arXiv:2110.10732v3 [quant-ph] 26 Mar 2022

A Logarithmic Bayesian Approach to Quantum Error Detection

Ian Convy[1,2] and K. Birgitta Whaley[1,2]

[1]Department of Chemistry, University of California, Berkeley, CA 94720, USA
[2]Berkeley Quantum Information and Computation Center, University of California, Berkeley, CA 94720, USA

We consider the problem of continuous quantum error correction from a Bayesian perspective, proposing a pair of digital filters using logarithmic probabilities that are able to achieve near-optimal performance on a three-qubit bit-flip code, while still being reasonable to implement on low-latency hardware. These practical filters are approximations of an optimal filter that we derive explicitly for finite time steps, in contrast with previous work that has relied on stochastic differential equations such as the Wonham filter. By utilizing logarithmic probabilities, we are able to eliminate the need for explicit normalization and can reduce the Gaussian noise distribution to a simple quadratic expression. The state transitions induced by the bit-flip errors are modeled using a Markov chain, which for log-probabilties must be evaluated using a LogSumExp function. We develop the two versions of our filter by constraining this LogSumExp to have either one or two inputs, which favors either simplicity or accuracy, respectively. Using simulated data, we demonstrate that the single-term and two-term filters are able to significantly outperform both a double threshold scheme and a linearized version of the Wonham filter in tests of error detection under a wide variety of error rates and time steps.

1 Introduction

One of the major obstacles to large-scale quantum computation is the high frequency of qubit errors. Small interactions with the environment or imperfections in gate implementation can perturb the underlying quantum state throughout a computation and ultimately render the output useless. These issues only magnify as the size and complexity of the quantum computer increases, jeopardizing any attempt to demonstrate quantum advantage in key tasks such as the factorization of large integers [1]. The obvious challenge posed by quantum errors has given rise to the field of *quantum error correction* [2] and spurred the development of numerous techniques to identify and correct errors when they occur. These schemes, referred to as error correction *codes*, typically operate by encoding the quantum state into a larger Hilbert space and then monitoring the location of the state within this space using a specific set of observables.

The present work is focused on continuous or "always-on" error correction, where the state of the system is continuously monitored through a noisy signal channel in order to diagnose errors and to make corrections if necessary. Since the signal is noisy, we cannot know with certainty whether an error has occurred, and must instead make subjective

Ian Convy: ian_convy@berkeley.edu

judgements about the condition of the system based on past measurements and any other information we might have. In order to quantify the uncertainty in this task it is natural to adopt a Bayesian framework, which allows for new information to be easily combined with prior knowledge of the system using Bayes' theorem.

Assuming that the errors experienced by a quantum computer are uncorrelated, we can model the system as a Markov chain whose state is imperfectly observed by a measurement apparatus. Optimal treatments of this problem originate with work done by Wonham and his derivation of what is today known as the *Wonham filter*, which describes the evolution of the Markov state probabilities conditioned on a particular set of noisy observations [3]. Despite its optimality, this filter has not found widespread use for error correction, in part because it assumes that the value of the signal is characterized as a continuous function of time, while in practice the signal is typically only observed as a discrete set of measurement samples. As a non-linear stochastic differential equation, the Wonham filter can also be challenging to evaluate numerically, even if the signal function is known. Due to these limitations, most continuous error correction schemes avoid using Bayesian theory and instead rely on various thresholding procedures which are easier to implement but known to provide suboptimal performance [4][5][6].

The purpose of this work is to introduce a discrete-time filter for continuous quantum error correction that significantly outperforms common thresholding schemes, while still being practical to implement on real hardware. We construct our filter using logarithmic probabilities, which are numerically stable, easy to normalize, and allow for straightforward evaluation of Gaussian noise distributions using only arithmetic operations. The filter has two different forms based on how it updates the posterior, with one method emphasizing accuracy while the other emphasizes computational simplicity. Both versions of the filter are able to achieve near optimal performance under a wide variety of error rates and time steps.

The remainder of the paper has the following structure. Sec. 2 reviews the three-qubit code that we will use as a simple error correction primitive, and describes the principles of continuous measurement. Sec. 3 derives an optimal Bayesian filter for identifying bit-flip errors, which cannot be practically implemented in its exact form. Sec. 4 discusses prior work in the field of continuous quantum error correction and outlines several obstacles to implementing a Bayesian filter. Sec. 5 proposes an approximate version of the optimal filter which overcomes those obstacles, and provides two different implementations which are both experimentally feasible using today's technology. Sec. 6 tests our filter on simulated data and compares the performances of the two implementations against other schemes found in the literature. Sec. 7 summarizes our findings and identifies areas for future work. An Appendix is also provided, which contains technical details that supplement the body of the paper.

2 Quantum Error Correction

2.1 Three-qubit coding scheme

To insulate a quantum system from errors, we must add some level of redundancy to its Hilbert space. This is achieved by assembling a set of physical qubits which is larger than necessary to perform the desired task, with the idea that these extra degrees of freedom will be used to implement an error correction scheme. The computation is then understood in terms of its effect on the *coding subspace* of the system, which describes the portion of the expanded Hilbert space where the intended quantum evolution will take place.

For simplicity we focus on a three-qubit repetition code, which is designed to protect the state of a single qubit from the effects of bit-flip errors. This encoded qubit is the true object of interest, so the computational task is conceptualized in terms of single-qubit operations which then need to be mapped to the full three-qubit system. We can perform this mapping by considering the so-called *logical qubit*, defined as

$$|0_L\rangle \equiv |000\rangle \text{ and } |1_L\rangle \equiv |111\rangle\,, \tag{1}$$

which occupies the two-dimensional subspace spanned by $|000\rangle$ and $|111\rangle$ that we designate as the coding subspace.

The logical states in Eq. (1) are used to encode the desired behavior of the system in the absence of errors, while the other parts of the Hilbert space serve to catch any bit-flips that occur. For example, a bit-flip error on the first qubit will transform $|000\rangle$ into $|100\rangle$ and $|111\rangle$ into $|011\rangle$, so the system will be shifted into a different two-dimensional subspace. As long as we can detect this change in the subspace, it will be possible for us to track and then correct the errors that have occurred on the system.

To perform this detection, we require a set of observables that uniquely identify the coding subspace and each of the three subspaces generated from bit-flips on the first, second, and third qubits. This is achieved using the *parity operators* Z_1Z_2 and Z_2Z_3, whose eigenstates and eigenvalues are specified by

$$(Z_1Z_2\,|\psi\rangle\,, Z_2Z_3\,|\psi\rangle) = \begin{cases} (|\psi\rangle\,, |\psi\rangle) & |\psi\rangle = a\,|000\rangle + b\,|111\rangle \\ (-|\psi\rangle\,, |\psi\rangle) & |\psi\rangle = a\,|100\rangle + b\,|011\rangle \\ (|\psi\rangle\,, -|\psi\rangle) & |\psi\rangle = a\,|001\rangle + b\,|110\rangle \\ (-|\psi\rangle\,, -|\psi\rangle) & |\psi\rangle = a\,|010\rangle + b\,|101\rangle\,, \end{cases} \tag{2}$$

for any complex numbers a and b. Eq. (2) shows that the combined action of the parity operators splits the three-qubit Hilbert space into four distinct 2-D subspaces, each characterized by a different pair of eigenvalues in $\{-1, 1\}$ which are referred to as *syndromes*. The $(1, 1)$ syndrome identifies the coding subspace, while the other three combinations correspond to states which result from a bit-flip on one of the qubits. We refer to the latter as the *error subspaces*, and note that the syndrome value identifies on which qubit the error has occurred.

The observables Z_1Z_2 and Z_2Z_3 offer us a simple method for bit-flip error detection, assuming that they can in fact be measured. After initializing the system into the coding subspace, we simply perform parity measurements of our system as needed and monitor the response of the system. If our results indicate that the system has moved out of the coding subspace, then we can record this deviation and apply a bit-flip correction based on which error subspace the system has been moved to. For example, if one error occurs on the first qubit then we have

$$a\,|000\rangle + b\,|111\rangle \xrightarrow[\text{qubit 1}]{\text{error}} a\,|100\rangle + b\,|011\rangle \xrightarrow[\text{qubit 1}]{\text{correction}} a\,|000\rangle + b\,|111\rangle\,, \tag{3}$$

and the system has been returned to its proper state. Each time a bit-flip is diagnosed to have occurred on a given qubit, we can simply apply another bit-flip to that same qubit and the error will be undone.

It is important to note that if *two* bit-flip errors occur simultaneously (or rapidly enough to appear indistinguishable) then the above detection scheme will mistakenly assume that only one error has occurred and perform the wrong correction. For example, if errors occur

on the second and third qubits then

$$a\,|000\rangle + b\,|111\rangle \xrightarrow[\text{qubits 2,3}]{\text{errors}} a\,|011\rangle + b\,|100\rangle \xrightarrow[\text{qubit 1}]{\text{correction}} a\,|111\rangle + b\,|000\rangle\,, \tag{4}$$

and the system will experience a *logical error* since the a and b coefficients have been exchanged. This type of misdiagnosis occurs because the errors in Eq. (3) and Eq. (4) both move the system into the same parity subspace and thus return the same syndrome values, so we must respond with the same correction in both cases. Given that single errors are more common then double errors under realistic conditions, the most reasonable response is to always assume that one bit-flip has occurred whenever the system enters an error subspace. This compromise is an inescapable consequence of using a short code, though we can reduce its impact by detecting errors more quickly and thus better resolving the difference between one bit-flip and two sequential bit-flips.

Throughout this paper we will focus our attention on *state tracking*, where the goal is to correctly identify the net number of bit-flips that have acted on the system at a given point in time. For simplicity, we will assume that the system is initialized into $|000\rangle$, and only evolves due to random bit-flip errors. That said, our results are not restricted to this quantum memory regime, as Eq. (2) shows that the parity operators act uniformly on states within a given coding or error subspace, meaning that the measurement signal will be identical regardless of whether the system is in a basis state or an arbitrary superposition. Even if a Hamiltonian is applied to the system, as would occur during a process like quantum annealing [7], the measurement signal will not be affected so long as the Hamiltonian does not mix together the four parity subspaces. As a result, any filter designed to handle quantum memory can be seamlessly applied to more complicated processes by just changing the metric used to evaluate its performance.

2.2 Error correction with continuous measurements

The three-qubit code described in Sec. 2.1 is agnostic to the actual method used to probe Z_1Z_2 and Z_2Z_3, since the separation of the Hilbert space into coding and error subspaces is a property of the observables themselves rather than of any particular measurement scheme. Most work on quantum errror detection has focused on *discrete* error correction, which utilizes projective measurements to resolve the exact syndrome of the state at periodic intervals [2]. However, recent advances in superconducting qubit architecture has led to increased interest in *continous* error correction, where the parity operators are monitored at all times and a noisy readout of the underlying syndrome values is generated [8][9]. Under these noisy conditions, a filter will be needed in order to extract relevant information from the signal.

In an idealized setting, we can formulate a simple mathematical description of the continuous measurement process using Gaussian POVMs. At time t, the probability of observing signal readout $\alpha(t)$ from a parity operator with syndrome $S(t)$ is

$$\text{Prob}(\alpha(t)) \propto \exp\left[-\frac{(S(t)-\alpha(t))^2}{2\sigma^2}\right], \tag{5}$$

which implies that the signal values are distributed as $\mathcal{N}(S(t),\sigma^2)$. The syndrome value $S(t) \in \{-1,+1\}$ is determined by the state of the system at time t, and will therefore change as errors occur. We assume that the variance σ^2 of the noisy signal is identical across all states and is thus independent of time.

The weak measurement described by Eq. (5) must occur over some finite time Δt, whose duration will determine the variance of the signal through the relation $\sigma^2 = \frac{k}{\Delta t}$ for

some positive k (the strength of the measurement is thus proportional to $\frac{1}{k}$). A longer measurement $\tilde{M}$ of time T can be constructed by averaging over a set of these weaker measurements

$$\tilde{M}(t,T,\Delta t) = \left(\frac{T}{\Delta t}\right)^{-1} \sum_{n=1}^{\frac{T}{\Delta t}} \alpha(t+n\Delta t) = \frac{1}{T}\sum_{n=1}^{\frac{T}{\Delta t}} [S(t+n\Delta t)+\epsilon]\Delta t, \tag{6}$$

where the signal at time increment $t+n\Delta t$ is the sum of the syndrome $S(t+n\Delta t)$ and a Gaussian noise term $\epsilon \sim \mathcal{N}(0, \frac{k}{\Delta t})$. The value of t represents the time at which the averaging process starts, while the integer n indexes the set of sequential weak measurements being averaged over, each of duration Δt. Since the average is taken over a time interval of length T, we can perform $\frac{T}{\Delta t}$ weak measurements within the interval.

Given that the Gaussian noise ϵ in Eq. (6) is additive, its average can be calculated separately from the syndrome average as

$$\xi(T) \equiv \frac{\Delta t}{T}\sum_{n=1}^{\frac{T}{\Delta t}} \epsilon = \frac{\Delta t}{T}\tilde{\epsilon}, \tag{7}$$

where $\tilde{\epsilon} \sim \mathcal{N}(0, \frac{kT}{(\Delta t)^2})$ due to the additive property of the variance and thus $\xi(T) \sim \mathcal{N}(0, \frac{k}{T})$ after scaling by $\frac{\Delta t}{T}$. The value of $\tilde{M}$ is therefore equal to the averaged syndrome plus a new noise term $\xi(T)$, whose variance is inversely proportional to T:

$$\tilde{M}(t,T,\Delta t) = \frac{1}{T}\sum_{n=1}^{\frac{T}{\Delta t}} [S(t+n\Delta t)\Delta t] + \xi(T). \tag{8}$$

From Eq. (8), we define the *continuous* measurement M as the limit of $\tilde{M}$ for $\Delta t \to 0$, which turns the discrete sum into an integral,

$$M(t,T) = \frac{1}{T}\int_0^T S(t+\tau)d\tau + \xi(T), \tag{9}$$

that is distributed as a Gaussian with variance $\frac{k}{T}$ and a mean value given by the integrated average of the syndrome over the interval.

In Figure 1 we provide examples of measurement sequences that could be observed as a three qubit system evolves under the influence of random bit-flip errors for 30 μs. For a system starting in $|000\rangle$, the measurements $M(t,T)$ for each syndrome will be centered at $+1$ and oscillate randomly due to the additive noise term. Comparing Figures 1A and 1B, it is clear that increasing the integration time T from 0.1 μs (panel B) to 1 μs (panel A) dramatically reduces the fluctuations of the signal, making it much easier to see visually the effects of the bit-flip errors on the measurements. Perhaps counterintuitively, this reduction in the variance does not actually make the task of identifying a syndrome value from its noisy measurements any easier, since the standard error (SE) of mean value estimation scales with both the noise strength $\frac{k}{T}$ *and* the number of samples to be averaged. Given that the number of measurements n in a time interval of length β is inversely proportional to T, i.e., $n = \frac{\beta}{T}$, we have

$$\mathrm{SE}^2 = \frac{\sigma^2}{n} = \frac{k}{T}\frac{T}{\beta} = \frac{k}{\beta} \tag{10}$$

and thus the uncertainty in our estimate of the syndrome will not depend on T.

