Finite temperature thresholds for stabilizer codes with few measurements

C. Daniel Freeman,[1,2,*] Mohan Sarovar,[3,†] C. M. Herdman,[4,5,6] and K. B. Whaley[1,7]

[1] *Berkeley Quantum Information & Computation Center, University of California, Berkeley, CA 94720, USA*
[2] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[3] *Sandia National Laboratories, Livermore, CA*
[4] *Institute for Quantum Computing, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[5] *Department of Chemistry, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[6] *Department of Physics & Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[7] *Department of Chemistry, University of California, Berkeley, CA 94720, USA*

(Dated: March 28, 2017)

We demonstrate the existence of a finite temperature threshold for a 1D stabilizer code under an error correcting protocol that uses only a fixed, constant fraction of the syndrome measurements. We sketch how this algorithm generalizes to higher dimensional stabilizer codes with string-like excitations, like the toric code.

I. INTRODUCTION

While an ideal finite temperature quantum memory would require no active error correcting elements, no such systems are known to exist in an experimentally accessible number of dimensions. Practically, the operation of a quantum memory is a balance between passive elements (i.e. dissipative cooling), and active measurement and correction cycles to keep quantum information protected. For near term architectures, stabilizer codes[?] have emerged as the leading candidate for active error correction in quantum hardware, with small scale architectures actively being developed and deployed[?] .

A tremendous amount of effort has gone into developing novel decoding schemes for stabilizer codes. Cite Adrian Hutter, Harrington, Others. More Background.

In previous work, we analyzed the finite temperature dynamics of the toric code, verifying the well-known no-go theorems for the upper bound to the lifetime of the toric code at finite temperature. Using this analysis, we were able to construct a measurement-free protocol for protecting the encoded qubits of the toric code, but these protocols again were limited by the no-go theorems, and only provided a multiplicative constant increase to the lifetime.

In this work, we examine the extent to which a limited amount of measurement can increase the lifetime of stabilizer codes with string-like excitations. In sum, we demonstrate an algorithm that, for any constant density of measurements for a stabilizer code with stringlike excitations undergoing dissipation at a fixed temperature, a threshold can be achieved. The threshold temperature scales with the amount of measurement used—fewer measurements result in a smaller threshold temperature, whereas more complete measurement raises the threshold temperature. This tradeoff is commensurate with and complements what is known about decoding the stabilizer codes in the presence of *noisy*, but complete measurements.

The remainder of the manuscript is structure as follows: Section II briefly reviews the theoretical tools used for performing simulation of stabilizer codes at finite temperature. The content of this section is also expanded upon in refs[? ?] . Section III includes the full description of our few-measurement algorithm, including a discussion of the expected low temperature error processes that cause the algorithm to fail, and a heuristic justification for the expectation of a threshold at low temperature. Section IV details our numerical investigations of our algorithm for the 1D Ising model. Finally, Sec. V sketches how this algorithm could be generalized to higher dimensions, and Sec. VI provides some concluding analysis and discussion.

II. STABILIZER CODES AT FINITE TEMPERATURE

A. Definitions

In this section, we briefly review the theory of the 1D Ising model, as well as the Lindblad formalism for evaluating its finite temperature dynamics. The Hamiltonian for the 1D Ising model is

$$H_{\text{Ising}} = -\Delta\Sigma_i \sigma_z^i \sigma_z^{i+1} \tag{1}$$

where, for the remainder of the manuscript, unless explicitly stated otherwise, we assume $\Delta = 1$. This also fixes our units for temperature. This is exactly the hamiltonian version of the repetition stabilizer *code*[?] . Note that the terms $\sigma_z^i \sigma_z^{i+1}$ exactly correspond to the parity check stabilizer operators of the repetition code. Confer with Table I for the correspondences between the stabilizer hamiltonian and stabilizer code versions of the 1d Ising model.

In the parlance of the 1D Ising model, bit flip errors are often also classified via the dual variables called *domain walls*. Domain walls are simply locations on the 1D Ising chain where a stabilizer operator yields a measurement

Repetition Stabilizer Code	1d Ising Stabilizer Hamiltonian
Encoded States	Ground States
Bit flip errors	Excited States
Error correction via decoder	Same or via cooling

TABLE I. A short summary of the similarities and differences between the Ising model considered as a code versus as a hamiltonian.

of -1—i.e., locations where neighboring spins point in different directions. With periodic boundary, the number of these locations is always a multiple of two, and a single bit flip event either creates a pair of such domain walls, deletes a pair of domain walls, or causes a domain wall to translate by one unit.

