity chan-

nel itself and not the noisy memory experiments, so we exclude noise in the initialization and logical measurement processes. In other words, our goal in this work is to find the optimal strategy for protecting logical quantum information encoded in idling surface code patches while some other patches may undergo logical operations by any of the proposed approaches [31, 32, 33, 34, 35, 36].

3 Error model

We consider three types of errors during quantum error correction rounds, to give a rough description of noise in current quantum devices [5, 12]: *idling errors*, describing the dissipative dynamics of data qubits; *gate errors*, modeling the noisy quantum gates; and *readout errors*, describing imperfect measurements. Among the error types we neglect are leakage errors, correlated errors (these two can be suppressed to some degree with various physical-level solutions [37, 38, 39, 40]), and crosstalk [41, 42] (which we plan to investigate in follow-up work). We also neglect coherence in the errors by considering only random Pauli error channels. This approximation (obtained, e.g., using Pauli twirling) is necessary for the large scale numerical simulations[43, 44]. Neglecting coherence in the physical errors is generally believed to be a reasonable approximation, as supported by analytical[16, 15], numerical [17, 18, 19], and experimental [5] work.

Idling errors describe the dissipative dynamics of qubits. They are here modeled based on the amplitude and phase damping channel [45], characterized by two time scales T_1 and T_ϕ: the relaxation and pure dephasing times. The corresponding Lindblad equation in the rotating frame reads:

$$\frac{\partial \rho}{\partial t} = \frac{1}{T_1}\Big(|0\rangle\langle 1|\,\rho\,|1\rangle\langle 0| - \frac{1}{2}\{|1\rangle\langle 1|\,,\rho\}\Big) + \frac{1}{T_\phi}\Big(|1\rangle\langle 1|\,\rho\,|1\rangle\langle 1| - \frac{1}{2}\{|1\rangle\langle 1|\,,\rho\}\Big). \quad (5)$$

After solving Eq. 5 and applying the Pauli twirl approximation we get

$$\rho(t) = p_0(t)\rho_0 + p_x(t)X\rho_0 X + p_y(t)Y\rho_0 Y + p_z(t)Z\rho_0 Z, \quad (6)$$

with time dependent Pauli error probabilities,

$$p_x(t) = p_y(t) = \frac{1}{4}(1 - e^{-t/T_1}) \quad (7)$$

$$p_z(t) = \frac{1}{2}(1 - e^{-t/T_2}) - \frac{1}{4}(1 - e^{-t/T_1}) \quad (8)$$

$$p_0(t) = 1 - p_x(t) - p_y(t) - p_z(t). \quad (9)$$

Here $1/T_2 = 1/2T_1 + 1/T_\phi$ is the rate of dephasing [46].

In the numerical simulations we apply idling errors via Eq. 6 on the data qubits, during the time when the ancilla qubits are being measured and reinitialized, see Fig. 3. We approximate this time as T/N in each cycle, where T is the total time of the of memory experiment, and N is the number of stabilizer measurement rounds. Our basic assumption is that the measurement time is much longer than gate times, which is realistic in current devices [5, 12], thus we neglect the dissipative dynamics of qubits during gate operations.

Gate errors are errors occuring during imperfect quantum gates, whose microscopic model depends very much on the implementation of the gate. We thus use a generic approach: a single-qubit depolarizing channel after each single qubit quantum gate, and a two-qubit depolarizing channel after each two-qubit quantum gate. These channels read:

$$\mathrm{Dep}_1(\rho) = (1-p)\rho + \frac{p}{3}\sum_{P\in\{X,Y,Z\}} P\rho P; \quad (10)$$

$$\mathrm{Dep}_2(\rho) = (1-p)\rho + \frac{p}{15}\sum_{P\in\{I,X,Y,Z\}^{\otimes 2}/\{II\}} P\rho P, \quad (11)$$

where I is the single qubit identity operator.

Readout errors model faulty measurements and faulty resets, which we treat on the phenomenological level: ideal measurementswith possibly erroneously recorded results. The measurement outcome m can suffer a classical bit-flip during the readout, therefore, we record the wrong measurement outcome with probability q,

$$P(m = 1 \to m = -1) = P(m = -1 \to m = 1) = q, \quad (12)$$

and correspondingly $P(m = 1 \to m = 1) = P(m = -1 \to m = -1) = 1 - q$. In Fig. 3 a segment of the noisy syndrome extraction circuit is shown, displaying all three error mechanisms.

4 Optimal number of stabilizer measurement rounds

We ran extensive numerical simulations to characterize the logical identity channel for a broad spectrum of parameters: idling error time scales T_1, T_ϕ, gate error parameter p, readout error probability q, and code distance d. We identified an optimal number of measurement rounds N^*, for each set of parameters, where the logical failure rate is minimal. In this Section we give a qualitative explanation, and the numerical results, of how N^* depends on different parameters. Finally, we discuss the implications of our results for two promising experimental platforms: superconducting qubits and neutral atoms.