This is not to say, however, that the length of the integration period has no effect on the state tracking problem. Indeed, Eq. (10) simply states that there is a proportionate

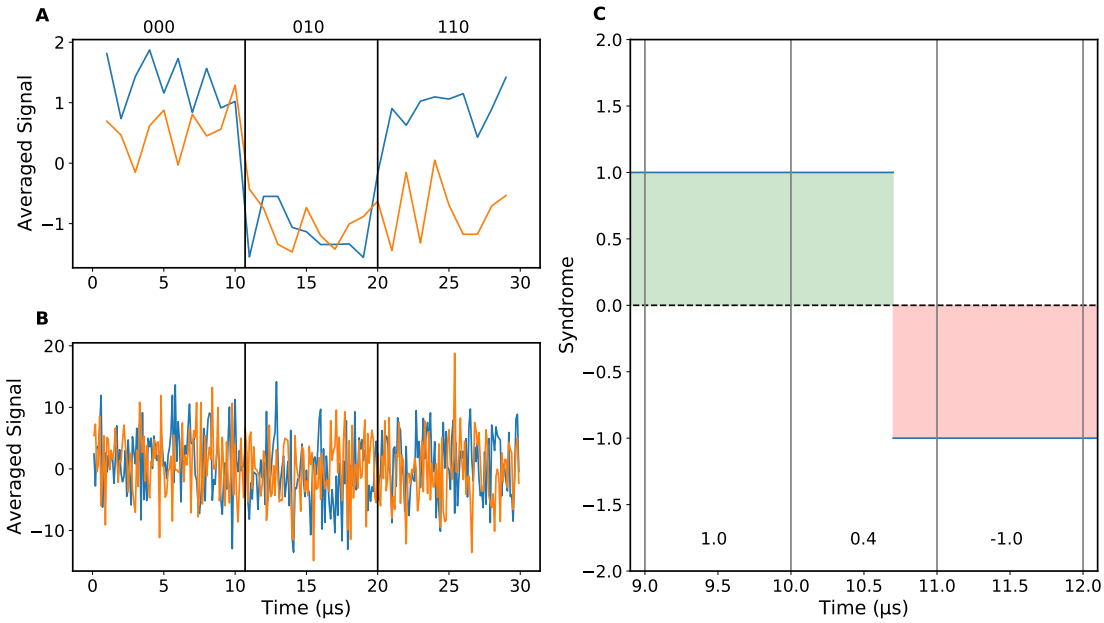

Figure 1: Plots A and B show examples of 30 μs measurement sequences for $T = 1$ μs and $T = 0.1$ μs respectively, with $k = 0.5$ μs for both plots. Each point represents the unitless value of $M(t, T)$ taken at the time shown on the x-axis for one of the two parity operators, with the blue lines tracking the noisy measurement of Z_1Z_2 and the orange lines tracking Z_2Z_3. The underlying state evolution in both plots is identical, with the vertical black lines indicating the times at which bit-flip errors occurred (10.7 μs and 20 μs). The state of the system before and after each error is given by the three-bit number above plot A. Plot C shows the true syndrome value from Z_1Z_2 without noise, zoomed in on times near the first bit-flip. The vertical gray lines indicate the integration intervals used to generate the averaged syndromes found in Eq. (9) when $T = 1$ μs, with the numerical values of these averages given at the bottom of the plot. The shaded regions indicate the "area under the curve" contribution at a given time, with green indicating a positive contribution and red a negative contribution. The bit-flip error occurs within the second interval, causing its syndrome average to deviate from the expected values of ± 1 and instead lie somewhere in between.

trade-off between the informativeness of a measurement ($\sigma^2 = \frac{k}{T}$) and the number of such measurements ($n = \frac{\beta}{T}$) when the syndrome value is *fixed*. When we introduce bit-flip errors into the system and thus consider situations where the syndrome changes with time, the analysis is significantly more complicated. Figure 1C shows that when an error occurs within an integration period, a portion of the syndrome average will consist of $+1$ contributions and the other portion will consist of -1 contributions. The resulting syndrome mean can therefore take on any value in the interval $[-1, +1]$, which suggests that the measurement will be inherently less informative than a measurement that occurs when the state of the system is constant. As the value of T grows larger, more measurements will fall into this category. The incorporation of these intermediate syndrome means into an optimal state tracking model is explored in Sec. 3.

2.3 Bayesian treatment of measurement

Determining the state of a quantum system from noisy signals is not a new problem, and the field of quantum mechanics has long grappled with how to properly describe the mechanism and effects of measurement [10]. One of the key challenges is knowing how to properly update the density matrix ρ_s of a system to reflect the fact that we have performed a measurement on it. A common procedure involves coupling the system to a detector and then tracing out the detector's degrees of freedom,

$$\rho_s = \mathrm{Tr}_d(\rho_{sd}) \tag{11}$$

where ρ_{sd} is the joint density matrix of the system and detector. By its very nature this approach cannot offer any information about the state of the system after a *specific* measurement, since we are averaging over all of the possible detector configurations without reference to the outcome that really occurred. Instead, what Eq. (11) describes is the behavior of an *ensemble* of systems after they are all measured, since the frequency of each measurement outcome can be specified without knowing the result for any given system. This approach is therefore insufficient for an error correction scheme, as the very purpose of error correction is to monitor a specific system and respond to the errors that actually occur, rather than to the average over all possible errors.

The issue of specific outcomes versus ensemble behavior is reminiscent of the disagreement between Bayesian and frequentist statistics [11], and it is therefore unsurprising that a Bayesian formalism for quantum measurement was developed. This formalism was pioneered by Korotkov, who proposed an update rule for the diagonal elements of the density matrix after a measurement M using Bayes' theorem that has the form

$$\hat{\rho}_{ii} = \frac{P(M|i)\rho_{ii}}{\sum_j P(M|j)\rho_{jj}}, \tag{12}$$

where ρ_{ii} is the original diagonal element acting as the prior and $\hat{\rho}_{ii}$ is the updated (posterior) element [12]. The term $P(M|i)$ describes the probability of observing measurement result M given that the system is in state i.

3 An Optimal Bayesian Filter

For our purposes, the significance of Korotkov's measurement formalism from Eq. (12) lies in its treatment of the density matrix elements as classical Bayesian probability distributions, which allows us to leverage the existing toolkit of Bayesian statistics [13][14]

to develop an error correction algorithm. Fundamentally, the goal of this algorithm is to determine, using all available information, the *most probable* state of the system at each time step in order to detect errors. As in Eq. (12), we seek to derive a posterior probability distribution for the state of our system after each measurement, using knowledge of the underlying bit-flip dynamics and of the Gaussian noise corrupting the measurement signal.

3.1 Recursive form of the posterior probability

For notational simplicity, we introduce the vector-valued quantity

$$\vec{M}_i \equiv \begin{bmatrix} M^{(1)}(iT, T) \\ M^{(2)}(iT, T) \end{bmatrix}, \tag{13}$$

where $M^{(1)}$ and $M^{(2)}$ are measurements of the parity operators Z_1Z_2 and Z_2Z_3 respectively and i is a non-negative integer. The time argument iT appears because the sets of (non-overlapping) sequential measurements will be spaced out in increments of T. To denote the state of the system after the ith measurement, i.e., at time $t = (i+1)T$, we employ the compound index $0 \leq \ell_i \leq 7$. Using this notation, we write the posterior probability of the ℓ_ith state as

$$\hat{P}(\ell_i) \equiv P(\ell_i|\vec{M}_0...\vec{M}_i), \tag{14}$$

where $P(C|AB)$ is generically the probability distribution of variable C given knowledge of the state of variables A and B. Using the chain rule of probability, Eq. (14) can be put into a recursive form as

$$\begin{aligned}
\hat{P}(\ell_i) &= \frac{P(\vec{M}_i\ell_i|\vec{M}_0...\vec{M}_{i-1})}{P(\vec{M}_i|\vec{M}_0...\vec{M}_{i-1})} \\
&= \frac{1}{P(\vec{M}_i|\vec{M}_0...\vec{M}_{i-1})} \sum_{\ell_{i-1}=0}^{7} \hat{P}(\ell_{i-1}) P(\vec{M}_i\ell_i|\ell_{i-1}\vec{M}_0...\vec{M}_{i-1}) \\
&= \frac{1}{P(\vec{M}_i|\vec{M}_0...\vec{M}_{i-1})} \sum_{\ell_{i-1}=0}^{7} \hat{P}(\ell_{i-1}) P(\vec{M}_i|\ell_i\ell_{i-1}\vec{M}_0...\vec{M}_{i-1}) P(\ell_i|\ell_{i-1}\vec{M}_0...\vec{M}_{i-1}) \\
&= \frac{1}{P(\vec{M}_i|\vec{M}_0...\vec{M}_{i-1})} \sum_{\ell_{i-1}=0}^{7} \hat{P}(\ell_{i-1}) P(\vec{M}_i|\ell_i\ell_{i-1}) P(\ell_i|\ell_{i-1}),
\end{aligned} \tag{15}$$

where in the last line we explicitly assume that $\vec{M}_i$ and ℓ_i are conditionally independent of $\{\vec{M}_0, ..., \vec{M}_{i-1}\}$ given knowledge of ℓ_{i-1}. This is equivalent to assuming that the state transitions depend only on the prior state (Markovian assumption), and that the additive measurement noise is uncorrelated across time. Since $P(\vec{M}_i|\vec{M}_0...\vec{M}_{i-1})$ has no explicit ℓ_i dependence and is not otherwise coupled to any of the terms in the sum, it will be ignored in our subsequent analysis and treated as simply a normalization factor.

If we assume that the system begins in a known state, which is reasonable at the start of a quantum experiment, then Eq. (15) provides a recipe for iteratively computing the probabilities of future states in light of new measurement results. The two quantities that must be solved for as part of this procedure are

1. $P(\ell_i|\ell_{i-1})$, which is the probability of jumping from state ℓ_{i-1} to state ℓ_i in time T due to bit-flip errors, and

2. $P(\vec{M}_i|\ell_i\ell_{i-1})$, which is the probability of measuring values $\vec{M}_i$ from the parity operators when we have knowledge of the system states at the beginning and end of the integration period.

The remainder of this section will be dedicated to deriving explicit functional forms for these two probability terms.

3.2 Analysis of the transition probability

Our treatment of bit-flip errors will assume that they act on the system independently of one another and occur at a fixed rate μ that is known to us. In saying that the errors are independent, we mean that an error occurring in one time interval gives us no information about whether an error will occur in any other non-overlapping interval. The number e_k of such errors acting on the kth qubit in a time interval of length T will obey the following Poisson distribution

$$P(e_k) = \frac{(\mu T)^{e_k}}{e_k!} e^{-\mu T}, \tag{16}$$

where we emphasize that the error rate is identical across all three qubits.

In order to use Eq. (16) to derive a functional form for $P(\ell_i|\ell_{i-1})$, we must specify the number of errors needed to connect state ℓ_{i-1} to state ℓ_i. Taking for example a system that starts as $|000\rangle$ at step $i-1$ and ends up as $|100\rangle$ at step i, corresponding to $\ell_{i-1} = 0$ and $\ell_i = 4$, it is clear that the first qubit must have been flipped at some point in the interval between the two steps. However, this does not mean that only a single error occurred on the first qubit, or that none of the other qubits experienced any errors. Rather, it means that the first qubit experienced an *odd* number of errors while the rest of the qubits experienced an *even* number of errors (including no error at all), such that the net number of bit-flips works out to be one for the first qubit and zero for the others. We can therefore express $P(\ell_i|\ell_{i-1})$ as a sum over the Poisson probabilities of every possible error combination consistent with the net number of flips between ℓ_i and ℓ_{i-1}.

Rather than evaluate this sum directly, it is easier to look at each qubit separately and calculate the probability that it will experience either an even or odd number of errors. Since the errors on each qubit are independent, we can then take appropriate products of these probabilities to calculate $P(\ell_i|\ell_{i-1})$. The probability that the kth qubit will experience an even number of errors is

$$P(e_k \text{ is even}) = \sum_{j=0}^{\infty} P(e_k = 2j) = e^{-\mu T} \sum_{j=0}^{\infty} \frac{(\mu T)^{2j}}{(2j)!} = e^{-\mu T} \cosh(\mu T), \tag{17}$$

and therefore the probability of an odd number of errors is

$$P(e_k \text{ is odd}) = 1 - P(e_k \text{ is even}) = e^{-\mu T} \sinh(\mu T). \tag{18}$$

Using Eq. (17) and Eq. (18), the transition probability $P(\ell_i|\ell_{i-1})$ is given by

$$P(\ell_i|\ell_{i-1}) = [\sinh(\mu T)]^{d(\ell_i,\ell_{i-1})} [\cosh(\mu T)]^{3-d(\ell_i,\ell_{i-1})} \exp[-3\mu T], \tag{19}$$

where $d(\ell_i, \ell_{i-1})$ is the Hamming distance between the bit representation of ℓ_i and the bit representation of ℓ_{i-1}. In words, the probability consists of an exponential multiplied by a sinh term for every qubit that experiences a net flip and a cosh term for every qubit whose state is ultimately left unchanged.

Since the error rate μ is independent of time, the value of index i is irrelevent to the value of $P(\ell_i|\ell_{i-1})$. This allows us to define a single 8×8 time-invariant transition matrix J

with elements $J_{\ell_{i-1},\ell_i} \equiv P(\ell_i|\ell_{i-1})$. This matrix can be understood as the parameterization of a discrete Markov chain [15] which describes how bit-flip errors can alter the state of our system over time intervals of length T. For the remainder of this paper we will use the elements of J to denote the transition probabilities $P(\ell_i|\ell_{i-1})$.

3.3 Analysis of the measurement density

From our discussion of continuous measurement in Sec. 2.2, we know that the value of the ith measurement will depend on the average value of the syndrome from time iT to $(i+1)T$, plus an additive Gaussian noise term. Since it will be necessary to discuss measurements of the two parity operators separately, we unpack the vector quantity $\vec{M}_i$ as per Eq. (13) into its components $M_i^{(1)}$ and $M_i^{(2)}$, which describe the results of measuring Z_1Z_2 and Z_2Z_3 respectively. The measurement distribution $P(\vec{M}_i|\ell_i\ell_{i-1})$ will be hereafter written as $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ to make this distinction explicit. For convenience we denote the averaged syndrome values as $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$, such that

$$\bar{S}_i^{(j)} \equiv \frac{1}{T}\int_0^T S^{(j)}(iT+\tau)d\tau \tag{20}$$

where $S^{(j)}(t)$ is the value of the jth syndrome at time t.

Using this labeling, we have from Eq. (9) that $M_i^{(j)} \sim \mathcal{N}(\bar{S}_i^{(j)}, \frac{k}{T})$, so $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ will broadly resemble a bivariate Gaussian distribution. However, since the means $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ are themselves random variables due to errors occurring within the integration period, the measurement distribution will consist of a Gaussian function integrated over mean values in the continuous interval [-1, +1]. This is given by

$$P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1}) = \int_{-1}^{1}\int_{-1}^{1}\prod_{j=1}^{2}\frac{\exp[\frac{-T}{2k}(M_i^{(j)}-\bar{S}_i^{(j)})^2]}{\sqrt{2\pi\frac{k}{T}}}P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})d\bar{S}_i^{(1)}d\bar{S}_i^{(2)}, \tag{21}$$

where we assume that the measurement noise of Z_1Z_2 and Z_2Z_3 is uncorrelated and of equal variance. Parsing Eq. (21), we see that the measurement distribution has the form of a continuous Gaussian mixture, with each component centered at a different point in the 2D interval $[-1,+1]\times[-1,+1]$. The mixture components are weighted by the probability $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$ of observing those mean values, assuming that ℓ_i and ℓ_{i-1} are known. Figure 2 illustrates the Gaussian-like behavior of $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$, which is especially clear under strong noise. We exploit this similarity in our Bayesian algorithm to efficiently approximate the measurement log-likelihood as described in Sec. 5.1.

The fact that Eq. (21) is not simply a bivariate Gaussian distribution with a fixed mean reflects the uncertainty that is inherent in our problem. Since $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$ is conditioned only on the state of the system at iT and $(i+1)T$, we must predict the values of the syndrome averages using this information alone. Figure 1C demonstrates the difficulty of such an inference, as the value of $\bar{S}_i^{(j)}$ depends not only on the precise *number* of errors in the integration interval but more importantly on the *locations* of these errors within the interval, all of which are unknown to us. There are an infinite number of possible values that $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ can take for any given (ℓ_{i-1},ℓ_i) pair, so we must marginalize over these degrees of freedom in order to derive the measurement distribution.