Whether considered as a code or as a hamiltonian, the condition for whether a state of the system with errors present can be reliably returned to an encoded state is the same: so long as less than have the system has had errors, a majority rule decoder that has access to measurements of the full set of stabilizers $\sigma_z^i\sigma_z^{i+1}$ will reliably be able to correctly remove errors. When errors are completely independent (i.e., at very high temperature), we can define random variables $x_i = 1$ when an error occurs on site i, and 0 otherwise. If these errors occur with probability p on each spin, independently at random every error detection cycle, then Chernoff's bound gives an upper bound to the probability of an error in the encoded space, $P(\Sigma_i x_i \geq L/2) \leq \exp[-\mathrm{Lp}\frac{\delta^2}{2+\delta}]$ for $\delta = 1/2p - 1$. Thus, for complete measurement, errors in the encoded subspace are exponentially suppressed in system size, so long as the error rate is sufficiently small.

For much of the remainder of the manuscript, we consider how the decoding scheme changes when one does *not* have access to the full set of stabilizer measurements.

Following **??**, we consider a simple local Ohmic, Markovian bath to model finite temperature effects. This amounts to solving the Lindblad equation,

$$\dot{\rho} = \sum_{\omega} 2c_\omega \rho c_\omega^\dagger - c_\omega^\dagger c_\omega \rho - \rho c_\omega^\dagger c_\omega, \tag{2}$$

for the density matrix ρ, with Lindblad operators c_ω chosen to take the form:

$$\{c_\omega\} = \left\{ \sqrt{\gamma(0)}T_b, \sqrt{\gamma(1)}D_b^\dagger, \sqrt{\gamma(-1)}D_b \right\} \tag{3}$$

where D_b dissipates a pair of domain walls, $D_b^\dagger$ creates a pair of domain walls, and where T_b translates a domain wall by one unit, and $\gamma(\cdot)$ is a rate function dependent on the details of the bath. This bath is chosen to model the dynamics of local, single bit-flip errors. In the Pauli basis, these operators take the following form.

$$\begin{aligned} D_b^\dagger =& \frac{1}{4}\left(I\sigma_x I\right)\left(1 + I\sigma_z\sigma_z\right)\left(1 + \sigma_z\sigma_z I\right) \\ D_b =& \frac{1}{4}\left(I\sigma_x I\right)\left(1 - I\sigma_z\sigma_z\right)\left(1 - \sigma_z\sigma_z I\right) \\ T_b =& \frac{1}{4}\left(I\sigma_x I\right)\left(1 - I\sigma_z\sigma_z\right)\left(1 + \sigma_z\sigma_z I\right) \end{aligned} \tag{4}$$

Finally, the remaining details of the bath are specified by the spectral density, which determines the rates with which the different Lindblad operators act:

$$\gamma\left(\omega\right) = \xi \left| \frac{\omega^n}{1 - e^{-\beta\omega}} \right| \tag{5}$$

Choosing $n = 1$ corresponds to an Ohmic spectral density, which is the choice we make for the remainder of the manuscript. With this choice, in the absence of any error correcting protocol, the 1D Ising model has a system size independent thermal error rate given by**?**

$$\Gamma_0 = \frac{\gamma(0)}{1 + e^{1/T}} \tag{6}$$

We define its bare lifetime to be Γ_0^{-1}.

B. Finite Temperature vs. Infinite Temperature

The vast majority of the error correction literature assumes an error model akin to an "infinite temperature limit". More precisely, an array of qubits receives errors from some set of error operators E_i independently at random with some probability p every error correction cycle. The threshold theorems state that there exists some critical error probability p_c below which it is possible to return an error correcting code to its encoded state with unit probability for asymptotically large systems. For the toric code, $p_c \approx .11$.

Contrariwise, thresholds at finite temperature are usually quoted in terms of a critical *temperature*. That is, there must exist some critical temperature T_c below which codes can be reliably corrected. Unfortunately, this definition obscures a great deal of physics—different choices of bath model can greatly affect the dynamics of the error processes, to the extent that a quoted "critical temperature" often implicitly specifies a choice of bath. Because different bath interactions can give rise to different system dynamics, the choice of bath also directly affects the strategy used for error correction. For example, it is known that the toric code's threshold temperature is altered by considering a correlated bath rather than an Ohmic one**?** .