We start with a qualitative explanation why one can expect that there is an optimal number of stabilizer measurement rounds, N^*, to be performed during a fixed time T, to minimize the logical failure rate

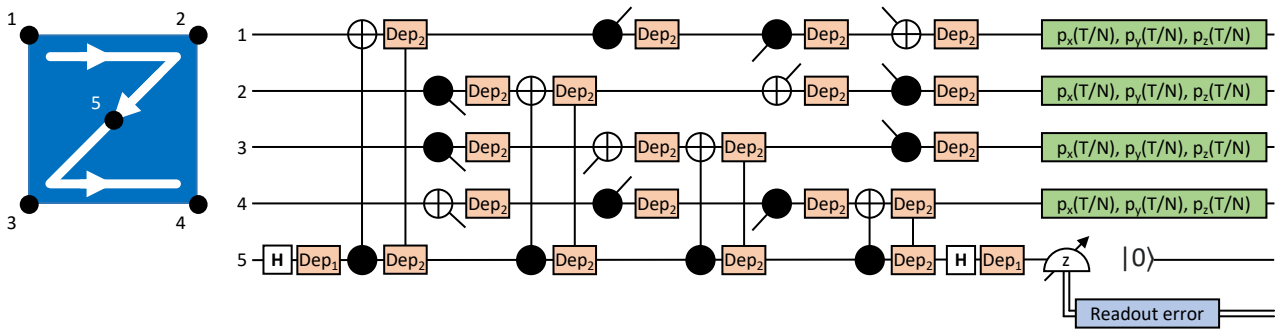

Figure 3: Quantum circuit for the noisy measurement of an X-stabilizer. Some CNOT gates are only partly drawn: these are parts of other stabilizer measurements, each free end points towards the ancilla qubit on which the other half of the CNOT acts. Idling errors are green rectangles acting on data qubits during the measurement and the reset of the ancilla. Readout error is a classical probabilistic bit-flip. Gate errors are orange rectangles.

p_L. With few measurement rounds, idling errors dominate the error budget, thus, decreasing N increases the chance that idling errors accumulate to cause an unobserved logical error. With many measurement rounds, gate and readout errors dominate the error budget, thus, increasing N increases the chance of a logical error. Examples of curves showing this behaviour are depicted in Fig. 4.

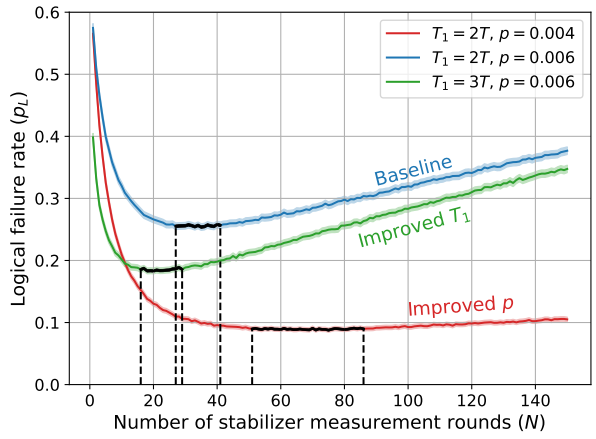

Figure 4: Numerically obtained logical failure rates, $p_L = p_x^L + p_y^L + p_z^L$, for three different sets of parameters, against the number of measurement rounds, N. The pure dephasing time ($T_\phi = 12T$), the readout error rate ($q = 0.02$), and the code distance ($d = 9$) were fixed, while the relaxation time (T_1) and the gate error rate (p) differ for the three curves. The broad minima (minimum to within Δp_L) are highlighted in black. At each stabilizer measurement round value, we ran ideal X, Y, and Z memory experiments with 100.000 execution rounds for each. The error bars are included as pale contours around the curves with width of $6\Delta p_L$.

It is quite intuitive to understand qualitatively how the optimal number of stabilizer measurement rounds, N^*, depends on the noise parameters. For higher quality qubits (longer T_1 and T_ϕ times) N^* is smaller, because fewer noisy rounds are sufficient to suppress the effect of idling errors. Conversely, for more precise gates and measurements (smaller p and q), N^* is larger, because each stabilizer measurement round introduces less noise, making it beneficial to perform more rounds. Numerical simulations confirm these intuitive expectations as shown in Fig. 6. We also find, as in Fig. 4, that the logical failure rate curves are flat around the minima. This results in a large uncertainty in N^*, a roughly identical performance in a wide range of N.