To better understand where the probability of $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ is concentrated, we can condition $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$ on the number of errors occurring in the ith interval, which

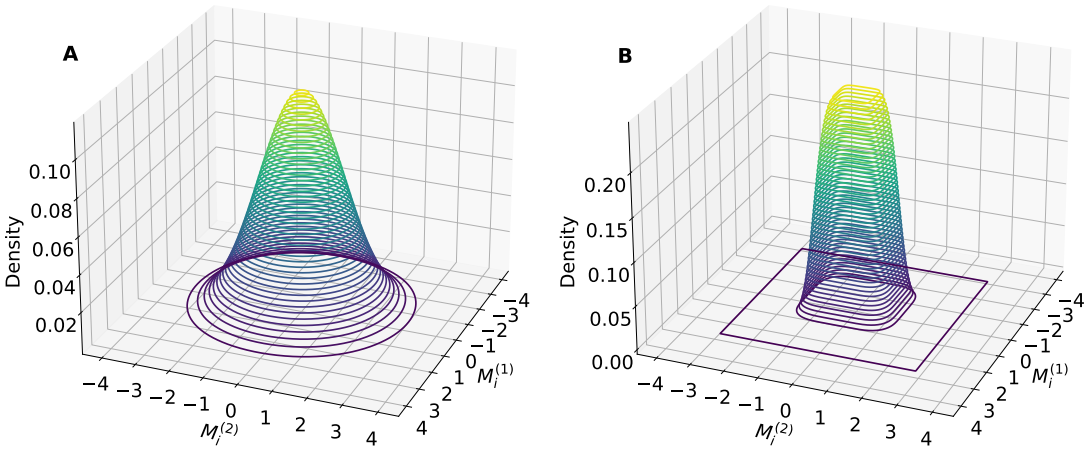

Figure 2: Contour plots of $P(M_i^{(1)} M_i^{(2)} | \ell_i \ell_{i-1})$ derived from Eq. (21) using a uniform distribution over the syndrome means $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$. Plot A was generated using a moderate amount of noise ($\frac{k}{T} = 1$), and possesses a Gaussian-like shape. Plot B contains far less noise ($\frac{k}{T} = 0.05$), so the square shape of the underlying uniform distribution starts to show through.

gives

$$\begin{aligned} P(\bar{S}_i^{(1)} \bar{S}_i^{(2)} | \ell_i \ell_{i-1}) &= \sum_{e_1,e_2,e_3=0}^{\infty} P(\bar{S}_i^{(1)} \bar{S}_i^{(2)} | \ell_i \ell_{i-1}; e_1 e_2 e_3) P(e_1 e_2 e_3 | \ell_i \ell_{i-1}) \\ &= \sum_{e_1,e_2,e_3=0}^{\infty} P(\bar{S}_i^{(1)} \bar{S}_i^{(2)} | \ell_{i-1}; e_1 e_2 e_3) P(e_1 | \ell_i \ell_{i-1}) P(e_2 | \ell_i \ell_{i-1}) P(e_3 | \ell_i \ell_{i-1}), \end{aligned} \tag{22}$$

where e_k is the number of errors occurring on the kth qubit during the ith interval (this time index is suppressed for notational convenience). In the final line we exploit the fact that errors on different qubits are independent, and that ℓ_i is a deterministic function of e_1, e_2, e_3, ℓ_{i-i} and therefore does not need to be explicitly conditioned on.

From our discussion in Sec. 3.2, we know that ℓ_i and ℓ_{i-1} only specify whether an even or odd number of errors occurred on each qubit, so $P(e_k | \ell_i \ell_{i-1})$ can be rewritten as $P(e_k|\text{odd})$ or $P(e_k|\text{even})$ depending on whether ℓ_i and ℓ_{i-1} differ in the kth qubit. These distributions are identical to the Poisson distribution from Eq. (16) except that the probability of either even or odd e_k is set to zero. After re-normalizing, we get

$$P(e_k|\text{odd}) = \begin{cases} \frac{(\mu T)^{e_k}}{e_k! \sinh(\mu T)} & e_k \text{ odd} \\ 0 & e_k \text{ even} \end{cases}, \qquad P(e_k|\text{even}) = \begin{cases} \frac{(\mu T)^{e_k}}{e_k! \cosh(\mu T)} & e_k \text{ even} \\ 0 & e_k \text{ odd} \end{cases}. \tag{23}$$

As an example of how these distributions are used, if $\ell_i = 4$ and $\ell_{i-1} = 0$ then the error distribution factorizes as $P(e_1 e_2 e_3 | 4, 0) = P(e_1|\text{odd}) P(e_2|\text{even}) P(e_3|\text{even})$.

Referring back to Eq. (22), the final term left to characterize is $P(\bar{S}_i^{(1)} \bar{S}_i^{(2)} | \ell_{i-1}; e_1 e_2 e_3)$, which describes how the syndrome averages are distributed given a starting state and the number of errors that occurred. This term is challenging to evaluate, since errors on the second qubit will flip both syndromes simultaneously and thus prevent the joint distribution

from factorizing. If we consider errors on the second qubit separately from errors on the first and third qubits, i.e., focus on cases where $e_1, e_3 = 0$ and where $e_2 = 0$, then we can derive (see Appendix A.1) the following expressions

$$P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}; e_1 0 e_3) = P(\bar{S}_i^{(1)}|\ell_{i-1}; e_1 00)P(\bar{S}_i^{(2)}|\ell_{i-1}; 00 e_3) \tag{24}$$

$$P(\bar{S}_i^{(1)}|\ell_{i-1}; e_1 00) = \begin{cases} \delta(1 \mp_1 \bar{S}_i^{(1)}) & e_1 = 0 \\ \dfrac{e_1!}{2a_1!b_1!}\left(\dfrac{1 \pm_1 \bar{S}_i^{(1)}}{2}\right)^{a_1}\left(\dfrac{1 \mp_1 \bar{S}_i^{(1)}}{2}\right)^{b_1} & e_1 > 0 \end{cases} \tag{25}$$

$$P(\bar{S}_i^{(2)}|\ell_{i-1}; 00 e_3) = \begin{cases} \delta(1 \mp_2 \bar{S}_i^{(2)}) & e_3 = 0 \\ \dfrac{e_3!}{2a_3!b_3!}\left(\dfrac{1 \pm_2 \bar{S}_i^{(2)}}{2}\right)^{a_3}\left(\dfrac{1 \mp_2 \bar{S}_i^{(2)}}{2}\right)^{b_3} & e_3 > 0 \end{cases} \tag{26}$$

$$P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}; 0 e_2 0) = \begin{cases} \delta(1 \mp_1 \bar{S}_i^{(1)})\delta(1 \mp_2 \bar{S}_i^{(2)}) & e_2 = 0 \\ \dfrac{e_2!}{2a_2!b_2!}\left(\dfrac{1 \pm_1 \bar{S}_i^{(1)}}{2}\right)^{a_2}\left(\dfrac{1 \mp_1 \bar{S}_i^{(1)}}{2}\right)^{b_2}\delta(\pm_1 \bar{S}_i^{(1)} \mp_2 \bar{S}_i^{(2)}) & e_2 > 0 \end{cases} \tag{27}$$

where a_k and b_k are the integer parts of $\frac{e_k}{2}$ and $\frac{e_k-1}{2}$ respectively, and $\delta(x)$ is the Dirac delta function. The "$\pm_j$" term indicates whether state ℓ_{i-1} has even "+" or odd "−" parity with respect to the jth measurement operator, and therefore whether the syndrome should be added or subtracted. We note that the syndrome density is uniform (on its support) when only a single error occurs (i.e., $a_k = b_k = 0$). This simplification will be utilized in Sec. 5.1 as we construct our logarithmic filter.

Unfortunately, when errors occur on the second qubit and on at least one of the other qubits, the joint distributions are non-smooth and appear to lack convenient analytic forms (see Appendix A.1.3). As an example, if the system starts in $|000\rangle$ and experiences one error on each qubit, then the syndrome distribution is given by

$$P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|0; 111) = \frac{1}{4}\left[\min(\bar{S}_i^{(1)}, \bar{S}_i^{(2)}) + \max(0, \bar{S}_i^{(1)} + \bar{S}_i^{(2)})\right], \tag{28}$$

which is a piecewise function with a discontinuous first derivative. Figure 3 shows a plot of $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|0; 111)$, together with $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|0; 011)$ which is even less smooth.

Due to the poor behavior of the syndrome densities, we are unable to make further progress toward an exact, analytic expression for the measurement probability $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$. Most of this difficulty stems from the fact that we must, in general, account for an infinite number of possible errors that could occur across all three qubits within a given integration interval. If, however, we restrict ourselves to situations where only a single error occurs, which is reasonable when μ and T are small, then the expressions in Eqs. (25 – 27) can be used directly and we will be able to evaluate the integrals in Eq. (21). This is the approach we take in Sec. 5.1, although it is no longer an exact treatment of the problem.

3.4 Constructing the Bayesian Filter

Although Sec. 3.3 suggests that the posterior distribution for the state tracking problem does not possess a convenient functional form, it is still useful to synthesize our results

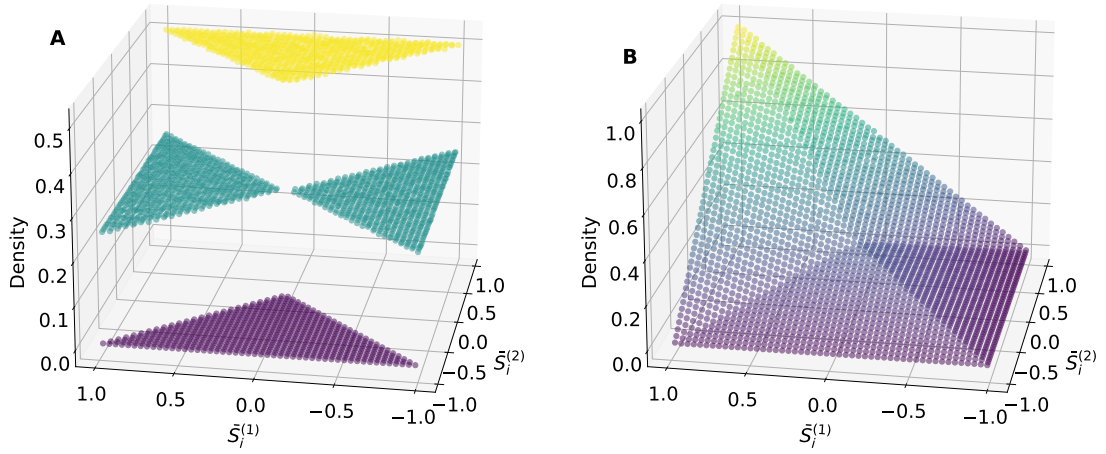

Figure 3: Plot A shows the probability density $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|0;011)$, which describes the syndrome means when the system starts in $|000\rangle$ and then experiences an error on the second and third qubits. Plot B shows $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|0;111)$, whose form is given in Eq. (28). As a general rule, the more errors that occur in the interval the more "smooth" the syndrome density will appear, with plot A being outright discontinuous while plot B has a discontinuous first derivative. These plots were generated from 10^8 samples using histograms with 50 equally spaced bins along each axis. The color gradient indicates the relative magnitude of each bin, with different scales for the two plots.

from this section into an algebraic description of how the optimal model would operate. Referring back to Eq. (15), the posterior state probabilities $\hat{P}(\ell_i)$ are given recursively in terms of the measurement probabilities $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ and transition probabilities $J_{\ell_{i-1}\ell_i} \equiv P(\ell_i|\ell_{i-1})$. Noting the dependence of both terms on ℓ_i and ℓ_{i-1}, the update rule is given by

$$\begin{aligned}\hat{P}(\ell_i) &\propto \sum_{\ell_{i-1}=0}^{7} \hat{P}(\ell_{i-1})P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})J_{\ell_{i-1}\ell_i} \\ &\propto \sum_{\ell_{i-1}=0}^{7} \hat{P}(\ell_{i-1})\tilde{J}(M_i^{(1)}M_i^{(2)})_{\ell_{i-1}\ell_i},\end{aligned} \tag{29}$$

where $\tilde{J}(M_i^{(1)}M_i^{(2)})_{\ell_{i-1}\ell_i} \equiv P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})J_{\ell_{i-1}\ell_i}$ can be viewed as a weighted transition matrix which combines the Markov dynamics of the error process with the likelihood of the observed measurement outcomes.

Due to the recursive nature of Eq. (29), the state tracking model operates naturally as a digital filter, taking in syndrome measurements at each time step and outputting a set of posterior probabilities that incorporate all of the information available to us. Setting aside the challenges of computing $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ from Eq. (21), this filter is optimal in the sense that it was constructed from the formal probability manipulations in Eq. (15), which assumed only that the measurements and errors were Markovian. In Sec. 3.3 we made further assumptions about the structure of the measurement signal, and insofar as these assumptions are valid there is nowhere for the algorithm to be improved. Of course, should the physical system not obey the idealized model outlined in Sec. 2.2, then the filter will need to be modified in order to remain optimal.

From a computational standpoint, Eq. (29) represents the filter as a simple matrix-vector product, with the difficulty centered on calculating $\tilde{J}$. More specifically, Sec. 3.3

made it clear that computing $P(M_i^{(1)} M_i^{(2)} | \ell_i \ell_{i-1})$ requires us to evaluate the integral in Eq. (21), which appears to be analytically intractable when the error combinations are not restricted. This limitation, together with other issues that will be discussed in Sec. 4, rules out using the optimal filter for real-time error tracking. That said, accurate numerical evaluation of the integral is still possible via sampling techniques, which are discussed further in Appendix A.2, so we will use this filter as a benchmark for other state tracking methods.

4 Obstacles and Prior Work

In the context of real-time quantum error tracking, where latency is measured in nanoseconds and hardware resources can be tight, the Bayesian filter outlined in Sec. 3 has four significant challenges:

1. The Gaussian integral in Eq. (21) cannot be evaluated analytically, and sampling methods are too resource-intensive for real-time filtering.

2. Computing even a single Gaussian requires the implementation of an exponential function, which can be challenging on low-latency, high-throughput devices such as FPGAs [16].

3. The outputs of the Gaussians will likely need to be represented using floating-point numbers to capture the necessary range and precision, which adds a further computational burden.

4. The probabilities generated by Eq. (29) will need to be periodically normalized in order to prevent overflow or underflow, which requires a dedicated division routine that will take up resources and be slow to run.

Overcoming all of these obstacles while preserving the underlying Bayesian framework is not trivial.

The first derivation of an optimal continuous-time filter for the three qubit bit-flip code was published by van Handel and Mabuchi, who recovered the well-known Wonham filter after solving for the least-squares estimator of the density matrix [17]. This classical filter was designed to output probabilities for the states of a Markov chain observed continuously by a signal with additive Gaussian noise. Using a modified version of our notation, the canonical Wonham filter has the form

$$d\hat{P}_\ell(t) = \sum_{\ell'=0}^{7} \hat{P}_{\ell'}(t) Q_{\ell'\ell} dt + \frac{1}{k} \sum_{j=1}^{2} [S_\ell^{(j)} - \bar{\bar{S}}^{(j)}(t)] \hat{P}_\ell(t) [dM^{(j)}(t) - \bar{\bar{S}}^{(j)}(t) dt], \tag{30}$$

where $\hat{P}_\ell(t)$ is the probability of the ℓth state at time t, $S_\ell^{(j)} \in \{-1, +1\}$ is the syndrome of the ℓth state from the jth parity operator, $\bar{\bar{S}}^{(j)}(t) \equiv \sum_\ell \hat{P}_\ell(t) S_\ell^{(j)}$ is the average of the jth syndrome at time t, and $Q = \frac{1}{T} \ln J$ is the rate matrix of the continuous Markov chain which describes the state transitions.

It is important to emphasize that Eq. (29) and Eq. (30) are each optimal with respect to different measurement regimes. The Wonham filter is optimal under the assumption that we have access to instantaneous signal readouts at arbitrary t, while our filter in Eq. (29) is optimal when the measurements are restricted to being integrated averages of the signal over some finite period T as in Eq. (9). As $T \to 0$ these two forms of measurement converge, and thus Eq. (29) converges to Eq. (30) after normalization (see Appendix A.3).

The general problem of filtering Markov jump processes using discrete observations has recently been explored by Borisov [18], who derives results that are in agreement with our work in Sec. 3.