The choice of an Ohmic bath operationally sets the scaling of Lindblad operators that do not exchange energy with the environemnt. While detailed balance

enforces the ratio of the rates of domain wall creation and annihilation to boltzmann-like scaling—i.e., $\gamma(1)/\gamma(-1) = \exp(-1/\mathrm{T})$—the rate of error processes that occur which do not involve energy exchange with the bath scales linearly with temperature. In other words, $\gamma(0) \sim T$. This follows from taking the $\omega \to 0$ limit of Eq. 5. Ultimately, this means that the hopping rate of domains walls is controlled by this choice of bath model. At non-infinite temperatures, this introduces subtle correlations into the patterns of errors that effect the system, so it's no longer possible to talk about something like an "independent error probability per site".

In the most extreme case, at sufficiently low temperature, when errors occur they're most often almost immediately dissipated by the bath. Because domain walls that are not immediately adjacent are not able to be dissipated by the bath, the system occasionally becomes "trapped" in the first excited state by a rare pair creation followed by pair hopping event—i.e. a $D_b^\dagger$ followed by a T_b—before the bath can dissipate the excitation. Then any subsequent errors are overwhelmingly more likely to act on existing pair of domain walls, translating them around in a one-dimensional random walk.

Thus, error correction for the 1D Ising model at low temperature with this sort of bath dynamics is equivalent to attempting to identify these randomly-migrating rare pairs of domain walls. While a majority-rule decoding scheme works in both low and high temperature limits for the Ising model, if the number of measurement resources is restricted, the standard majority rule scheme breaks down.

III. FEW MEASUREMENT ERROR CORRECTION ALGORITHM

A. The Algorithm

In this section, we sketch a new algorithm which reliably dissipates defects in the 1D Ising model below a threshold temperature which we determine numerically. The primary technical innovation of this algorithm is that it does not require measurement of a complete set of stabilizer operators for a given stabilizer code—only a fixed subset. We assume that the system is subject to periodic measurements on periodically spaced measurement "patches", that measurement readout and processing occurs much faster than any system timescale, and that the system is subject to a thermal bath as described in Sec. II A.

Algorithm:

1. Measure stabilizers on patches, keeping record of the age of defects that are already on patches, as well as defect locations.

2. Perform "centering" on patches with defects (see Fig. 1).

3. Calculate probability of fusion (i.e., Eq. 13) for all pairs of measured defects residing on the measurement sites.

4. Probablistically perform error correction based on probabilities calculated in 3.

5. Repeat

Step 2) encourages defects to remain "stuck" on measurement patches. This can be performed entirely unitarily by the $DSWAP$ operator, which takes the following form in the pauli basis,

$$\mathrm{DSWAP} = \frac{1}{2}\left(III + I\sigma_x I + \sigma_z I \sigma_z - \sigma_z\sigma_x\sigma_z\right) \qquad (7)$$

This centering process aids the probability of fusion calculation by ensuring that the coordinates and measurement times are representative of when and where defects are actually created. If defects escape from measurement patches, then upon being measured again, the time recorded by the measurement patch now underestimates how old the defect actually is, biasing the probability estimate. This centering operation greatly reduces the probability of defect escape.

Note the pattern of DSWAPs used in Fig. 1 also swaps the neighboring, *unmeasured* defect onto the measurement patch. This is to ensure that the protocol does not inadvertently create a new separated pair of domain walls in the system by shifting only one domain wall in a potentially adjacent pair.

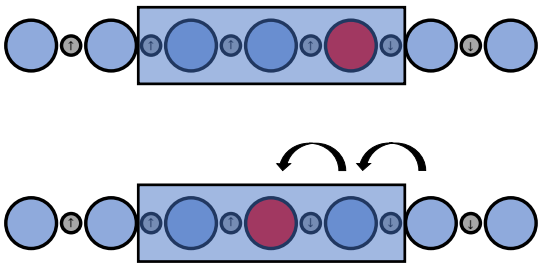

FIG. 1. This cartoon illustrates the "centering" procedure for detected defects. Spin variables are in gray, and domain wall variables are in blue (no defect present) and red (defect present). When a defect is detected on a measurement patch (blue box), it is swapped to the center of the measurement patch via the DSWAP operator (black arrows). The defect immediately adjacent to it is also swapped onto the measurement patch so as not to pull apart defect pairs that would have otherwise dissipated.