We find that the optimal number of stabilizer measurement rounds also depends on the code distance: N^* increases as the surface code patch is scaled up, see Fig. 5. To understand this behaviour we approximate the probability of logical X and Z errors with the contributions from the minimum weight error strings,

$$p_{x/z}^L = AN\Big(p_{x/z}(T/N) + p_y(T/N) + kp\Big)^{(d+1)/2}. \tag{13}$$

Here A is some function of the code distance d, but independent of N, and k is independent of both d and N. The factor AN is the multiplicity of error strings of weight $(d+1)/2$, composed of data qubit errors. These errors originate from idling errors with probabilities $p_x + p_y$ (for logical X errors) or $p_z + p_y$ (for logical Z errors), and from gate errors with probability kp. By counting the possible error events (see Appendix A) we estimate k as $56/15$. In Eq, (13) we neglect ancilla qubit errors (originating from readout errors and gate errors affecting ancilla qubits), because these usually do not contribute to minimum weight error strings. A more detailed explanation can be found in Appendix A.

By approximating the idling error rates as

$$p_{x/z}(T/N) + p_y(T/N) \approx \frac{T}{2T_{1/2}N}, \tag{14}$$

and differentiating p_x^L and p_z^L separately we derive the optimal numbers of stabilizer measurement rounds for

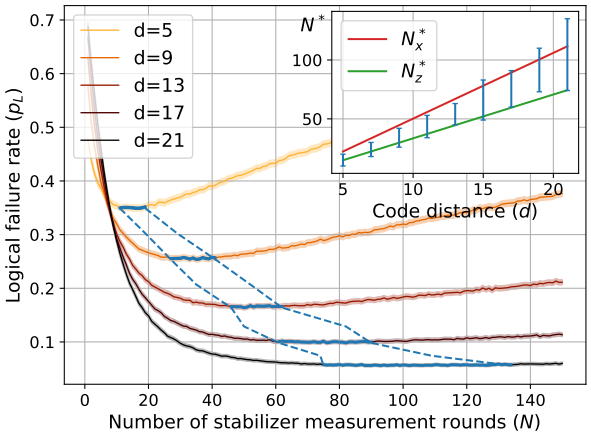

Figure 5: Logical failure rate, $p_L = p_x^L + p_y^L + p_z^L$, as the function of the number of stabilizer measurement rounds for code sizes $d = 5$ to 21. The optimal intervals are colored blue. In the inset, these optimal intervals (blue) can be seen as the function of the code distance. Our estimations in Eq. (15) for the optimal number of rounds during memory experiments (N_x^* and N_z^*) are shown as red and green curves. The noise parameters were $T_1 = 2T$, $T_\phi = 12T$,$p = 0.006$ and $q = 0.02$.The error bars are included as pale contours around the curves with width of $6\Delta p_L$.

Z and X memory experiments,

$$N_{x/z}^* = \frac{(d-1)T}{4kpT_{1/2}}. \tag{15}$$

We have plotted N_x^* and N_z^* as the function of d in Fig. 5 for the parameter set $T_1 = 2T$, $T_\phi = 12T$, $p = 0.006$, $q = 0.02$, and $k = 56/15$.

The optimal number of stabilizer measurement rounds N^* lays between N_x^* and N_z^*, and increases with d, because both $N_x^*(d)$ and $N_z^*(d)$ are monotonically increasing functions of d. Fig. 5 confirms this behaviour and validates Eq. (15). A more detailed analysis of the d dependence of the optimal number can be found in Appendix C. In addition, Eq. (15) verifies the intuitive expectations on the T_1, T_ϕ, and p dependence of N^*. To express the effect of readout errors, ancilla qubit errors have to be taken into consideration in Eq. (13) as well.

4.1 Comparison with Current Experiments

In the numerical simulations we assumed that the number of measurement rounds can be increased boundlessly. However, in real devices stabilizer measurements take some time to execute, therefore, it may happen that the optimal number cannot be reached in an experiment.

To compare the theoretically optimal and the experimentally feasible number of stabilizer measurement rounds, we consider two quantum error correction experiments from the last few years, one from Google Quantum AI with superconducting qubits [5] and one from Harvard/QuEra with neutral atoms [12]. The error parameters and the execution times of stabilizer measurement cycles of these experiments are summarized in Tab.1. In our previous choice of

	Superconducting qubits [5]	Neutral atoms [12]
T_1	20 µs	4 s
T_ϕ	120 µs	1.85 s
p	0.006	0.006
q	0.02	0.002
Cycle time	$\sim$ 900 ns	$\sim$ 2 ms

Table 1: Error parameters and quantum error correction cycle times for superconducting and neutral atom devices. The values are taken from [5, 12, 7].

parameters in Fig. 6, the relaxation-dominated family of parameter sets ($T_1 = 1 - 3T, T_\phi = 12T, p = 0.004 - 0.008, q = 0 - 0.04$) corresponds to superconducting qubits for experiment time $T = 10$ µs, while the dephasing-dominated family of parameter sets ($T_1 = 4T, T_\phi = 1.2T - 2T, p = 0.004 - 0.008, q = 0.002 - 0.042$) corresponds to neutral atoms for experiment time $T = 1$ s.