As a non-linear stochastic differential equation, the Wonham filter is not practical to use in its exact form, though it is still possible to create discretized approximations of Eq. (30) which are effective. For example, Gange George et al. applied the Euler-Maruyama method to a logarithmic transformation of the Wonham filter in order to derive a first-order update rule that was numerically stable [19]. In the specific context of continous error correction, Mohseninia et al. [5] proposed a linearized version of Eq. (30) equivalent to

$$\hat{P}(\ell_i) \propto \hat{P}(\ell_{i-1} = \ell_i) + T \sum_{\ell_{i-1}=0}^{7} \left[Q_{\ell_{i-1}\ell_i} + \frac{M_i^{(1)} S_{\ell_{i-1}}^{(1)} + M_i^{(2)} S_{\ell_{i-1}}^{(2)}}{k} \delta_{\ell_{i-1}\ell_i} \right] \hat{P}(\ell_{i-1}), \quad (31)$$

where $S_{\ell_{i-1}}^{(j)}$ and Q are defined as in Eq. (30). The filter described by Eq. (31) does not involve any Gaussian functions and therefore avoids the first three obstacles in our list, but still requires periodic normalization. Additionally, as T grows larger the accuracy of this first-order approximation will decline.

The filters discussed so far have all been motivated at some level by Bayesian probability analysis, but there exists another class of algorithm, referred to as *threshold* or *boxcar* filters, which eschews connections to probability theory in favor of computational simplicity [20]. At a basic level, these models take a pair of integrated measurement values $(M_i^{(1)}, M_i^{(2)})$ and compare them to a set of predefined thresholds, after which an appropriate action is taken. Atalaya et al. developed a double-threshold algorithm for the Bacond-Shor code using non-commuting observables [4], while Mosheninia et al. proposed a non-Markovian boxcar filter and a double threshold scheme for the three-qubit repetition code which were easy to implement and attained good performance [5]. Atalaya and Zhang et al. applied a flexible double thresholding scheme to a system undergoing quantum annealing, demonstrating that these algorithms were effective at error correction even in the presence of Hamiltonian evolution [6].

5 A Practical Bayesian Filter

To address the obstacles identified in Sec. 4, we propose an algorithm that avoids any exponentiation or division operations, while also approximating the analytically intractable integral in Eq. (21). In Sec. 5.1 we simplify Eq. (21) by assuming that only a single error occurs during each integration period, which yields simple expressions for $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$. In Sec. 5.2 we eliminate the need for exponentiation and division by shifting our analysis into log-probability space, where the multiplication of probabilities becomes addition, normalization becomes subtraction, and Gaussian distributions transform into simple quadratic functions. Finally, Sec. 5.4 introduces what we call "single-term" and "two-term" strategies for approximating the LogSumExp functions needed for Markov chain evolution using log-probabilities. The performance of our algorithm is tested numerically in Sec. 6.

5.1 Single-error approximation

The largest obstacle to using Eq. (29) as a filter is that $P(M^{(1)}M^{(2)}|\ell_i\ell_{i-1})$ cannot be easily evaluated. As explored in Sec. 3.3, this difficulty arises because the syndrome density terms $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}; e_1e_2e_3)$ do not in general have convenient analytic forms, especially when the number of errors is large. This in turn prevents us from solving the integral in

Eq. (21). However, if we were to assume that only a single error can occur within a given measurement interval, then Eqs. (25 – 27) would reduce to

$$P(\bar{S}_i^{(1)}|\ell_{i-1};e_1 00) = \begin{cases} \delta(1 \mp_1 \bar{S}^{(1)}) & e_1 = 0 \\ \frac{1}{2} & e_1 = 1 \end{cases} \tag{32}$$

$$P(\bar{S}_i^{(1)}|\ell_{i-1};00 e_3) = \begin{cases} \delta(1 \mp_2 \bar{S}_i^{(2)}) & e_3 = 0 \\ \frac{1}{2} & e_3 = 1 \end{cases} \tag{33}$$

$$P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1};0 e_2 0) = \begin{cases} \delta(1 \mp_1 \bar{S}_i^{(1)})\delta(1 \mp_2 \bar{S}_i^{(2)}) & e_2 = 0 \\ \frac{1}{2}\delta(\pm_1 \bar{S}_i^{(1)} \mp_2 \bar{S}_i^{(2)}) & e_2 = 1, \end{cases} \tag{34}$$

which all lead to tractable integrals when substituted into Eq. (21). This single-error approximation, which is reasonable when the average number of errors per interval, μT, is small, leads to Gaussian-like measurement distributions which can be easily incorporated into our log-probability filter.

For simplicity, we first consider an error that occurs on either the first ($e_1 = 1$) or third ($e_3 = 1$) qubit, since it will affect only one of the syndromes and thus allow the joint distribution to separate as in Eq. (24). By substituting $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1};100)$ into Eq. (21) and requiring that ℓ_i and ℓ_{i-1} differ only in the first qubit, we get

$$P(M_i^{(1)}M_i^{(2)}|\ell_{i-1};100) = \frac{1}{4}\left[\operatorname{erf}\left(\frac{M_i^{(1)}+1}{\frac{2k}{T}}\right) - \operatorname{erf}\left(\frac{M_i^{(1)}-1}{\frac{2k}{T}}\right)\right] \cdot \frac{\exp[\frac{-T}{2k}(M_i^{(2)} \mp_2 1)^2]}{\sqrt{2\pi\frac{k}{T}}}, \tag{35}$$

where $\operatorname{erf}(x) \equiv \frac{2}{\sqrt{\pi}}\int_0^x e^{-y^2}dy$ is the error function. The last line of Eq. (35) consists of two distinct terms, one of which is a function of $M_i^{(1)}$ and the other of $M_i^{(2)}$. The $M_i^{(2)}$ term is simply a Gaussian centered at $\pm_2 1$, since the second syndrome is not sensitive to an error on the first qubit. The $M_i^{(1)}$ term, by contrast, arises because the mean of the first syndrome is distributed uniformly over the range $[-1,+1]$, which yields a pair of error functions. Notably, this term does not depend on $\pm_1$, so it is independent of the state of the first two qubits.

While the first term in Eq. (35) is an exact result from of our single-error integration, it is not very convenient to evaluate numerically. As described in Appendix A.4, the measurement distribution can be approximated as

$$P(M_i^{(1)}M_i^{(2)}|\ell_{i-1};100) = \frac{\exp[\frac{-3}{2(1+3\frac{k}{T})}(M_i^{(1)})^2]}{\sqrt{2\pi(\frac{1}{3}+\frac{k}{T})}} \cdot \frac{\exp[\frac{-T}{2k}(M_i^{(2)} \mp_2 1)^2]}{\sqrt{2\pi\frac{k}{T}}}, \tag{36}$$

which will be easy to evaluate numerically after we shift to log-probabilities in Sec. 5.2. The measurement density for an error on the third qubit can be obtained from Eq. (36) by simply swapping $M_i^{(1)}$ and $M_i^{(2)}$ and replacing $\mp_2$ with $\mp_1$. Carrying out a similar procedure on the second qubit gives

$$P(M_i^{(1)}M_i^{(2)}|\ell_{i-1};010) = \frac{1}{2}\frac{\exp[\frac{-T}{k}(\frac{M_i^{(1)} \mp_c M_i^{(2)}}{2})^2]}{\sqrt{\pi\frac{k}{T}}} \cdot \frac{\exp[\frac{-3}{2(1+3\frac{k}{2T})}(\frac{M_i^{(1)} \pm_c M_i^{(2)}}{2})^2]}{\sqrt{2\pi(\frac{1}{3}+\frac{k}{2T})}}, \tag{37}$$

where $\pm_c$ is positive when ℓ_{i-1} has the same parity with respect to Z_1Z_2 and Z_2Z_3 but negative when the parities are opposite (roughly, "$\pm_c \equiv \pm_1 \cdot \pm_2$").

The final case to consider is when no errors occur in an interval at all. With reasonable values of μ and T this is the most likely outcome for any given interval, and it's measurement distribution can be solved for easily by substituting $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1}) = P(\bar{S}_i^{(1)}|\ell_{i-1};000)P(\bar{S}_i^{(2)}|\ell_{i-1};000)$ into Eq. (21). This gives

$$P(M_i^{(1)}M_i^{(2)}|\ell_{i-1};000) = \frac{\exp[\frac{-T}{2k}(M_i^{(1)} \mp_1 1)^2]}{\sqrt{2\pi\frac{k}{T}}} \cdot \frac{\exp[\frac{-T}{2k}(M_i^{(2)} \mp_2 1)^2]}{\sqrt{2\pi\frac{k}{T}}}, \tag{38}$$

which is simply the product of Gaussian distributions with means corresponding to the parities of ℓ_{i-1}.

Taken together, Eqs. (36, 37, 38) constitute the approximate description of $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ that we will use to construct our log-probability filter. The Gaussian form of each equation is especially important on a practical level, as the log-densities will all reduce to sums of simple quadratic equations that can be easily computed.

5.2 Moving to logarithmic probability

To begin the transformation from probabilities to log-probabilities, we take the logarithm of the optimal Bayesian filter from Eq. (29)

$$\begin{aligned}\log\hat{P}(\ell_i) &= \log\sum_{\ell_{i-1}=0}^{7}\hat{P}(\ell_{i-1})P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})J_{\ell_{i-1}\ell_i} \\ &= \log\sum_{\ell_{i-1}=0}^{7}\exp[\log\hat{P}(\ell_{i-1}) + \log\tilde{J}(M_i^{(1)}M_i^{(2)})_{\ell_{i-1}\ell_i}],\end{aligned} \tag{39}$$

where in the last line we exponentiate the logarithm of each term in the sum in order to preserve the recursive structure of the filter. Since $\log\tilde{J}_{\ell_{i-1}\ell_i} \equiv \log J_{\ell_{i-1}\ell_i} + \log P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$, the terms in Eq. (39) can be evaluated using our results from Sec. 3.2 and Sec. 5.1.

The transition log-probability is given by the logarithm of Eq. (19), which yields

$$\log J_{\ell_{i-1}\ell_i} = d(\ell_i,\ell_{i-1})\log\sinh(\mu T) + [3 - d(\ell_i,\ell_{i-1})]\log\cosh(\mu T) - 3\mu T \tag{40}$$

where $d(\ell_i,\ell_{i-1})$ is the Hamming distance between ℓ_i and ℓ_{i-1}. While the form of Eq. (40) is not especially illuminating, the value of μT will be known to us and assumed to be constant across an experiment. We are therefore able to compute the values of $\log J_{\ell_{i-1}\ell_i}$ in advance and use them in the filter without needing to evaluate the log sinh or log cosh functions in real time.

The measurement log-probabilities, by contrast, must be calculated anew each time a measurement is recorded. Such rapid computation is feasible because Eqs. (36, 37, 38) are

all in Gaussian form, and therefore have the following quadratic log-probabilities

$$\log P(M_i^{(1)} M_i^{(2)} | \ell_{i-1}; 000) = -\frac{T}{2k}[(M_i^{(1)} \mp_1 1)^2 - (M_i^{(2)} \mp_2 1)^2] + \log N \tag{41}$$

$$\log P(M_i^{(1)} M_i^{(2)} | \ell_{i-1}; 100) = -\frac{3}{2(1+3\frac{k}{T})}(M_i^{(1)})^2 - \frac{T}{2k}(M_i^{(2)} \mp_2 1)^2 + \log N' \tag{42}$$

$$\log P(M_i^{(1)} M_i^{(2)} | \ell_{i-1}; 010) = -\frac{T}{k}(\frac{M_i^{(1)} \mp_c M_i^{(2)}}{2})^2 - \frac{3}{2(1+3\frac{k}{2T})}(\frac{M_i^{(1)} \pm_c M_i^{(2)}}{2})^2 + \log N'' \tag{43}$$

$$\log P(M_i^{(1)} M_i^{(2)} | \ell_{i-1}; 001) = -\frac{T}{2k}(M_i^{(1)} \mp_1 1)^2 - \frac{3}{2(1+3\frac{k}{T})}(M_i^{(2)})^2 + \log N''', \tag{44}$$

where the normalization constants have been left unspecified for convenience. The different variances, all functions of k and T, are constant during an experiment and therefore do not need to be computed at each step.

Using Eqs. (40 – 44), we can calculate $\log \tilde{J}$ and therefore evaluate the exponent of each term in the sum of Eq. (39). The last line of Eq. (39) is known as a *LogSumExp* [21], and its evaluation is the final obstacle to implementing our filter. For convenience we adopt the more compact notation $L_{\ell_{i-1}}^{\ell_i} \equiv \log \tilde{J}_{\ell_{i-1}\ell_i} + \log \hat{P}(\ell_{i-1})$, which serve as the inputs of the LogSumExp function,

$$\log \hat{P}(\ell_i) = \log \sum_{\ell_{i-1}=0}^{7} \exp[L_{\ell_{i-1}}^{\ell_i}]. \tag{45}$$

Using this form, we exploit a well-known trick for evaluating LogSumExp functions by pulling the largest exponential outside of the sum

$$\log \sum_{\ell_{i-1}=0}^{7} \exp[L_{\ell_{i-1}}^{\ell_i}] = L_*^{\ell_i} + \log(1 + \sum_{\ell_{i-1} \neq *} \exp[L_{\ell_{i-1}}^{\ell_i} - L_*^{\ell_i}]), \tag{46}$$

where $L_*^{\ell_i}$ is the largest input to the LogSumExp function. The second term in Eq. (46) is now the logarithm of one plus a set of terms whose exponents are all non-positive. Figure 4 shows that as the difference between $L_*^{\ell_i}$ and the other inputs grows, the output of Eq. (46) will converge to $L_*^{\ell_i}$. If the difference is large for some subset of the inputs, then the expression can be simplified by removing that subset from the sum.

5.3 Evaluating the LogSumExp

To take advantage of Eq. (46), we must know the relative magnitudes of the various $L_{\ell_{i-1}}^{\ell_i}$ terms in a typical run. For convenience, we introduce the non-positive quantity $\Delta L_{\ell_{i-1}}^{\ell_i} \equiv L_{\ell_{i-1}}^{\ell_i} - L_*^{\ell_i}$, which is simplified to just ΔL when the specific indices are not relevant. Using Figure 4 as a guide, it is clear that when

$$\Delta L \leq -4 \tag{47}$$

the magnitude of $\log(1+e^{\Delta L})$ is negligible. This provides a threshold to determine whether a given $L_{\ell_{i-1}}^{\ell_i}$ contributes significantly to Eq. (46) or if it can instead be reasonably ignored.

In Figure 5, we plot the average values of $L_{\ell_{i-1}}^{\ell_i}$ for $\ell_i = 0$ and $\ell_i = 1$ before and after a bit-flip error occurs on a state initialized to $|000\rangle$. These values are taken from

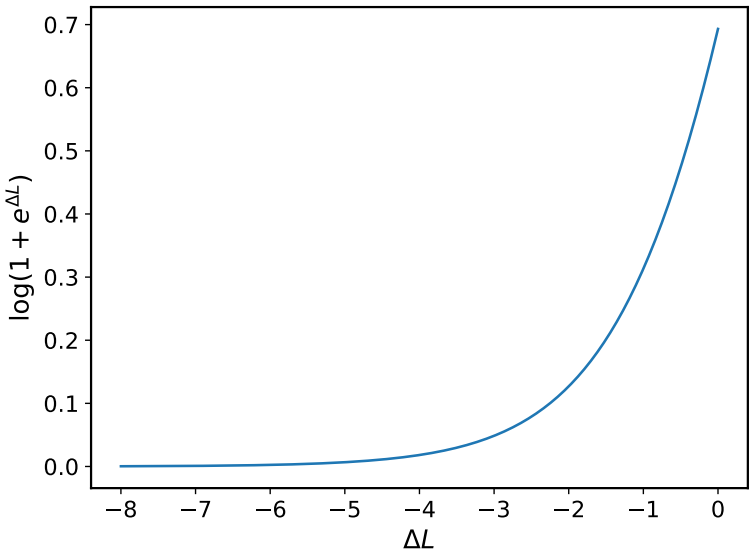

Figure 4: Plot of $\log(1+e^{\Delta L})$, which goes to zero as the magnitude of $\Delta L \equiv L^{\ell_i}_{\ell_{i-1}} - L^{\ell_i}_{*}$ increases. This implies that $L^{\ell_i}_{*} + \log(1+e^{\Delta L})$ rapidly approaches $L^{\ell_i}_{*}$ as the gap between $L^{\ell_i}_{*}$ and the other terms increases. Note that ΔL is never positive, since $L^{\ell_i}_{*}$ is the largest of all $L^{\ell_i}_{\ell_{i-1}}$ by definition.

the optimal filter in Eq. (29), which we simulate numerically using the method described in Appendix A.2. The plots of Figure 5 demonstrate clearly that only one $L^{\ell_i}_{\ell_{i-1}}$ term is significant when ℓ_i is equal to the true system state and two of the terms are significant when ℓ_i is separated from the true state by one bit-flip. For example, the gap ΔL between L^0_0 and the second-largest $L^{\ell_i}_{\ell_{i-1}}$ in plot 5A is roughly -15 before the bit-flip, which is far smaller than the cutoff threshold of -4 described in Eq. (47). Plots of the other ℓ_i indicate that states separated from the true state by two bit-flips have four significant terms and the state separated by three bit-flips has eight significant terms.