B. Fusion Probability Calculation

To properly perform error correction, we need to be able to estimate the probability that two given measured defects are a pair, given that they have been measured

at two particular measurement patches at two different times. For notational convenience, we define:

$$d_1 : d_2 \equiv \text{defect } \mathrm{d}_1 \text{ and } \mathrm{d}_2 \text{ are a pair} \tag{8}$$

and

$$d_i^{x_1,t_1} \equiv \text{defect } \mathrm{d_i} \text{ measured at time } \mathrm{t_1} \text{ at patch } \mathrm{x_1} \tag{9}$$

Then, we aim to calculate:

$$P(d_1 : d_2 | d_1^{x_1,t_1} \wedge d_2^{x_2,t_2}) \tag{10}$$

Because this probability is cumbersome to calculate explicitly, we proceed via Bayes rule:

$$P(d_1 : d_2 | d_1^{x_1,t_1} \wedge d_2^{x_2,t_2}) = \frac{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2) P(d_1 : d_2)}{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2}} \tag{11}$$

The individual terms in the right hand side of equation 11 are straightforward to interpret. $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)$ represents the probability that two measured defects would be at spacetime coordinates (x_1, t_1) and (x_2, t_2) given that they are indeed a pair. This can be related to the probability that a one dimensional diffusion process with diffusion constant D will have performed an excursion of length $|x_2 - x_1|$ in time $|t_2 - t_1|$, and is equal to $1 - \mathrm{erf}(|\mathrm{x_2} - \mathrm{x_1}|/(2\sqrt{\mathrm{D}|\mathrm{t_2} - \mathrm{t_1}|}))$ for a random walk with diffusion constant D. This represents the probability that a random walk will not perform an excursion of distance $|x_2 - x_1|$ in a time $t_2 - t_1$. For our analysis, we will choose $D \propto \gamma_0$. The exact correspondence between D and γ_0 depends on the details of the error correction algorithm itself, so, in practice, we treat the constant of proportionality as an empirically tuned parameter.

$P(d_1 : d_2)$ represents the probability that two measured defects, d_1 and d_2, are in fact a pair. We decompose this into two pieces: a combinatorial piece, and a dynamical piece.

For the combinatorial piece, note that a necessary condition for pairing to be possible is for both defects belonging to a pair to actually be measured. That is, there might be a large number of measured defects $d_1, d_2, ..., d_{n_m}$, but the pairing defect for some of these defects might not be measured. Among those defects which are both measured, and which have their pair also measured, then the probability of selecting two defects that are a pair is simply the combinatorial factor $\frac{1}{\binom{N_{\text{measured pair defects}}}{2}}$ where $N_{\text{measured pair defects}}$ counts the average number of measured defects for which their pair is also measured.

The dynamical piece is the probability that d_1 and d_2 are defects whose pairs are also measured. This probability depends on how quickly defects make excursions to measurement sites, as well as how quickly defects are being paired—either erroneously or correctly—by the protocol. We can crudely lower bound this by taking the equilibrium defect distribution, and calculating the probability that a pair of defects lands on a measurement patch. λ_m/λ_t sites have measurement operators, thus $(\lambda_m/\lambda_t)L\gamma_+$ is an underestimate of the number of defects on measurement patches. This is an underestimate because the protocol is actually more efficient at concentrating defects on measurement patches than equilibrium dynamics is. Given $L\gamma_+$ pairs, this amounts to a binomial counting argument, and the expected number of measured pairs is simply $L\gamma_+(\lambda_m/\lambda_t)^2$. Thus, a lower bound to the equilibrium probability of two selected defects being a measured pair is simply $\frac{L\gamma_+(\lambda_m/\lambda_t)^2}{L\gamma_+} = (\lambda_m/\lambda_t)^2$.

NEED TO DO: Argue why this equilibrium estimate is *still* a lowerbound of the nonequilibrium probability. Intuitively it makes sense, but hard to show math-wise.

Finally, $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2})$ is the probability that two defects are measured, one at (x_1, t_1), and the other at (x_2, t_2). This can be decomposed:

$$\begin{aligned} &P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2}) = \\ &P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2) P(d_1 : d_2) + \\ &P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 ! d_2)(1 - P(d_1 : d_2)) \end{aligned} \tag{12}$$

The first term we calculated above. The second term represents the probability of two measurement events, conditioned on those events *not* being part of a pair. This is essentially the probability that two independent measurement events have occurred, which is approximately the probability that two independent creation events have occurred (assuming defects are measured suitably efficiently). For a suitably large system at moderately low temperature, this probability can be estimated as $\propto \frac{|t_2 - t_1|}{(L\gamma_+)^{-2}} := \delta(L, T, \Delta t)$.