While the optimal numbers of stabilizer measurement rounds are of the same order of magnitude for both types of devices (see Fig. 6), the experimentally feasible error correction rounds differ significantly. With the superconductor experiment, under 10 µs at most 11 stabilizer measurement rounds can be performed, in contrast to that on the neutral atom based quantum device, under 1 s at most 500 rounds. Consequently, the optimal strategy differs for these two devices: For $d = 9$ (Fig. 6), for superconducting qubits it is beneficial to do the maximum number of measurement rounds with no idling time, however, for neutral atoms, measuring the stabilizers less frequently is suggested. Leaving some idling time in between stabilizer measurements provides optimal performance here.

Although above we discussed the case of a fix|changeed idling time T, we believe our conclusions regarding the usefulness of idling time between stabilizer measurements are valid more generally. If the errors are rare enough it is reasonable to assume that logical failure rate is proportional to the number of stabilizer measurement rounds with fixed error rates, i.e.,

$$p_L = Nf(p_x, p_y, p_z, p, q, d). \tag{16}$$

Therefore, the logical failure rate over unit time, p_L/T, does not explicitly depend on T or N, only on the combination N/T. As a consequence the optimal frequency of interventions N^*/T minimizes this failure rate, meaning that our findings are valid for arbitrary idling time. This is true using the logical failure rate over unit time as the efficiency metric.

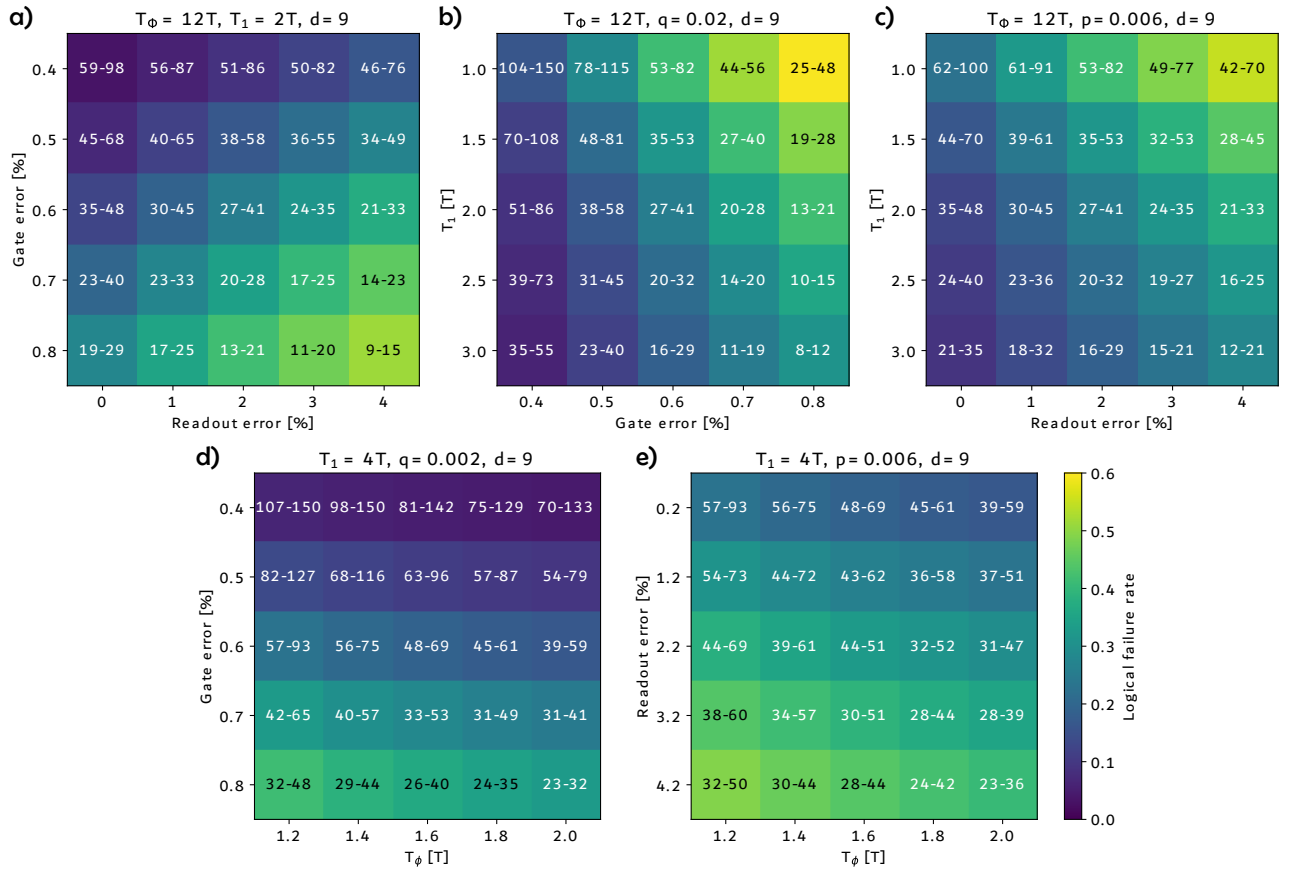

Figure 6: The optimal number of stabilizer measurement rounds and the corresponding logical failure rates for a distance 9 surface code patch over a wide range of physical error parameters. Numbers indicate the borders of the optimal intervals, while the heatmap indicates the logical failure rates. (a),(b), and (c) correspond to the relaxation-dominated parameter sets, conversely, (d) and (e) to the dephasing-dominated parameter sets.