To understand the behavior shown in Figure 5, it is necessary to determine the typical magnitudes of Eq. (40) and Eqs.(41 – 44) under realistic conditions. Starting first with the measurement log-densities, we are interested in their expected values when the system is in some state ℓ', so we take the average of Eqs.(41 – 44) with respect to measurements distributed as $P(M_i^{(1)}M_i^{(2)}|\ell'\ell')$. For convenience we label this quantity as

$$G^{\ell_i}_{\ell_{i-1}} \equiv E[\log P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})], \tag{48}$$

and after some algebra the average log-densities are shown to scale as

$$G^{\ell_i}_{\ell_{i-1}} = \begin{cases} \mathcal{O}(-\log\frac{T}{k}) & \frac{T}{k} \ll 1, \text{ all } \ell_i, \ell_{i-1} \\ \mathcal{O}(-\log\frac{T}{k}) & \frac{T}{k} \gg 1, \ \ell_i \text{ or } \ell_{i-1} \in \{\ell', \ell' \oplus 7\} \\ \mathcal{O}(-\frac{T}{k}) & \frac{T}{k} \gg 1, \ \ell_i \text{ or } \ell_{i-1} \notin \{\ell', \ell' \oplus 7\}, \end{cases} \tag{49}$$

where $\oplus$ is the bitwise exclusive-or and $\frac{T}{k}$ is the *precision* of the noise (reciprocal of its variance). In words, Eq. (49) states that the log-densities all have a similar scale when the noise is large ($\frac{T}{k} \ll 1$) but begin to diverge when the noise is small ($\frac{T}{k} \gg 1$), with transitions to/from ℓ' and its complement having a significantly greater likelihood. This is expected, as the measurement densities differ primarily in the locations of their mean values, which become more obvious as the widths of the distributions shrink. With respect

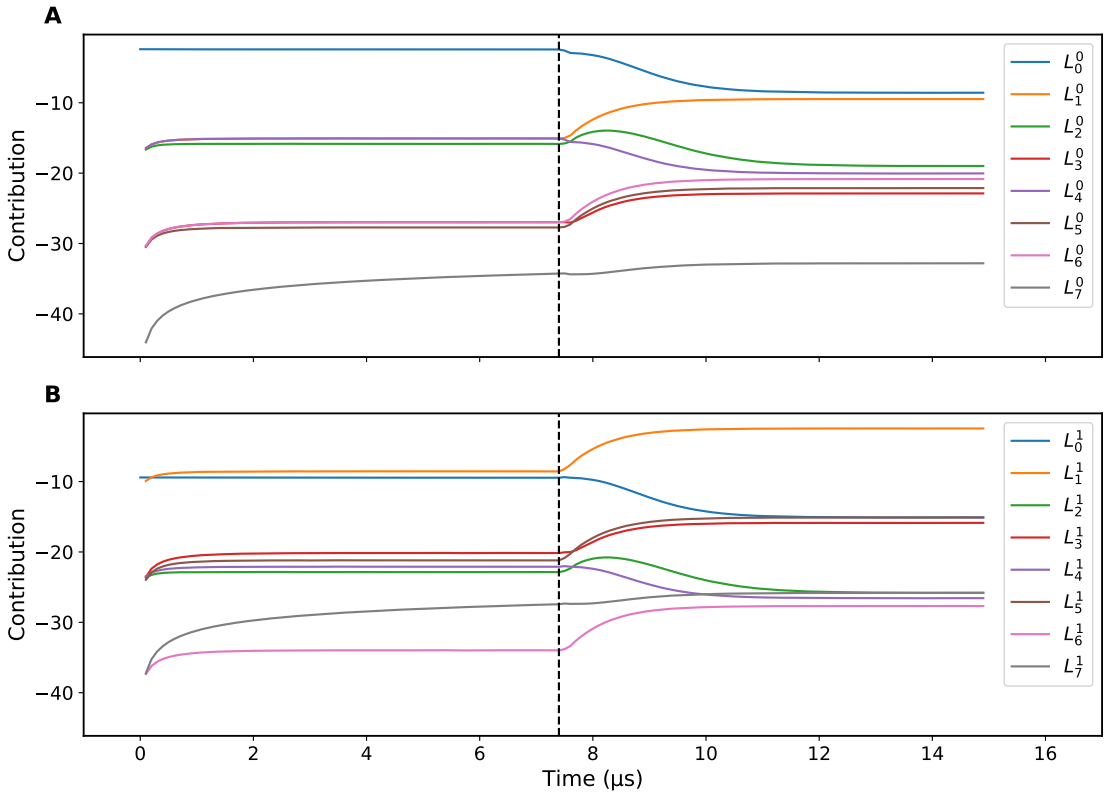

Figure 5: Plots of average $L_{\ell_{i-1}}^{\ell_i}$ contributions to the exponent of Eq. (46) for $|000\rangle$ (plot A) and $|001\rangle$ (plot B) taken from the optimal filter during a 15 μs run, with $\frac{k}{T} = 4$ and a state that was initialized to $|000\rangle$. The dashed black lines mark the occurrence of a bit-flip error at 7.5 μs on the third qubit, which takes the system to $|001\rangle$. The plotted data represent averages over the measurement noise taken from 10^5 different runs. When the system is $|000\rangle$, plot A shows clearly that only L_0^0 (blue line) is significant, while plot B shows that both L_0^1 (blue line) and L_1^1 (orange line) are significant. After the bit-flip occurs the state is $|001\rangle$ and the situation is reversed, with L_1^1, L_0^0, and L_1^0 contributing significantly.

to what is shown in Figure 5, Eq. (49) indicates that unless the noise magnitude is very small, the $L^{\ell_i}_{\ell_{i-1}}$ will not differ greatly from one another based on their $G^{\ell_i}_{\ell_{i-1}}$ values.

Turning our attention to the transition probabilities, we know from the numerical tests in Sec. 6.2 that when μT is on the order of 10^{-2} it becomes impossible to accurately track the state of the system at moderate run times, even for the optimal filter. We can therefore expand Eq. (40) for small μT as

$$\log J_{\ell_{i-1}\ell_i} = d(\ell_i, \ell_{i-1}) \log(\mu T) + \mathcal{O}(\mu T), \tag{50}$$

and keep only the leading μT term. For fixed μT, the values of $\log J_{\ell_{i-1}\ell_i}$ differ only in how many factors of $\log(\mu T)$ are included based on the Hamming distance $d(\ell_i, \ell_{i-1})$. Since $\log(10^{-2}) \approx -4.6$, the magnitude of a given $L^{\ell_i}_{\ell_{i-1}}$ decreases rapidly with every bit-flip that separates ℓ_i from ℓ_{i-1}. Indeed, $\log(10^{-2})$ lies below the threshold identified in Eq. (47), so if two $L^{\ell_i}_{\ell_{i-1}}$ terms differ by this amount then only the larger of the two will be significant.

Using Eq. (50) and the fact that $L^{\ell_i}_{\ell_{i-1}} \equiv G^{\ell_i}_{\ell_{i-1}} + \log J_{\ell_{i-1}\ell_i} + \log \hat{P}(\ell_{i-1})$, it is possible to explain the broad patterns observed in Figure 5. Beginning with the contributions to $|000\rangle$, the magnitude of L^0_0 (blue line in plot 5A) will dominate, since the system begins in $|000\rangle$ (large prior) and has a high probability of remaining in $|000\rangle$ (large transition element). All other L^0_{i-1} will involve at least one bit-flip and therefore will not be significant relative to L^0_0. By contrast, among the contributions to $|001\rangle$ (plot 5B), both L^1_0 and L^1_1 (blue and orange lines respectively) are significant, as each term has one large component (prior for L^1_0 and transition element for L^1_1) and one small component (transition element for L^1_0 and prior for L^1_1). After the bit-flip occurs the measurement likelihoods begin to favor $|001\rangle$ over $|000\rangle$, so the prior terms all shift toward $|001\rangle$ as well.

5.4 Single-term and two-term filters

The combined results of Sec. 5.2 and Sec. 5.3 offer a straightforward recipe for implementing the log-probability filter of Eq. (39). First, the measurement values are passed through the quadratic expressions in Eqs. (41–44) to compute the measurement log-densities $\log P(M^{(1)}_1 M^{(2)} | \ell_i \ell_{i-1})$, which are then added to the transition log-probabilities $\log J_{\ell_{i-1}\ell_i}$ and log-priors $\log \hat{P}(\ell_{i-1})$ to generate the $L^{\ell_i}_{\ell_{i-1}}$ terms. These terms must then be fed into Eq. (46), but we are free to choose how many of them to include based on their relative magnitudes. It is here that the algorithm splits into two different paths depending on whether we favor accuracy or simplicity.

If we desire a filter that is highly accurate, then we can take the two largest values of $L^{\ell_i}_{\ell_{i-1}}$ for each ℓ_i and substitute them into Eq. (46). In this scenario the LogSumExp reduces to

$$\log \hat{P}(\ell_i) = \log \sum_{\text{pair}} \exp[L^{\ell_i}_{\ell_{i-1}}] = L^{\ell_i}_* + \log(1 + \exp[\Delta L]), \tag{51}$$

where ΔL is the (negative) difference between the largest and second-largest $L^{\ell_i}_{\ell_{i-1}}$ terms. While Eq. (51) can be inconvenient to evaluate exactly, methods have been developed in the context of logarithmic number systems that utilize lookup tables and various interpolation schemes to yield efficient and accurate estimates [22]. By keeping the two largest values of $L^{\ell_i}_{\ell_{i-1}}$ we guarantee that the log-probabilities of the true state and its three nearest states are all correctly propagated into the next time step. The log-probabilities of the other four states will not be computed as accurately, although this is acceptable since these states are not involved in the single-flip transitions which the three-qubit code is designed to detect.

Although the logarithm term in Eq. (51) can be evaluated in a reasonable manner, it still represents an extra computational step that increases the overall complexity of the

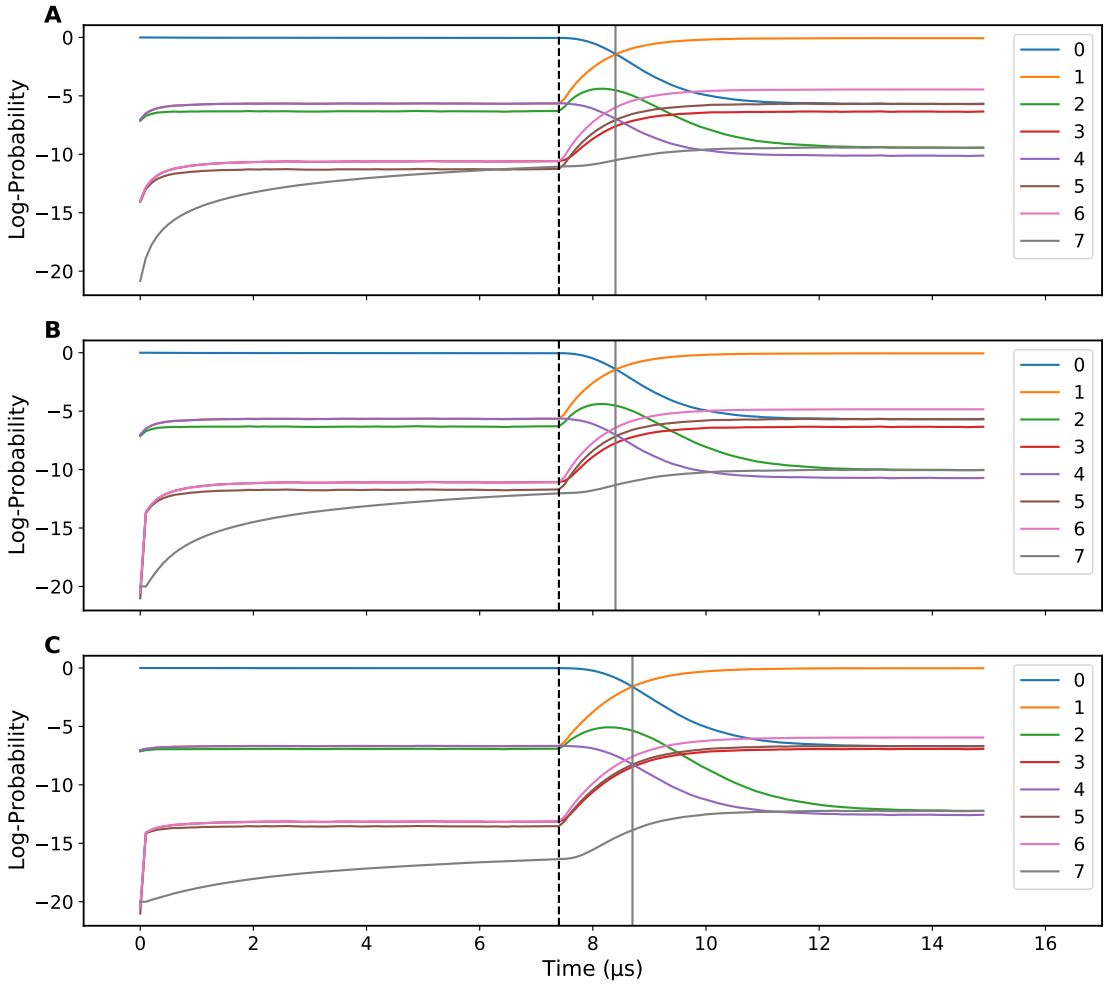

Figure 6: Plots of average posterior log-probabilities (normalized) for the optimal filter (plot A, Eq. (39)), two-term filter (plot B, Eq. (51)), and single-term filter (plot C, Eq. (52)), generated from the same parameters and measurement set used in Figure 5. The solid gray line indicates the time at which the filter correctly identifies that an error has occurred on the third qubit (dashed black line marks the time of the error). The optimal and two-term filter show almost identical behavior, while the single-term filter takes longer to detect the error.

algorithm. If simplicity is valued over accuracy, then we can alternatively choose to keep only the largest of the $L^{\ell_i}_{\ell_{i-1}}$ after each time step. In this extreme case the LogSumExp collapses to

$$\log \hat{P}(\ell_i) = \log \exp[L^{\ell_i}_*] = L^{\ell_i}_*, \tag{52}$$

and $L^{\ell_i}_*$ simply becomes the new posterior. Figure 5 shows that when ℓ_i equals the true state only one $L^{\ell_i}_{\ell_{i-1}}$ will be significant, so nothing important is being ignored there. For states adjacent to the true state, however, we will necessarily be ignoring one of the two significant contributions. While this single-term simplification appears quite restrictive, we show in Sec. 6 that Eq. (52) achieves high accuracy, though as expected it performs somewhat worse than the two-term approach.