In practice, at low temperature for moderately sized systems, $P(d_1 : d_2)$ is very nearly 1. This arises from the low density of defects meaning that only rarely are there even a pair of defects in the system. Of course, if system size is made sufficiently large, this bare probability will become diminished, but it is still the case that defects within a separation distance $2\sqrt{D|t_2 - t_1|}$ are, more often than not, a pair at low temperature. For the same reason, $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 ! d_2)$ is very nearly 0 because this probability is roughly equivalent to the probability that two independent pair creation events have occurred, which is unlikely at low temperature and moderate system size. Again, for sufficiently large systems this probability grows, but it is likewise the case that this probability is small for defects within a distance $2\sqrt{D|t_2 - t_1|}$. Then, if we write $P(d_1 : d_2) = 1 - \epsilon(L, T)$, and perform some rearranging:

$$P(d_1 : d_2 | d_1^{x_1,t_1} \wedge d_2^{x_2,t_2}) = \tag{13}$$
$$\frac{1}{1 + \frac{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 \not| d_2) \epsilon(L,T)}{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)(1-\epsilon(T))}}$$
$$\geq \frac{1}{1 + \frac{\frac{\delta(L,T,\Delta t)}{1-\epsilon(L,T)}}{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)}}$$
$$\approx \frac{1}{1 + \frac{\delta(L,T,\Delta t)}{P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)}}$$

Thus, only when $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2) << \delta(L,T,\Delta t)$ is this factor not equal to 1. This naturally occurs when comparing defects that are much farther apart than diffusive motion would usually allow. For example: for a very large system, if one defect of a pair is measured at site 0 and another defect belonging to another independent pair is measured at site $L/2$ shortly thereafter (compared to the timescale for defect motion), it is exceedingly unlikely for these two measured defects to be a pair because it's exponentially unlikely for such a long random excursion to occur. In this way, the factor $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)$ serves as an indicator function which answers the question, "Could these two defects have arisen from a random walk starting in the same place?". The factor $\delta(L,T,\Delta t)$ sets the cutoff for a plausible excursion—i.e., when the error function is much less than this term, the denominator of 13 blows up, and the probability of performing that fusion is essentially zero.

In practice, the precise details of these additional factors arising from Bayes theorem aren't too important for the protocol to function, and we find that using the conditional probability $P(d_1^{x_1,t_1} \wedge d_2^{x_2,t_2} | d_1 : d_2)$ itself as a proxy for the full expression from Bayes theorem is sufficient to reliably correct errors. We provide some heuristic comparisons of different decoding schemes in appendix TODOGOHERE.

C. Common error modes and analysis

The two primary failure modes of the error correction procedure as temperature is made large are (1) defects escaping measurement patches, and (2) defects being paired incorrectly. (1) can be suppressed by working with large measurement patches, as the escape escape timescale is exponential in the number of sites within the measurement patch—i.e., it scales like $(1/\gamma_0)^{\#\text{ of sites on patch}}$. (2) can be suppressed to a degree, but at high enough temperature, will eventually happen with high probability, as is the case in the full-measurement version of the protocol. If too many errors are introduced between measurement cycles, it becomes impossible to correct the system even in principle because the syndrome measurements can no longer be unambiguously decoded. At sufficiently low temperature, however, with some assumptions, our algorithm can extend the lifetime of the system exponentially in system size.

Suppose that (a) defects are efficiently trapped by the measurement patches, and (b) temperature is low so that there is $\rho << 1$ defect pair per measurement patch. In this regime, condition (a) guarantees that defects don't escape measurement patches. Then, the only error pathway is a sequence of incorrect, but nontheless "plausible" error correction operations, whereby defects are paired with defects that are not their own pair.

To upper bound this error rate, we need to consider lowest order error processes more carefully. At low temperatures—i.e., low densities of errors—the primary breakdown of the corrective step occurs when a defect is paired with another defect that is not its true pair. This occurs because its pairing defect has not yet been measured. In other words, one defect of a pair is measured, and the other defect spends "too long" migrating around in the unmeasured bulk region. Suppose, then, that the average amount of time a defect spends sitting on a measurement patch is τ_s, and the average amount of time it takes a defect to be detected on a measurement patch, after creation, is τ_d. Further suppose $\tau_s/\tau_d = \delta << 1$—that is, on average, defects are unlikely to escape from a measurement patch over the timescale it takes a defect to migrate within the bulk region to a measurement patch.