5 Conclusion

In this work we numerically characterized the logical noise channel of an idling surface code patch across various code distances, number of stabilizer measurement rounds, and noise parameters. For each set of noise parameters and code distances, we determined an optimal number for stabilizer measurement rounds, N^*, that maximizes the performance of the surface code patch during a fixed idling time. Our circuit-level noise model incorporated both time-dependent dynamic errors, amplitude and phase damping of data qubits, and time-independent static gate and readout errors.

Our results indicate that N^* decreases with higher qubit quality, while it increases with higher gate quality and larger code sizes. We also explored the applicability of our findings to superconducting and neutral atom devices, revealing distinct optimal strategies for these two architectures. Specifically, for superconducting qubits, the optimal strategy involves performing the maximum feasible number of stabilizer measurement rounds, whereas for neutral atoms, the optimal number is significantly lower than the maximum feasible.

Including more sophisticated errors, such as leakage [37], crosstalk [41, 42], or correlated errors [40] would be an interesting extension of our work, as would considering more parameters to optimize on (e.g., code, code distance, reset after measurements) to get the best performance, as in [47, 48].

We believe our results will be valuable for instrumentalists aiming to further optimize the performances of logical qubits encoded into quantum error-correcting codes.

Our data is available at [22].

6 Acknowledgement

This research was supported by the Ministry of Culture and Innovation and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary (Grant No. 2022-2.1.1-NL-2022-00004), and by the Horizon Europe research and innovation programme of the European Union through the HORIZON-CL4-2022-QUANTUM01-SGA project 101113946 OpenSuperQ-Plus100 of the EU Flagship on Quantum Technologies. We acknowledge the support of the Wigner Re-

search Centre for Physics' trainee program. This work was supported by the HUN-REN Hungarian Research Network through the Supported Research Groups Programme, HUN-REN-BME-BCE Quantum Technology Research Group (TKCS-2024/34).

References

[1] A.Yu. Kitaev. "Fault-tolerant quantum computation by anyons". Annals of Physics **303**, 2–30 (2003).

[2] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. "Surface codes: Towards practical large-scale quantum computation". Physical Review A**86** (2012).

[3] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. "Topological quantum memory". Journal of Mathematical Physics **43**, 4452–4505 (2002).

[4] Chenyang Wang, Jim Harrington, and John Preskill. "Confinement-Higgs transition in a disordered gauge theory and the accuracy threshold for quantum memory". Annals of Physics **303**, 31–58 (2003).

[5] Google Quantum AI. "Suppressing quantum errors by scaling a surface code logical qubit". Nature **614**, 676 – 681 (2022).

[6] Google Quantum AI and Collaborators. "Quantum error correction below the surface code threshold". Nature **638**, 920–926 (2024).

[7] Dolev Bluvstein, Harry Levine, Giulia Semeghini, Tout T. Wang, Sepehr Ebadi, Marcin Kalinowski, Alexander Keesling, Nishad Maskara, Hannes Pichler, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. "A quantum processor based on coherent transport of entangled atom arrays". Nature **604**, 451–456 (2022).

[8] Sebastian Krinner, Nathan Lacroix, Ants Remm, Agustin Di Paolo, Elie Genois, Catherine Leroux, Christoph Hellings, Stefania Lazar, Francois Swiadek, Johannes Herrmann, et al. "Realizing repeated quantum error correction in a distance-three surface code". Nature **605**, 669–674 (2022).

[9] J. F. Marques, B. M. Varbanov, M. S. Moreira, H. Ali, N. Muthusubramanian, C. Zachariadis, F. Battistel, M. Beekman, N. Haider, W. Vlothuizen, A. Bruno, B. M. Terhal, and L. DiCarlo. "Logical-qubit operations in an error-detecting surface code". Nature Physics **18**, 80–86 (2021).

[10] Youwei Zhao et al. "Realization of an error-correcting surface code with superconducting qubits". Physical Review Letters**129** (2022).

[11] M. H. Abobeih, Y. Wang, J. Randall, S. J. H. Loenen, C. E. Bradley, M. Markham, D. J. Twitchen, B. M. Terhal, and T. H. Taminiau. "Fault-tolerant operation of a logical qubit in a diamond quantum processor". Nature **606**, 884–889 (2022).

[12] Dolev Bluvstein et al. "Logical quantum processor based on reconfigurable atom arrays". Nature (2023).