In Figure 6 we plot the average evolution of the posterior log-probabilities in the presence of an error using outputs from the optimal filter versus those of the single-term and two-term filters. The plots were generated from the same data used to create Figure 5, and

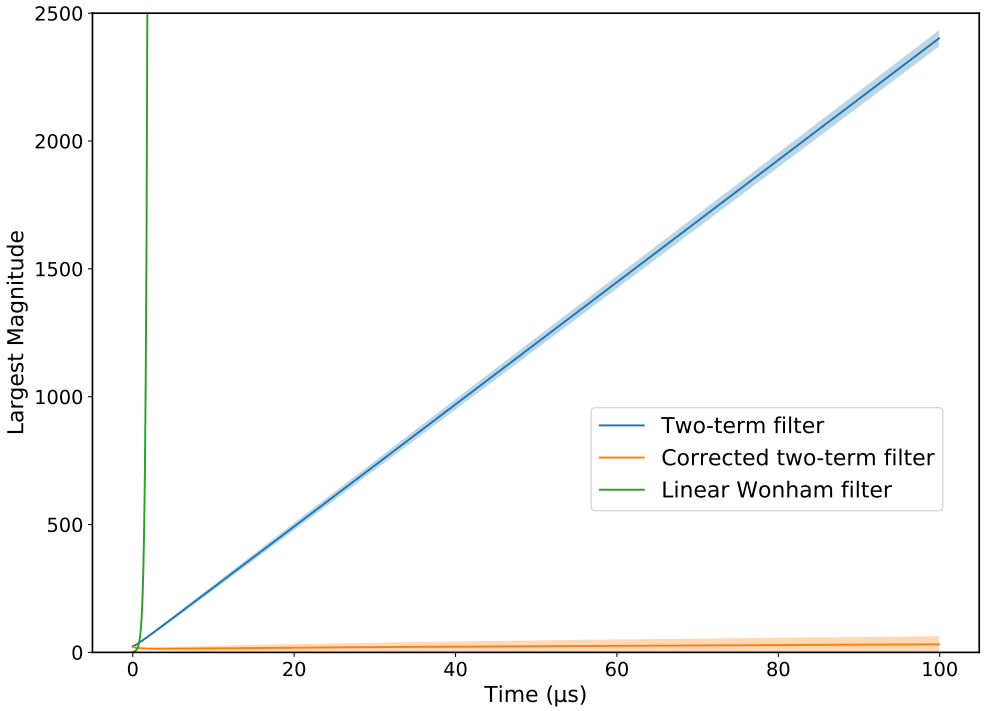

Figure 7: This plot tracks the magnitude of the largest probability output as a function of time for the linear Wonham filter and two different versions of our two-term filter (the single-term filter behaves identically), with $T = 0.1$ μs, $k = 0.4$ μs, and $\mu = 0$. The solid lines are averages across 10^5 trials, while the shaded regions enclose one standard deviation (omitted for the Wonham filter). The blue line records the magnitude of the two-term filter without any normalization, demonstrating a clear linear pattern. The orange line shows the same algorithm with the analytic correction factor from Eq. (53) included, which dramatically reduces the largest magnitude (mean and standard deviation of about 30 after 100 μs).

clearly show that all three filters are able to detect the error on the third qubit. The single-term filter performs worst as expected, taking longer to identify the error. Impressively, the two-term filter behaves almost identically to the optimal filter, with small discrepancies emerging for $\ell_i \in \{3, 5, 6, 7\}$ since these states all have more than two significant LogSumExp terms.

5.5 Normalization

As discussed in Sec. 4, one of the practical challenges of using Bayesian methods is the need for periodic normalization of the probability outputs, without which the values will grow or shrink rapidly. This behavior is shown in Figure 7, where the outputs of the linearized Wonham filter from Eq. (31) grow exponentially with time. Since exponential changes become linear on a logarithmic scale, the outputs of our logarithmic filters instead scale linearly with time. In situations where error tracking needs to be done for only a short duration, the linear growth of the outputs is likely tolerable and can simply be ignored. This represents a significant improvement over the linearized Wonham filter, which must incorporate a costly normalization routine to be viable over virtually any time scale.

For longer runs, or in cases where the range of output values is limited due to hardware restrictions, we may wish to slow the growth of the output magnitudes even further. This can be achieved by calculating the average rate of change per time step and simply

subtracting this quantity at the end of each step. The average change Δ is given by

$$\Delta = -[1 + \log(2\pi \frac{k}{T}) - 3 \log \cosh(\mu T) + 3\mu T], \tag{53}$$

which is the sum of $\log J_{\ell\ell}$ and G^{ℓ}_{ℓ} when the true state of the system is ℓ (since this is the only significant contribution). The corrected outputs are represented by the orange line in Figure 7. Applying this correction greatly slows the growth of the unnormalized log-probabilities, which allows the filters to be run within a very narrow numerical range.

6 Performance

To evaluate the performance of our single-term and two-term filters in an error-correction setting, we simulate a large number of trajectories (see Appendix A.5) and then record how accurately the filters are able to correctly identify the final state. For our definition of "accuracy" we adopt a binary measure which is 1 when the filter predicts either the true state or a state that differs from the true state by a single bit-flip, and is 0 otherwise. At any given time there will be four states considered to be correct and another four states considered to be incorrect, which gives an expected accuracy of 50% when guessing randomly. We choose this measure of accuracy because errors on a single qubit can be corrected via simple majority vote, while errors on two or more qubits will signify a logical error which the filter was supposed to have prevented.

In the following subsections, we present and discuss the performance of the filters under different run durations (Sec. 6.1), error rates μ (Sec. 6.2), and time steps T (Sec. 6.3). In all simulations we set $k = 0.4$ μs, and the state was initialized to $|000\rangle$. For reference, the performances of our single-term and two-term filters are plotted alongside those of the optimal filter from Eq. (29) to assess the practical impact of the simplifying assumptions made in Sec. 5.4. We also include plots for the linearized Wonham filter from Eq. (31) and the double-threshold scheme from Atalaya and Zhang et al. [6] to see how well our filters perform relative to existing algorithms.

6.1 Experiment duration

Figure 8 shows the performance of our single-term and two-term filters alongside the reference algorithms as a function of run duration, with the longest runs lasting a full millisecond. The accuracy is a decreasing function of time, since the probability of a logical error will increase as the run grows longer. Mohseninia et al. [5] showed that the accuracy will decrease linearly with run duration at small error rates, a pattern that is clearly visible in our data across all filters.

As expected, the optimal filter achieves the highest accuracy, though it is almost exactly matched by the two-term filter. Despite its greater simplicity, the single-term filter performs within a percentage point of the optimal model. By contrast, the linear Wonham filter and double threshold algorithm both perform significantly worse, though for different reasons. The linearized Wonham filter is constructed explicitly from a first-order approximation, which is exact as $T \to 0$ but suboptimal for any finite step size. The double threshold scheme, by contrast, is simply suboptimal by design, as it does not directly incorporate the error probabilities or the Gaussian distribution of the noise into its filter.

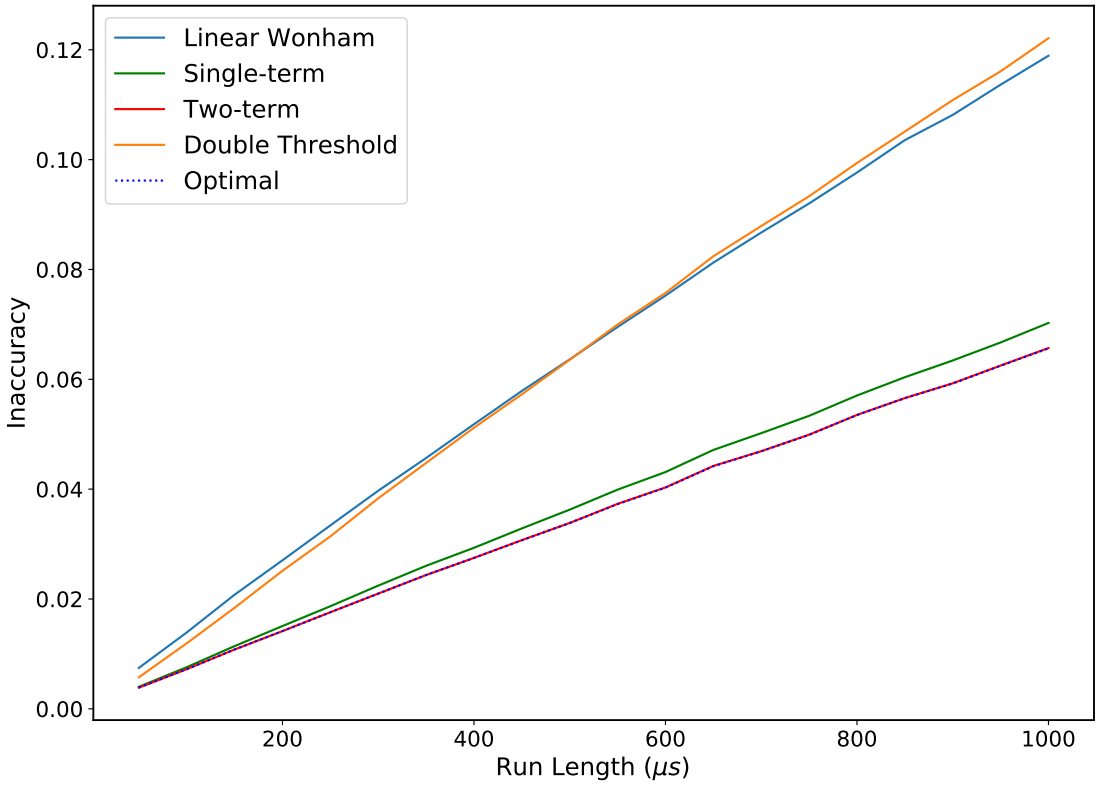

Figure 8: Plot of average inaccuracy (1 - accuracy) of the five different filters as a function of run time, with $k = 0.4$ μs and an initial state of $|000\rangle$. The error rate was fixed at $\mu = 2.5 \times 10^{-3}$ $(\mu\text{s})^{-1}$ and the time step at $T = 10^{-1}$ μs. The accuracy of the two-term filter is effectively equal to that of the optimal filter, such that the two curves completely overlap, while the single-term version performs slightly worse. The linear Wonham filter and double threshold performed comparably, though both did significantly worse than the logarithmic filters.

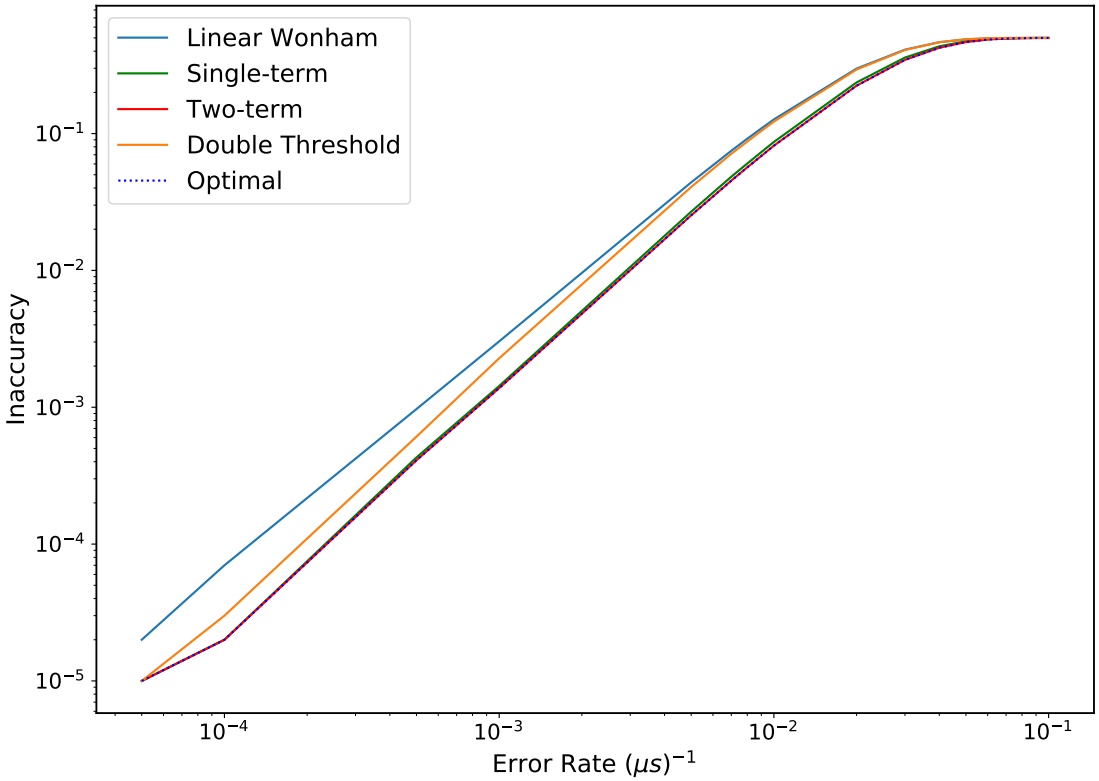

Figure 9: Plot of average inaccuracy (1 - accuracy) of the five filters as a function of error rate, with both axes on a log scale. The duration of the runs was fixed at 100 μs and the time step at $T = 10^{-1}$ μs, with $k = 0.4$ μs and an initial state of $|000\rangle$. The accuracy of the two-term filter is effectively equal to that of the optimal filter (curves overlap completely), with the single-term version performing slightly worse. The linear Wonham filter performed worse than the double threshold, with the gap widening at low error rates.

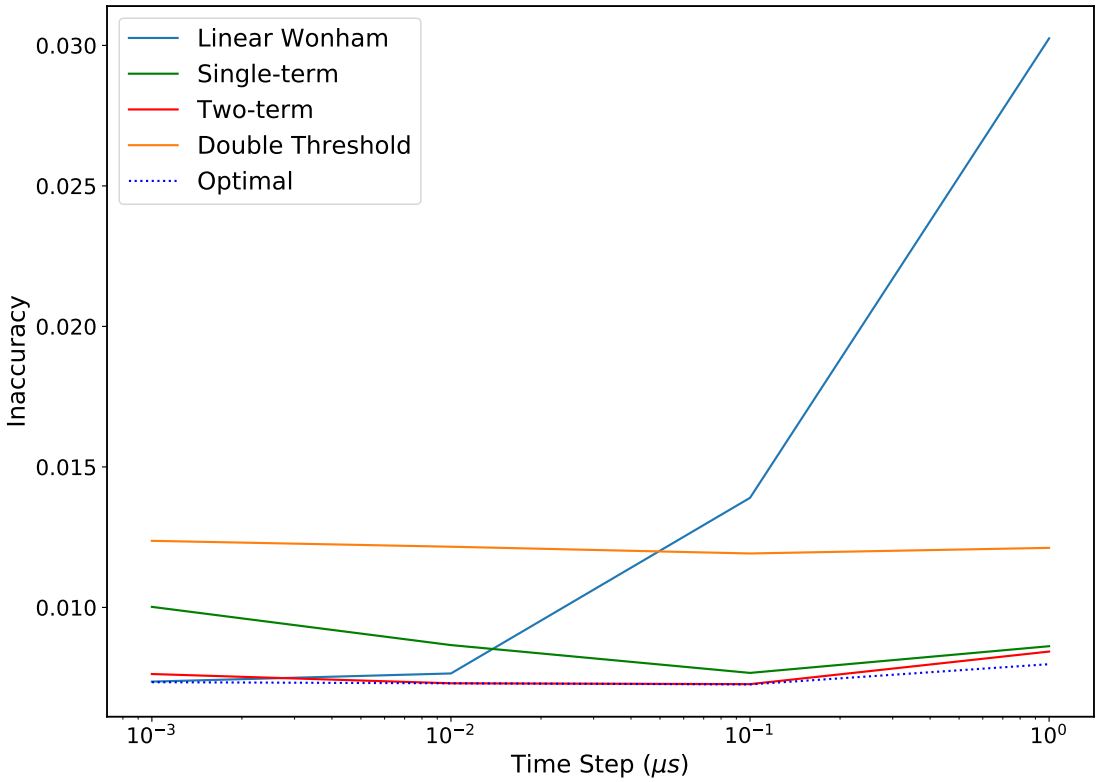

Figure 10: Plot of average inaccuracy (1 - accuracy) of the five different filters as a function of time step T, with $k = 0.4$ μs and an initial state of $|000\rangle$. The duration of the run was fixed at 100 μs and the error rate at $\mu = 2.5 \times 10^{-3}$ $(\mu s)^{-1}$. The accuracy of the two-term filter matches that of the optimal curve at most step sizes, but performs slightly worse at very small and very large values of T. The single-term filter shows a significant improvement as the time step increases, while the accuracy of the linear Wonham filter decreases rapidly. The double threshold appears largely unaffected by the step size.

6.2 Error rate

Figure 9 shows the accuracy of the five filters as a function of the per-qubit bit-flip error rate μ. On a log-log plot the dependence between the accuracy and error rate is linear for small error rates, but levels off at 0.5 as the error rate approaches 0.1 $(\mu s)^{-1}$. In this high-error regime the system experiences enough logical errors that all memory of the initial state is lost, and the system is therefore unable to distinguish the true state from its complement. As before, the optimal and two-term filters perform almost identically, while the single-term term algorithm is slightly less accurate. The negative impact of non-zero T on the linear Wonham filter is especially obvious at small error rates, where the accuracies of the other filters all converge together while the Wonham filter performs significantly worse.