To estimate the rate at which these spurious error-correction events occur, we first consider a slightly altered model. Suppose that a single pair of defects is in a system, and that no new pairs will be introduced. Further, suppose that there's an error process that occasionally increases the distance between this pair by an amount $2\sqrt{\gamma_0 \delta t}$, for hopping rate γ_0 and for system time δt. Note that this means the error process acts on longer distance scales as system time increases. The particular form of this error process is chosen to mimic a spurious error correction event, where a pair of defects is "corrected", but this pair of defects was not a real pair. As time increases, the characteristic distance over which this error process could occur also increases, in accordance with the typical pair-wise separation between two defects performing a random walk. This typical distance is exactly the factor used by the error correction algorithm to determine if a pair of defects should be corrected or not.

If this error process happens every τ_ϵ, then an uncorrectable error will occur if the bulk defect remains undetected up until it crosses half the system. For patches of size λ, the probability that this occurs is roughly $(\tau_\epsilon/\tau_d)^k$, where k is the number of times the error process must occur for the error process to have separated the defects a distance $(L/2\lambda)$ and τ_d is the timescale associated with a non-erroneous error correction operation. Because this error process happens approximately every τ_ϵ, after a timescale $q * \tau_\epsilon$, the defect pairs will have been separated a distance $\Sigma_{i=1}^{q} \sqrt{i * \gamma_0 \tau_\epsilon}$, assuming that they are never correctly paired. This sum $\Sigma_{i=1}^{q} \sqrt{i}$ grows asymptotically like $q^{3/2}$. Thus, k then scales approximately like $(L/\lambda)^{2/3}$.

Finally, if we have L/λ such simultaneous error processes independently in the system—one for each measurement patch—then the total error probability scales like

$$P(\text{uncorrectable error}) \leq (\mathrm{L}/\lambda)(\tau_\epsilon/\tau_\mathrm{d})^{(\mathrm{L}/\lambda)^{2/3}}. \quad (14)$$

For this system, τ_ϵ is much smaller than τ_d, thus the full probability of erroneous corrective operations is exponentially small in system size.

This probability serves as an upper bound because it is extremely pessimistic about the sorts of error processes occurring. Note that in the real system, a spurious error correction process that effectively increases the distance between a pair of defects also reduces the total number of defects by two. Thus, while our toy error model is "non-interacting"—that is, it assumes L/λ independent error processes which, in sum, take the form described in equation 14—a more careful treatment of the error process, including interactions between defects, as in the real model, would result in an error probability *smaller* than the one calculated here.

In the following section, we provide numerical evidence that the lifetime of the Ising model in the presence of the protocol scales quasi-exponentially with the number of measurement patches, as anticipated by our upper bound 14.

IV. FINITE TEMPERATURE SIMULATION EXPERIMENTS

In this section, we examine the behaviour of the protocol at different system sizes, temperatures, and measurement fractions. System size scalings are generally reported in terms of scaling with the number of *measurement patches*. We make this choice because the size of the measurement patches is the smallest length scale relevant to the algorithm, and thus controls the majority of the scaling behavior. Perhaps unsurprisingly, systems of radically different size but with the same number of measurement patches have fairly comparable lifetimes (e.g., size 200 system with 10 patches versus size 100 system with 10 patches). We elaborate on the scaling of the measurement fractions further in the discussion.

Fig. 2 depicts the scaling of the system lifetime with temperature. At low temperatures, the system lifetime increases dramatically with both lower temperatures, and larger system size. Due to finite size effects, it's difficult to extract an unambiguous threshold temperature, but below $T \approx .13$, the lifetime manifestly increases with larger system size.

This is clearer in Fig. 3, which depicts the scaling of the lifetime with system size directly. Note that, for models above $T \approx .13$, larger system sizes asymptote to a fixed lifetime improvement, whereas for models below $T \approx .13$, the lifetime grows quasimonotonically with system size. Beyond $\lambda \approx 15$, finite size effects are significantly reduced. This is because small-system sizes can-

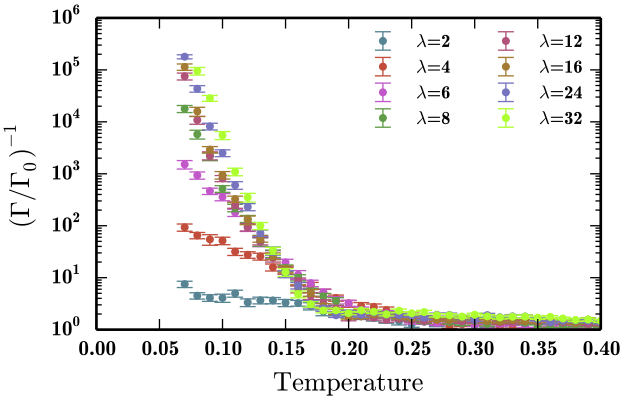

FIG. 2. Lifetime of various system sizes as a function of Temperature. Notice the transition between no-effect, and significant lifetime improvement near $T = .16$.