[13] Jiaxuan Zhang, Zhao-Yun Chen, Yun-Jie Wang, Bin-Han Lu, Hai-Feng Zhang, Jia-Ning Li, Peng Duan, Yu-Chun Wu, and Guo-Ping Guo. "Demonstrating a universal logical gate set on a superconducting quantum processor" (2024). arXiv:2405.09035.

[14] Bence Hetényi and James R. Wootton. "Creating entangled logical qubits in the heavy-hex lattice with topological codes". PRX Quantum **5**, 040334 (2024).

[15] Joseph K Iverson and John Preskill. "Coherence in logical quantum channels". New Journal of Physics **22**, 073066 (2020).

[16] Stefanie J. Beale, Joel J. Wallman, Mauricio Gutiérrez, Kenneth R. Brown, and Raymond Laflamme. "Quantum error correction decoheres noise". Physical Review Letters **121**, 190501 (2018).

[17] Sergey Bravyi, Matthias Englbrecht, Robert König, and Nolan Peard. "Correcting coherent errors with surface codes". npj Quantum Information**4** (2018).

[18] Áron Márton and János K. Asbóth. "Coherent errors and readout errors in the surface code". Quantum **7**, 1116 (2023).

[19] Dávid Pataki, Áron Márton, János K. Asbóth, and András Pályi. "Coherent errors in stabilizer codes caused by quasistatic phase damping". Phys. Rev. A **110**, 012417 (2024).

[20] H. Bombin and M. A. Martin-Delgado. "Optimal resources for topological two-dimensional stabilizer codes: Comparative study". Physical Review A**76** (2007).

[21] Craig Gidney. "Stim: a fast stabilizer circuit simulator". Quantum **5**, 497 (2021).

[22] Áron Márton and János K. Asbóth. "Data for "optimal number of stabilizer measurement rounds in an idling surface code patch"". https://doi.org/10.5281/zenodo.15124237 (2025).

[23] Austin G. Fowler. "Minimum weight perfect matching of fault-tolerant topological quantum error correction in average $o(1)$ parallel time" (2014). arXiv:1307.1740.

[24] Oscar Higgott. "Pymatching: A python package for decoding quantum codes with minimum-weight perfect matching". ACM Transactions on Quantum Computing **3**, 1–16 (2022).

[25] Oscar Higgott and Craig Gidney. "Sparse Blossom: correcting a million errors per core second with minimum-weight matching". Quantum **9**, 1600 (2025).

[26] Ben Barber, Kenton M. Barnes, Tomasz Bialas, Okan Buğdaycı, Earl T. Campbell, Neil I. Gillespie, Kauser Johar, Ram Rajan, Adam W. Richardson, Luka Skoric, Canberk Topal, Mark L. Turner, and Abbas B. Ziad. "A real-time, scalable, fast and resource-efficient decoder for a quantum computer". Nature Electronics **8**, 84–91 (2025).

[27] Johannes Bausch et al. "Learning high-accuracy error decoding for quantum processors". Nature **635**, 834–840 (2024).

[28] Oscar Higgott, Thomas C. Bohdanowicz, Aleksander Kubica, Steven T. Flammia, and Earl T. Campbell. "Improved decoding of circuit noise and fragile boundaries of tailored surface codes". Phys. Rev. X **13**, 031007 (2023).

[29] Noah Shutty, Michael Newman, and Benjamin Villalonga. "Efficient near-optimal decoding of the surface code through ensembling" (2024). arXiv:2401.12434.

[30] Craig Gidney. "Inplace access to the surface code y basis". Quantum **8**, 1310 (2024).

[31] Dominic Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter. "Surface code quantum computing by lattice surgery". New Journal of Physics **14**, 123011 (2012).

[32] Benjamin J. Brown, Katharina Laubscher, Markus S. Kesselring, and James R. Wootton. "Poking holes and cutting corners to achieve Clifford gates with the surface code". Physical Review X**7** (2017).

[33] Sergey Bravyi and Alexei Kitaev. "Universal quantum computation with ideal Clifford gates and noisy ancillas". Physical Review A**71** (2005).

[34] Benjamin J. Brown. "A fault-tolerant non-Clifford gate for the surface code in two dimensions". Science Advances **6**, eaay4929 (2020).

[35] Daniel Litinski. "A game of surface codes: Large-scale quantum computing with lattice surgery". Quantum **3**, 128 (2019).

[36] Héctor Bombín, Chris Dawson, Ryan V. Mishmash, Naomi Nickerson, Fernando Pastawski, and Sam Roberts. "Logical blocks for fault-tolerant topological quantum computation". PRX Quantum**4** (2023).

[37] Kevin C. Miao et al. "Overcoming leakage in quantum error correction". Nature Physics **19**, 1780 – 1786 (2022).

[38] Nathan Lacroix et al. "Fast flux-activated leakage reduction for superconducting quantum circuits". Phys. Rev. Lett. **134**, 120601 (2025).