6.3 Time step

Lastly, Figure 10 shows the accuracy of the filters as a function of the measurement integration time T, which reveals many different trends. The time steps range from 1 ns to 1 μs, mapping to noise variances on the order of 100 to 0.1 respectively, which allows us to explore the behavior of these filters under both high-noise and low-noise conditions.

Focusing first on the double-threshold scheme, it is clear that the step size does not have a significant effect on its performance. This is unsurprising, as the thresholding procedure of [6] already includes an exponential moving average which smooths over the Gaussian noise for all step sizes. Indeed, since smaller values of T result in more measurements per time interval, the stronger noise is balanced out by the greater sample size of the average. This is the same trade-off described in Eq. (10) in the context of mean value estimation.

Looking next at the linear Wonham filter, Figure 10 provides a very clear demonstration of its dependence on T. At a time step of 1 ns the filter is effectively optimal, which is expected given that the differential Wonham filter of Eq. (30) is known to be exact. However, under more realistic conditions in which hardware latency and other factors generate longer time steps, the first-order nature of the filter starts to negatively impact its performance.

For the single-term and two-term filters, several different trends emerge. The single-term filter (green line) has a much poorer performance at small time steps than the two-term filter (red line), while at larger time steps their performances converge. This occurs because the transition probabilities are suppressed as $T \to 0$, which means that almost all of the contributions to the LogSumExp are small. For the two-term filter this is not a significant problem, since taking both the largest and second-largest terms allows for the log-probability to accumulate over time. The outputs of the single-term filter cannot build up in this manner, so the log-probabilities are reduced far below their true values. As T increases this effect is diminished, so the performance of the two filters will begin to converge. Figure 10 also shows that the two-term filter diverges from the optimal filter at large T, since the single-error Gaussian approximation made in Sec. 5.1 starts to break down when μT is large and $\frac{k}{T}$ is small.

7 Discussion

The core objective of our work here was to devise a quantum error detection filter that is grounded in formal Bayesian analysis, yet also practical to implement on real-world quantum hardware. We have assumed that the continuous measurements are represented as a set of discrete signal averages taken over a time T, which then serve as inputs to the filter. This finite time step T cannot be avoided when working with actual hardware, since even if weak measurements could be made on arbitrarily short time scales there would still be latency associated with transferring and processing the data. As discussed in Sec. 2.2, the fact that T must be non-zero has important effects on the theoretical distribution of the measurement values, which become more significant as both T and the error rate μ increase.

The filter derived from our Bayesian analysis in Sec. 3 can be understood as as a discretized version of the Wonham filter for finite T, though our derivation did not originate from its stochastic differential equation. Much of the difficulty in that analysis stemmed from the fact that errors can occur within the integration period of the measurement, which causes the underlying syndrome means to be distributed across the entire $[-1, 1]$ interval instead of being constrained to only ± 1. Even with this additional degree of freedom, a closed-form solution for the marginal distribution of either syndrome can be derived by simply summing Eq. (32) over $0 \leq e_k < \infty$. The truly challenging part of the analysis comes when the two syndrome distributions are coupled together by errors on the second qubit, which induces complex inter-dependencies between $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$. One avenue for future work could lie in analyzing the form of these joint distributions, with the goal of determining whether a convenient or illuminating analytic form exists. In the absence of

such a form, some type of simplification will always be necessary in order to utilize the filter.

Indeed, the single-error assumption that we made when deriving our log-probability filters in Sec. 5 was designed explicitly to avoid the mathematical challenges associated with the joint syndrome distributions. This assumption was the most restrictive that we made when developing the filters, and it becomes increasingly inaccurate as the average number of errors per step (μT) increases. While it is true that the error tracking problem as a whole becomes quite challenging when μT is large, the question of how a filter could best be designed for such a situation warrants further exploration.

However, in the preferred regime where μT is small, the numerical results in Sec. 6 show that our single-term and two-term log-probability filters are highly effective. In every test that we conducted, these two filters outperformed both the first-order Wonham filter and double threshold algorithm by a significant margin, with the two-term filter being virtually optimal over a wide range of μ and T values. The single-term filter performs slightly worse, but is surprisingly effective given its simplicity. Furthermore, Figure 10 shows that the performance of the single-term filter increases as T grows larger, up to a value of about 10^{-1} μs. This, combined with the fact that the optimal filter experiences only a negligible improvement in performance when moving to smaller T, suggests that effort spent on latency reduction is likely to provide diminishing returns for filter accuracy.

Given the rapidly growing size of modern quantum hardware, it is safe to assume that error correction will remain an integral component of quantum computation for the foreseeable future. The three-qubit toy model analyzed here is insufficient to protect against arbitrary errors, so an obvious extension of our work would focus on developing continuous error correction filters for larger codes e.g., the Shor, Steane, or subsystem codes which provide more comprehensive protection [23][24]. Our results suggest that Bayesian filters similar to those presented here would find great success on these more complex systems.

Acknowledgements

I.C. was supported by the National Aeronautics and Space Administration under Grant/Contract/ Agreement No. 80NSSC19K1123 issued through the Aeronautics Research Mission Directorate. This material is also based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator.

References

[1] Peter W. Shor. "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer". In: *SIAM Journal on Computing* 26.5 (Oct. 1997), pp. 1484–1509. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/S0097539795293172.

[2] *Quantum Error Correction*. Cambridge University Press, 2013. ISBN: 978-0-521-89787-7. DOI: 10.1017/CBO9781139034807.

[3] W. M. Wonham. "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering". In: *Journal of the Society for Industrial and Applied Mathematics Series A Control* 2.3 (Jan. 1964), pp. 347–369. ISSN: 0887-4603. DOI: 10.1137/0302028.

[4] Juan Atalaya et al. “Bacon-Shor code with continuous measurement of non-commuting operators”. In: *Physical Review A* 95.3 (Mar. 2017), p. 032317. DOI: 10.1103/PhysRevA.95.032317.

[5] Razieh Mohseninia et al. “Always-On Quantum Error Tracking with Continuous Parity Measurements”. In: *Quantum* 4 (Nov. 2020), p. 358. DOI: 10.22331/q-2020-11-04-358.

[6] J. Atalaya et al. “Continuous quantum error correction for evolution under time-dependent Hamiltonians”. In: *Physical Review A* 103.4 (Apr. 2021), p. 042406. DOI: 10.1103/PhysRevA.103.042406.

[7] Satoshi Morita and Hidetoshi Nishimori. “Mathematical foundation of quantum annealing”. In: *Journal of Mathematical Physics* 49.12 (Dec. 2008), p. 125210. ISSN: 0022-2488. DOI: 10.1063/1.2995837.

[8] Kurt Jacobs and Daniel A. Steck. “A straightforward introduction to continuous quantum measurement”. In: *Contemporary Physics* 47.5 (Sept. 2006), pp. 279–303. ISSN: 0010-7514, 1366-5812. DOI: 10.1080/00107510601101934.

[9] Alexander N. Korotkov. “Quantum Bayesian approach to circuit QED measurement with moderate bandwidth”. In: *Physical Review A* 94.4 (Oct. 2016), p. 042326. ISSN: 2469-9926, 2469-9934. DOI: 10.1103/PhysRevA.94.042326.

[10] Howard M. Wiseman and Gerard J. Milburn. *Quantum Measurement and Control*. Cambridge University Press, 2009. ISBN: 978-0-521-80442-4. DOI: 10.1017/CBO9780511813948.

[11] Vic Barnett. *Comparative statistical inference*. 3rd ed. Wiley series in probability and mathematical statistics. Wiley, 1999. ISBN: 978-0-471-97643-1. DOI: 10.1002/9780470316955.

[12] Alexander N. Korotkov. “Continuous quantum measurement of a double dot”. In: *Physical Review B* 60.8 (Aug. 1999), pp. 5737–5742. DOI: 10.1103/PhysRevB.60.5737.

[13] Bruno De Finetti et al. *Theory of probability: a critical introductory treatment*. John Wiley & Sons, 2017. ISBN: 978-1-119-28634-9. DOI: 10.1002/9781119286387.

[14] D. S. Sivia and J. Skilling. *Data analysis: a Bayesian tutorial*. 2nd ed. Oxford science publications. Oxford University Press, 2006. ISBN: 978-0-19-856831-5.

[15] J. R. Norris. *Markov Chains*. 1st ed. Cambridge University Press, Feb. 1997. ISBN: 978-0-521-48181-6. DOI: 10.1017/CBO9780511810633.

[16] Aiken Pang and Peter Membrey. *Beginning FPGA: Programming Metal*. Apress, 2017. ISBN: 978-1-4302-6247-3. DOI: 10.1007/978-1-4302-6248-0.

[17] Hideo Mabuchi. “Continuous quantum error correction as classical hybrid control”. In: *New Journal of Physics* 11.10 (Oct. 2009), p. 105044. ISSN: 1367-2630. DOI: 10.1088/1367-2630/11/10/105044.

[18] A. V. Borisov. “L1-Optimal Filtering of Markov Jump Processes. I. Exact Solution and Numerical Implementation Schemes”. In: *Automation and Remote Control* 81.11 (Nov. 2020), pp. 1945–1962. ISSN: 0005-1179, 1608-3032. DOI: 10.1134/S0005117920110016.

[19] Yin Gang George, Qing Zhang, and Yuanjin Liu. “Discrete-time approximation of Wonham filters”. In: *Journal of Control Theory and Applications* 2.1 (Feb. 2004), pp. 1–10. ISSN: 1993-0623. DOI: 10.1007/s11768-004-0017-7.

[20] G. M. Hieftje. "Signal-to-noise enhancement through instrumental techniques. II. Signal averaging, boxcar integration, and correlation techniques". In: *Analytical Chemistry* 44.7 (June 1972), 69A–78a. ISSN: 0003-2700, 1520-6882. DOI: 10.1021/ac60315a036.

[21] Pierre Blanchard, Desmond J Higham, and Nicholas J Higham. "Accurately computing the log-sum-exp and softmax functions". In: *IMA Journal of Numerical Analysis* (Aug. 2020), draa038. ISSN: 0272-4979, 1464-3642. DOI: 10.1093/imanum/draa038.

[22] M. Haselman et al. "A Comparison of Floating Point and Logarithmic Number Systems for FPGAs". In: *13th Annual IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM'05)*. IEEE, 2005, pp. 181–190. ISBN: 978-0-7695-2445-0. DOI: 10.1109/FCCM.2005.6.

[23] Peter W. Shor. "Scheme for reducing decoherence in quantum computer memory". In: *Physical Review A* 52.4 (Oct. 1995), R2493–R2496. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.52.R2493.

[24] A. M. Steane. "Error Correcting Codes in Quantum Theory". In: *Physical Review Letters* 77.5 (July 1996), pp. 793–797. DOI: 10.1103/PhysRevLett.77.793.

A Appendix

A.1 Syndrome distributions

A.1.1 General form

In Eq. (22), we expand the syndrome distribution $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$ in terms of the error probabilities $p(e_k|\ell_i\ell_{i-1})$ and the conditional distribution $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1};e_1e_2e_3)$ of observing a particular pair of syndrome means given a number of errors e_k on the kth qubit. Figure 1C shows that the values of $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ depend explicitly on the number and *location* of the bit-flip errors within the integration interval, so we introduce a new set of variables $\{x_j^{k_j}\}$ which denote the positions of the $N \equiv e_1 + e_2 + e_3$ errors. The index j indicates the order that the errors occur, such that $x_{j-1}^{k_{j-1}} \leq x_j^{k_j}$, while $k_j \in \{1,2,3\}$ denotes on which qubit the error occurs. Note that these "positions" are in fact times t during the run, but for convenience we define the beginning of the ith integration period to be at $t = 0$ and then divide the times by T so that they all lie on the unit-less interval $[0,1]$. The value $0 \leq x_j^{k_j} \leq 1$ then denotes a location within this interval. Using this new set of variables, the syndrome density is given by the marginalization over all error locations and all possible qubit assignments:

$$P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1};e_1e_2e_3) = \sum_{k_1=1}^{3}\cdots\sum_{k_N=1}^{3}\int_0^1 dx_1^{k_1}\int_{x_1^{k_1}}^1 dx_2^{k_2}\cdots\int_{x_{N-1}^{k_{N-1}}}^1 dx_N^{k_N}\, P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}\{x_j^{k_j}\})P(\{x_j^{k_j}\}|e_1e_2e_3), \tag{54}$$

where we have now further factorized the syndrome density into the product of two new distributions.

Starting first with $P(\{x_j^{k_j}\}|e_1e_2e_3)$, we know that the location of an error and the qubit that it is assigned to are independent, so we have

$$P(\{x_j^{k_j}\}|e_1e_2e_3) = P(\{x_j\}|e_1e_2e_3)P(\{k_j\}|e_1e_2e_3). \tag{55}$$

Given that each qubit is equally likely to experience an error, $P(\{k_j\}|e_1e_2e_3)$ must be uniform over all *valid* assignments of $\{k_j\}$, where an assigment is valid if exactly e_k errors are assigned to the kth qubit. Since the distribution must be normalized, each configuration will have a probability equal to one over the total number of valid configurations:

$$P(\{k_j\}|e_1e_2e_3) = \begin{cases} \dfrac{e_1!e_2!e_3!}{N!} & \text{when valid} \\ 0 & \text{otherwise.} \end{cases} \tag{56}$$

Similarly, the errors themselves will be uniformly distributed within the interval, so the joint distribution of all error locations $\{x_j\}$ is one over the integral across all positions

$$P(\{x_j\}|e_1e_2e_3) = \left[\int_0^1 dx_1^{k_1} \int_{x_1^{k_1}}^1 dx_2^{k_2} ... \int_{x_{N-1}^{k_{N-1}}}^1 dx_N^{k_N}\right]^{-1} = N!, \tag{57}$$

which depends only on the total number of errors.

Turning now to $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}\{x_j^{k_j}\})$, the location of the errors and their qubit assignments will completely determine the value of the syndromes, so we will have a product of Dirac delta distributions. Specifically, the average value over an interval depends on the gaps in time between successive errors, as well as the gap between the start of the interval and the first error and the gap between the end of the interval and the last error. As shown in Figure 1C, the syndrome value is given by the sum of these gaps, with the sign of each contribution alternating due to the parity flips caused by the errors. Therefore, the syndrome distribution has the form

$$\begin{aligned} P(\bar{S}_i^{(m)}|\ell_{i-1}\{x_j^{(m)}\}) &= \delta\left(\pm_m \bar{S}_i^{(m)} - x_1^{(m)} - \left[\sum_{j=2}^{N^{(m)}} (-1)^{j-1}(x_j^{(m)} - x_{j-1}^{(m)})\right] - (-1)^{N^{(m)}}(1 - x_{N^{(m)}}^{(k)})\right) \\ &= \delta\left(\pm_m \bar{S}_i^{(m)} - (-1)^{N^{(m)}} - 2\sum_{j=1}^{N^{(m)}} (-1)^{j-1} x_j^{(m)}\right), \end{aligned} \tag{58}$$

where $\{x_j^{(m)}\}$ is the set of errors which affect the mth parity operator and $N^{(m)}$ is the number of errors in this set. For the second interval in Figure 1C, the syndrome distributions would then be given by

$$P(\bar{S}_i^{(m)}|\ell_{i-1}\{x_j^{(m)}\}) = \delta\left(\pm_m \bar{S}_i^{(m)} - x_1 + (1 - x_1)\right) = \delta\left(\pm_m \bar{S}_i^{(m)} - 0.4\right), \tag{59}$$

where $x_1 = 0.7$ is the normalized position of the error on the second qubit.