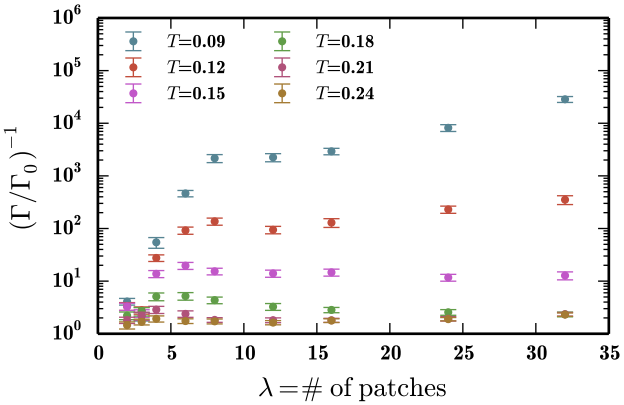

FIG. 3. Lifetime of the Ising model at various temperatures plotted against system size. Note the monotonic growth in lifetime for low temperature models, as well as the plateau in lifetime for moderately sized systems.

not easily suppress second order errors, such as defects escaping from measurement patches or multiple pairs of defects in the system. Such errors are actually uncorrectable for systems of size $\lambda \leq 4$—hence the plateau appearing around $\lambda = 7$ to $\lambda = 12$. This plateau is of height $O((\Gamma_0^2)^{-1})$—the characteristic timescale of these "second-order" events. This scaling is more clear when plotting the liftime divided by the bare lifetime squared—i.e. Γ/Γ_0^2—as in Fig. 4.

Another free parameter in the algorithm is the ratio of measured sites λ_m to bulk sites λ_b. For this work, we fixed $\lambda_m = 3$ and varied λ_b. Fig. 5 depicts the effect measuring fewer sites on the system lifetime as a function of Temperature for several different measurement fractions. In sum, measuring a smaller fraction of the lattice causes the threshold temperature to shift down-

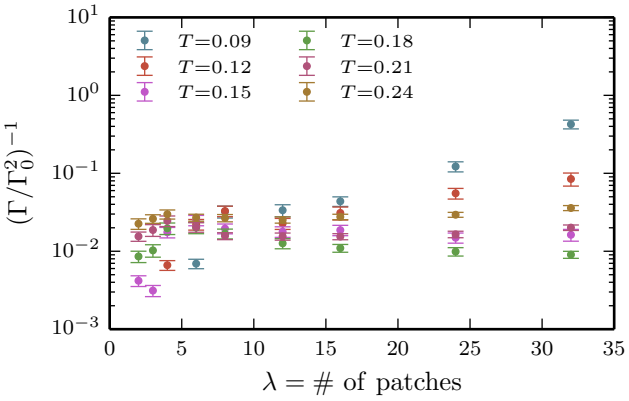

FIG. 4. The same data in Fig. 3 rescaled by an additional factor Γ_0^{-1}, to emphasize the origin and scaling of the plateau in system lifetimes.

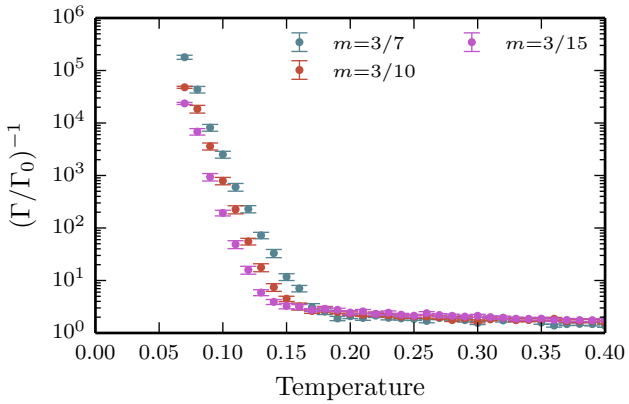

FIG. 5. Several measurement fractions plotted as a function of temperature at fixed number of measurement patches. Note the shift in the critical temperature as function of measurement fraction.