[39] J. F. Marques et al. "All-microwave leakage reduction units for quantum error correction with superconducting transmon qubits". Phys. Rev. Lett. **130**, 250602 (2023).

[40] Matt McEwen et al. "Resisting high-energy impact events through gap engineering in superconducting qubit arrays". Phys. Rev. Lett. **133**, 240601 (2024).

[41] Andreas Ketterer and Thomas Wellens. "Characterizing crosstalk of superconducting transmon processors". Physical Review Applied**20** (2023).

[42] Mohan Sarovar, Timothy Proctor, Kenneth Rudinger, Kevin Young, Erik Nielsen, and Robin Blume-Kohout. "Detecting crosstalk errors in quantum information processors". Quantum **4**, 321 (2020).

[43] Daniel Gottesman. "The Heisenberg representation of quantum computers" (1998). arXiv:quant-ph/9807006.

[44] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". Physical Review A **70**, 052328 (2004).

[45] Michael A. Nielsen and Isaac L. Chuang. "Quantum computation and quantum information: 10th anniversary edition". Cambridge University Press. (2010).

[46] Yu Tomita and Krysta M. Svore. "Low-distance surface codes under realistic quantum noise". Physical Review A**90** (2014).

[47] Avimita Chatterjee, Debarshi Kundu, and Swaroop Ghosh. "Mits: A quantum sorcerer stone for designing surface codes" (2024). arXiv:2402.11027.

[48] György P Gehér, Marcin Jastrzebski, Earl T Campbell, and Ophelia Crawford. "To reset, or not to reset—that is the question". npj Quantum Information **11**, 39 (2025).

A Minimum weight error strings

We approximated the probabilities of logical X and Z errors with the contributions coming from the minimum weight error strings in Eq. (13). Here, we provide a detailed explanation of this approximation, and also motivate our choice of neglecting error strings that contain ancilla qubit errors.

Ordering the CNOT gates properly (see Fig. 3) does not reduce the code distance at the circuit level. So,

at least d error events are required for an undetected logical error. As a consequence the minimum number of independent error events that can cause a logical error (after correction) is $(d+1)/2$. We call sets of error events that lead to a logical error and consist of $(d+1)/2$ elements, minimum weight error strings.

There are two types of minimum weight error strings: the first one contains error events only on data qubits, while the second one contains one error event on an ancilla qubit and $(d-1)/2$ error events on data qubits. The minimum weight error strings of the second type only count as minimum weight error strings (they only get corrected "wrongly") if the correction has a smaller weight than the error, which happens if the error events on the ancilla qubits have a smaller probability than on the data qubits. For large enough measurement rounds the error events on the ancilla qubits are more likely, therefore, we neglected the contribution of the minimum weight error strings of the second type. However, a more detailed analysis should contain the contribution of these error strings as well.

The error events on data qubits originate from gate or idling errors, therefore, the probabilities can be written in a form which we already used in Eq. (13),

$$p_{data}^{x/z} = p_{x/z}(T/N) + p_y(T/N) + kp + ..., \qquad (17)$$

here we neglected higher order terms.

The multiplicity of the minimum weight error strings is denoted as AN in Eq. 13. Here we considered this multiplicity to be proportional to N, for which we used the rarity of error strings as in Eq. (16).

B Counting error events

To estimate k (the weight of gate errors in physical error probability) we approximately counted the possible error events that can occur during one stabilizer measurement round. A data qubit participates in 4 CNOT operations which can cause data qubit error events. Also, gate errors on ancilla qubits can propagate to a data qubit from the other 3 CNOT gates. Therefore, the total number of data qubit error events is 7 in both the X and Z cases. Possible error events are shown in Fig. 7. The probability of an X or Z data qubit error event, caused by a gate error, is $8p/15$, because 8 two-qubit Pauli errors consist an X or Z error on a given (first or second) qubit. Therefore,

$$k = 7\frac{8}{15} = \frac{56}{15}. \qquad (18)$$

C The d dependence of N^*

Analytical expressions can be derived for the optimal number of stabilizer measurement rounds for X and Z memory experiments separately; see Eq. (15). These

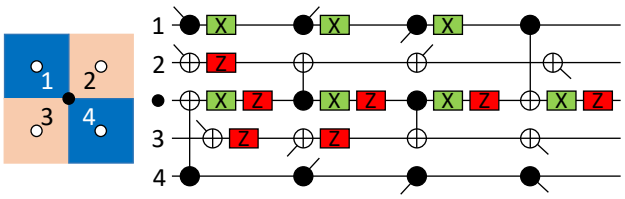

Figure 7: Possible gate error events during a stabilizer measurement round causing an error on a specific data qubit. A segment of the code patch is depicted with a single data qubit (black) and 4 ancilla qubits (white). The relevant part of the syndrome extraction circuit with the possible X (green) and Z (red) error events that cause an error on the data qubit at the end of the given round.

expressions are derived from the minimum weight error string contributions in the logical failure rate, Eq. (13). To see how the theoretically derived expressions of N_x^* and N_z^* match with the numerically determined optimal intervals of X and Z memory experiments, we have plotted these optimal intervals separately for X and Z memory experiments in Fig. 8.