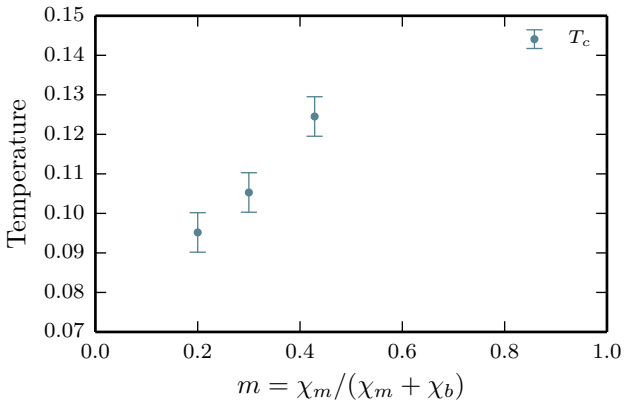

FIG. 6. The critical temperature plotted as a function of measurement fraction. At zero measurement fraction, the critical temperature is 0, and at unit measurement fraction, the critical temperature is ~ 1.

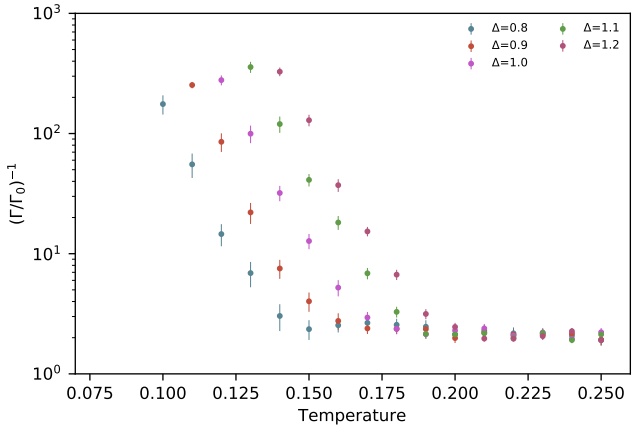

FIG. 7. Here we plot the lifetime of the system as a function of temperature for a variety of energy scales Δ (see Eq. **??**).

wards. This scaling of the threshold temperature itself is depicted explicitly in Fig. 6.

Finally, in Fig. 7, we plot the lifetime of the system in the presence of the protocol at various energy scales Δ. By contrasting Figures 5 and 7, one can deduce the relative benefits of error suppression via more measurement resources versus error suppression via reservoir engineering (i.e., a larger gap to excitation).

V. GENERALIZATION TO HIGHER DIMENSION

In this section, we sketch how the algorithm presented in Sec. III generalizes to a higher dimensional stabilizer quantum memory—the toric code. Where the dynamics of the 1D Ising model are typified by one dimensional random walks of domain walls, the nonequilibrium dynamics of the toric code are driven by two dimensional random walks of "quasiparticle" excitations. Consider the toric code hamiltonian, Eq. 15.

$$H_{\mathrm{TC}} = -\Delta_e \sum_v A_v - \Delta_m \sum_p B_p, \tag{15}$$

$$A_v \equiv \prod_{j\in v} \sigma_j^z, \quad B_p \equiv \prod_{j\in p} \sigma_j^x, \tag{16}$$

Where domain wall excitations in the 1D ising model are associated with -1 eigenstates of the $\sigma_z^i \sigma_z^{i+1}$ stabilizers, quasiparticle excitations for the toric code are associated with -1 eigenstates of the A_v and B_p stabilizers as defined in Eq. 16.

Broadly speaking, the algorithm is identical, but instead of having "patches" of measurement, there are mea-

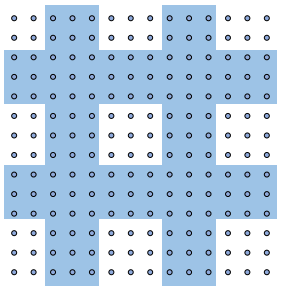

FIG. 8. Here we sketch one possible $m = 3/4$ geometry for the measurement rails for a realization of our protocol. More sparse geometries can be realized simply by moving the rails of measurement farther apart. Measured sites are in light blue, and vertex locations for the toric code are circles. Spin variables (not pictured) reside directly between neighboring vertices.

surement "rails", as indicated in Fig. 8. "Centering" of defects on rails amounts to shift-swaping defects into the center of the measurement rail. We conjecture that a sparse measurement strategy with randomly placed measurement patches of fixed diameter might exist for sufficiently large/cold systems, but we do not comment further on that here. The rail geometry of Fig. 8 is the simplest geometry that allows us to rely on the heuristic argument for a threshold sketched in Sec. III C.

VI. DISCUSSION

We have provided numerical and theoretical evidence of a few-measurement error correction protocol for a stabilizer code with string-like excitations. The primary technical innovation of our algorithm is the bayesian decoding scheme for pairing defects, sketched in Sec. III.

* daniel.freeman@berkeley.edu

† mnsarov@sandia.gov