Considering the minimum weight contributions in p_x^L and p_z^L, as in Eq. (13), not only offers analytical expressions for N_x^* and N_z^*, but also gives a theoretical prediction for N^*, the minimum of $p_L = p_x^L + p_z^L + ...$ We numerically determined the minima of the logical failure rate (contains only minimum weight error string contributions) for the relevant set of parameters and plotted it in Fig. 9, together with N_x^*, N_z^* and the optimal intervals. In Fig. 9 the line of N^* lays between N_x^* and N_z^* and matches with the numerically determined optimal intervals, as mentioned in the main text.

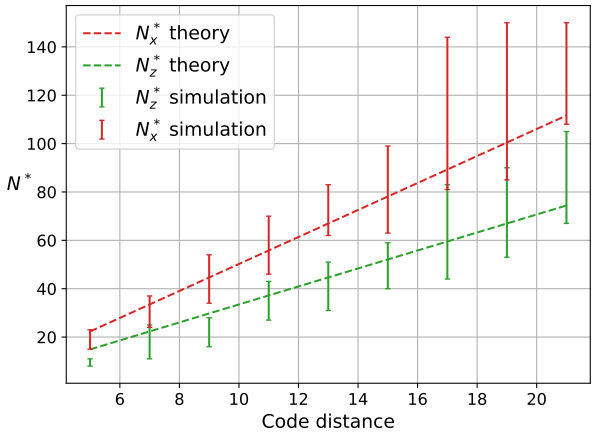

Figure 8: The optimal intervals and the theoretical predictions for the optimal number of measurement rounds separately for X and Z memory experiments as the functions of the code distance.

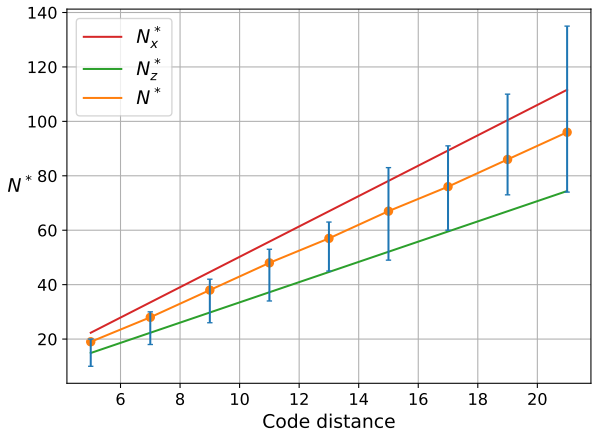

Figure 9: The optimal intervals (blue) and theoretical predictions for the optimal number of measurement rounds as the functions of the code distance. N_x^* and N_z^* are the optimal numbers for the X and Z memory experiments, as described in Eq. (15). N^* is the numerically determined minimum of $p_L = p_x^L + p_z^L + ...$ composed of minimum weight error string contributions as in Eq. (13).

D Determining the logical error rate

To fully characterize the noisy logical identity channel, we simulated ideal memory experiments as discussed in Sec. 2. In the absence of noise, the measurement outcome of the appropriate logical Pauli operator at the end of the experiment should be $+1$. However, errors that cause logical failures flip the sign of this final logical measurement. The probabilities of obtaining a -1 measurement outcome for the logical Pauli operators can be expressed in terms of the parameters of the noisy logical identity channel:

$$\begin{aligned} P(\langle X_L\rangle = -1) &= p_y^L + p_z^L; \\ P(\langle Y_L\rangle = -1) &= p_x^L + p_z^L; \\ P(\langle Z_L\rangle = -1) &= p_x^L + p_y^L. \end{aligned} \tag{19}$$

Therefore, the logical failure rate, as defined in Eq. (4), can be calculated as:

$$p_L = \frac{P(\langle X_L\rangle = -1) + P(\langle Y_L\rangle = -1) + P(\langle Z_L\rangle = -1)}{2}. \tag{20}$$

We determined the uncertainty in the logical failure rate,

$$\Delta p_L = \sum_{\mathcal{P}\in X_L, Y_L, Z_L} \frac{\Delta P(\langle \mathcal{P}\rangle = -1)}{2} \tag{21}$$

$$= \sum_{\mathcal{P}\in X_L, Y_L, Z_L} \sqrt{\frac{P(\langle \mathcal{P}\rangle = -1) P(\langle \mathcal{P}\rangle = 1)}{4N}}, \tag{22}$$

assuming that the number of failures in each memory experiment follows a binomial distribution. The total number of execution rounds is N for each memory experiment.