Using Eqs. (55–58), we can rewrite Eq. (54) as

$$\begin{aligned} &P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}; e_1e_2e_3) = \\ &\quad e_1!e_2!e_3! \sum_{\{k_j\}\text{ valid}} \int_0^1 dx_1^{k_1} \int_{x_1^{k_1}}^1 dx_2^{k_2} ... \int_{x_{N-1}^{k_{N-1}}}^1 dx_N^{k_N} P(\bar{S}_i^{(1)}|\ell_{i-1}\{x_j^{(1)}\}) P(\bar{S}_i^{(2)}|\ell_{i-1}\{x_j^{(2)}\}), \end{aligned} \tag{60}$$

which can be understood as the sum of volume integrals over the values of $\{x_j^{(k_j)}\}$ which satisfy the Dirac distributions from Eq. (58). In principle, all that remains is to solve for the appropriate integral boundaries and then evaluate the integrals, though this is difficult to do for an arbitrary distribution of errors. In Sec. A.1.2 we describe how Eq. (60) can be evaluated when errors occur on only a single qubit, while in Sec. A.1.3 we discuss the challenges of generalizing this to errors on multiple qubits, specifically when the second qubit is involved.

A.1.2 Distributions for single-qubit errors

If errors are constrained to occur on only a single qubit, then one of the syndrome means will behave in a trivial manner. For errors on the first qubit we will have $\bar{S}_i^{(2)} = \pm_2 1$, and for errors on the the third qubit we will have $\bar{S}_i^{(1)} = \pm_1 1$. With errors on the second qubit both syndromes will change in an identical manner, so we can choose to model explicitly the behavior of $\bar{S}_i^{(1)}$ and then just set $\bar{S}_i^{(2)} = \bar{S}_i^{(1)}$.

We begin by considering a scenario where the first qubit experiences N errors, while $e_2 = e_3 = 0$. Under these conditions, $P(\bar{S}_i^{(2)}|\ell_{i-1}\{x_j^{(2)}\}) = \delta(\bar{S}_i^{(2)} \mp_2 1)$, as $\{x_j^{(2)}\}$ is empty and thus there are no errors on the second or third qubits to flip the initial syndrome of the second parity operator. Indeed, the syndrome distribution of Eq. (60) factorizes completely, with the marginal distribution of $\bar{S}_i^{(1)}$ given by

$$P(\bar{S}_i^{(1)}|\ell_{i-1}; N00) = N! \int_0^1 dx_1 \int_{x_1}^1 dx_2 ... \int_{x_{N-1}}^1 dx_N \delta\left(\pm_1 \bar{S}_i^{(1)} - (-1)^N - 2\sum_{j=1}^{N}(-1)^{j-1}x_j\right). \tag{61}$$

Note that the sum of Eq. (60) has collapsed into a single term, since there is only one way to assign N errors to a single qubit. Given that $k_j = 1$ for all j, we have removed the superscripts from the error locations.

As mentioned previously, we can view Eq. (61) as a volume integral over the regions of the variable space which give the desired value of $\bar{S}_i^{(1)}$. To derive the integration bounds, we consider first the constraints placed on x_1. The first error will contribute $\pm_1(x_1 - 0) = \pm_1 x_1$ to the value of the syndrome, which cannot be altered by any subsequent errors. Therefore, the range of $\bar{S}_i^{(1)}$ after the first error is

$$\pm_1 x_1 - (1 - x_1) \leq \bar{S}_i^{(1)} \leq \pm_1 x_1 + (1 - x_1), \tag{62}$$

since the rest of interval can be entirely negative or entirely positive at the two extremes. Writing the limits with respect to x_1 gives

$$x_1 \leq \frac{1 \pm_1 \bar{S}_i^{(1)}}{2}, \tag{63}$$

which becomes the upper integration bound of the first integral. We can repeat this same process for each x_j, taking the sum of the contributions from errors up to x_j and either adding or subtracting the remaining interval length to get the upper and lower bounds respectively on $\bar{S}_i^{(1)}$. These bounds are then rearranged to derive an upper bound on x_j, which has the generic form

$$B_j \equiv \frac{1}{2} + (-1)^{j-1}\left(\frac{\pm_1 \bar{S}_i^{(1)}}{2} + \sum_{n=1}^{j-1}(-1)^n x_n\right), \tag{64}$$

such that the integration limits are $x_{j-1} \leq x_j \leq B_j$. These bounds have the recursion relation

$$B_j = (1 + x_{j-1}) - B_{j-1} \tag{65}$$

which is key to evaluating the integrals.

Using Eq. (64) and Eq. (65), the volume integral in Eq. (61) can be easily solved. The integral over x_N will collapse to $\frac{1}{2}$ due to the Dirac delta function, and contribute nothing

further to the volume. Considering the next two integrals we have

$$\begin{aligned}\int_{x_{N-3}}^{B_{N-2}} dx_{N-2} \int_{x_{N-2}}^{B_{N-1}} dx_{N-1} &= \int_{x_{N-3}}^{B_{N-2}} dx_{N-2}(B_{N_1} - x_{N-2}) \\ &= (1 - B_{N-2}) \int_{x_{N-3}}^{B_{N-2}} dx_{N-2} \\ &= (1 - B_{N-2})(1 - B_{N-3}),\end{aligned} \tag{66}$$

where we used the recursive expression $B_j - x_{j-1} = 1 - B_{j-1}$. The result of the first integral does not have any dependence on x_{N-2} and therefore is unaffected by the second integral. Indeed, the generic action of the integral for x_{j-1} on $(1 - B_j)^a$ is

$$\int_{x_{j-2}}^{B_{j-1}} dx_{j-1}(1 - B_j)^a = \frac{1}{a+1}(1 - B_{j-2})^{a+1}, \tag{67}$$

which maps $B_j \to B_{j-2}$. By repeated application of Eq. (67), the probability density in Eq. (61) reduces to

$$P(\bar{S}_i^{(1)}|\ell_{i-1}; N00) \propto (1 - B_0)^{N//2}(1 - B_1)^{(N-1)//2}, \tag{68}$$

where $x//y$ is x floor-divided by y, and where B_0 is defined from the recursion relation in Eq. (65) with $j = 1$ and $x_0 \equiv 0$. Substituting the expressions for B_1 and B_0 into Eq. (68) gives

$$P(\bar{S}_i^{(1)}|\ell_{i-1}; N00) = \frac{N!}{2(N//2)!((N-1)//2)!}\left(\frac{1 \pm_1 \bar{S}_i^{(1)}}{2}\right)^{N//2}\left(\frac{1 \mp_1 \bar{S}_i^{(1)}}{2}\right)^{(N-1)//2} \tag{69}$$

after normalization, which is simply a beta distribution with respect to $\frac{1 \pm \bar{S}_i^{(1)}}{2}$ on the interval $[-1, 1]$. This result can be easily generalized to N errors on the third qubit by replacing $\bar{S}_i^{(1)}$ with $\bar{S}_i^{(2)}$ and $\pm_1$ with $\pm_2$.

A.1.3 Challenges of multi-qubit errors

When errors occur on the second qubit and either the first or third qubits (or both), the evaluation of Eq. (54) is far more challenging. First, the sum over $\{k_j\}$ will not collapse to a single term, so we must evaluate multiple volume integrals. More importantly, the error locations assigned to the second qubit in each of those volume integrals will have to simultaneously satisfy constraints imposed by both $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$, which means that the bounds B_j on those errors will be given by

$$B_j = \min\left[\frac{1}{2} + (-1)^{j-1}\left(\frac{\pm_1 \bar{S}_i^{(1)}}{2} + \sum_{n=1}^{N_j^{(1)}}(-1)^n x_n^{(1)}\right), \frac{1}{2} + (-1)^{j-1}\left(\frac{\pm_2 \bar{S}_i^{(2)}}{2} + \sum_{n=1}^{N_j^{(2)}}(-1)^n x_n^{(2)}\right)\right], \tag{70}$$

where $N_j^{(m)}$ is the number of errors occurring earlier than $x_j^{k_j}$ that affect the mth parity operator. These bounds create dependencies between the variables in $\{x_j^{(1)}\}$ and $\{x_j^{(2)}\}$ which prevent us from using the simple recursion strategy employed in Eq. (67). That said, an expression for $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_{i-1}; e_1e_2e_3)$ can still be calculated from Eq. (54) in a brute-force manner using symbolic computation tools, though this becomes increasingly impractical as the values of e_1, e_2, and e_3 grow.

A.2 Implementing the optimal filter

From Eq. (29), the optimal Bayesian filter is a function of the measurement density $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ and transition elements $J_{\ell_{i-1}\ell_i}$. The matrix J is easy to compute, but in order to calculate $P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1})$ we must evaluate the integral in Eq. (21). Although the analytic solution is not available to us, we can approximate the value of the integral using a Riemann sum

$$P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1}) \approx \frac{1}{n^2}\sum_{m^{(1)}=-n}^{n-1}\sum_{m^{(2)}=-n}^{n-1}\prod_{j=1}^{2}\frac{\exp[\frac{-T}{2k}(M_i^{(j)}-(\frac{m^{(j)}+\frac{1}{2}}{n}))^2]}{\sqrt{2\pi\frac{k}{T}}}P(\bar{S}_i^{(1)}=\frac{m^{(1)}+\frac{1}{2}}{n},\bar{S}_i^{(2)}=\frac{m^{(2)}+\frac{1}{2}}{n}|\ell_i\ell_{i-1}), \tag{71}$$

where the interval $[-1,1]$ has been discretized into $4n^2$ evenly-spaced segments.

The syndrome density is estimated by using Eq. (22) to factorize it into $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|e_1e_2e_3)$ and $P(e_1e_2e_3|\ell_i\ell_{i-1})$, the latter of which can be evaluated analytically. To approximate the value of $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|e_1e_2e_3)$, we construct histograms of $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ with $4n^2$ bins for values of e_1, e_2, and e_3 such that $e_1+e_2+e_3 \leq N$. The value of N is chosen so that the probability of experiencing more than N errors in a given interval is below some threshold value. Since μ and T only impact $P(e_1e_2e_e|\ell_i\ell_{i-1})$, these histograms can be reused for multiple error rates and integration times. To reconstruct $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$, we perform the sum in Eq. (22) using the histograms and the analytic error probabilities to generate a single combined histogram.

The sum in Eq. (71) must be evaluated at every time step, which becomes a significant computational bottleneck for large n. To speed up the filter, we constructed a lookup table for the Gaussian distribution at discrete values across $M_i^{(1)}$ and $M_i^{(2)}$ and then used bi-linear interpolation to compute values for continuous inputs. Given the smoothness of the Gaussian function, these interpolations were highly accurate and very fast to compute.

A.3 Convergence to the Wonham filter

The (unnormalized) Bayesian filter given in Eq. (29) can be shown to converge to the linear Wonham filter of Mohseninia et al. [5] as $T \to 0$, and thus to converge to the Wonham filter after normalization. Up to the smallest order in T, the transition matrix elements are given by

$$J_{\ell_{i-1}\ell_i} \approx (\mu T)^{d(\ell_{i-1},\ell_i)}(1-3\mu T), \tag{72}$$

where $d(\ell_{i-1},\ell_i)$ is the Hamming distance between the three-bit representations of ℓ_i and ℓ_{i-1}. Ignoring the normalization factor, the measurement density can be expanded up to its smallest order in T as

$$P(M_i^{(1)}M_i^{(2)}|\ell_i\ell_{i-1}) \propto \int_{-1}^{1}\int_{-1}^{1}\left(1-\frac{T}{2k}\left[(M_i^{(1)}-\bar{S}_i^{(1)})^2+(M_i^{(2)}-\bar{S}_i^{(2)})^2\right]\right)\tilde{P}(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})d\bar{S}_i^{(1)}d\bar{S}_i^{(2)}, \tag{73}$$

where $\tilde{P}(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1})$ is equal to $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)})|e_1e_2e_3)$ such that $e_k=1$ if there is a net flip on the kth qubit when moving from ℓ_{i-1} to ℓ_i and $e_k=0$ otherwise. This is the syndrome density with the fewest errors that is still consistent with the state transition.

Substituting Eq. (72) and Eq. (73) into Eq. (29) gives

$$\hat{P}(\ell_i) \propto \sum_{\ell_{i-1}} (\mu T)^{d(\ell_{i-1},\ell_i)} (1-3\mu T)(1-\frac{T}{2k} I_{\ell_{i-1}\ell_i}) P(\ell_{i-1}), \tag{74}$$

where I is a matrix with elements given by

$$I_{\ell_{i-1}\ell_i} = \int_{-1}^{1}\int_{-1}^{1} \left[(M_i^{(1)} - \bar{S}_i^{(1)})^2 + (M_i^{(2)} - \bar{S}_i^{(2)})^2\right] \tilde{P}(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|\ell_i\ell_{i-1}) d\bar{S}_i^{(1)} d\bar{S}_i^{(2)}. \tag{75}$$

The sum in Eq. (74) can be simplified by keeping only those terms which are at most first-order in T. With $\mathcal{I} \equiv \{1,2,4\}$ containing the set of indices corresponding to bit-flips on the first, second, and third qubits respectively, we have

$$\hat{P}(\ell_i) \propto P(\ell_{i-1} = \ell_i) + T\left[\mu \sum_{i\in\mathcal{I}} P(\ell_{i-1} = \ell_i \oplus i) - (3\mu + \frac{1}{2k} I_{\ell_i\ell_i}) P(\ell_{i-1} = \ell_i)\right], \tag{76}$$

which is equivalent to Eq. (31) after introducing Q and evaluating $I_{\ell_i\ell_i}$.

A.4 Gaussian fit of single-error distribution

In Eq. (35) we considered the distribution of $M_i^{(1)}$ when a single error occurs on the first qubit, deriving the expression

$$P(M_i^{(1)}|\ell_{i-1}; 100) = \frac{1}{4}\left[\text{erf}(\frac{M_i^{(1)}+1}{\frac{2k}{T}}) - \text{erf}(\frac{M_i^{(1)}-1}{\frac{2k}{T}})\right]. \tag{77}$$

From a practical perspective it would be convenient if the distribution were Gaussian, so we compute the mean value s and variance σ^2 of Eq. (77) and then construct a Gaussian distribution off of these parameters. This can be done by going back to the integral which generated Eq. (77) and simply switching the order of integration

$$s = \frac{1}{2}\int_{-1}^{1} d\bar{S}_i^{(1)} \int_{-\infty}^{\infty} dM_i^{(1)} M_i^{(1)} \frac{\exp[\frac{-T}{2k}(M_i^{(1)} - \bar{S}_i^{(1)})^2]}{\sqrt{2\pi\frac{k}{T}}} = \frac{1}{2}\int_{-1}^{1} \bar{S}_i^{(1)} d\bar{S}_i^{(1)} = 0 \tag{78}$$

$$\begin{aligned}\sigma^2 &= \frac{1}{2}\int_{-1}^{1} d\bar{S}_i^{(1)} \int_{-\infty}^{\infty} dM_i^{(1)} (M_i^{(1)})^2 \frac{\exp[\frac{-T}{2k}(M_i^{(1)} - \bar{S}_i^{(1)})^2]}{\sqrt{2\pi\frac{k}{T}}} = \frac{1}{2}\int_{-1}^{1} [(\bar{S}_i^{(1)})^2 + \frac{k}{T}] d\bar{S}_i^{(1)} \\ &= \frac{1}{3} + \frac{k}{T}.\end{aligned} \tag{79}$$

A.5 Simulating trajectories

To generate synthetic data of duration nT, we must correctly evolve the state of the system across each of the n integration intervals (of length T), and then generate an appropriate measurement record. To perform this evolution we utilize the so-called "jump, no-jump" approach, where individual bit-flips are sampled at each time step and then applied to the system such that the state remains pure across the run. For particular values of μ and T the probability of experiencing e_k bit-flip errors on the kth qubit is given by the Poisson distribution in Eq. (16), which can be easily sampled from using standard mathematical libraries. After e_1, e_2, and e_3 have been determined for each of the n time steps, we apply the errors to our initial state $|000\rangle$ in the proper order to evolve the system.

To generate the corresponding measurements, we first sample the syndrome means $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$ from the measurement distribution $P(\bar{S}_i^{(1)}\bar{S}_i^{(2)}|e_1e_2e_3)$ conditioned on the number of errors that occurred in the step. This can be done in straightforward manner by sampling e_k uniform values on the interval $[0, 1]$ for each qubit and then sorting them by value in ascending order. To calculate $\bar{S}_i^{(1)}$, the time gaps between errors on the first and second qubit are added and subtracted in an alternating order, with the first gap given by the value of the first error and the last gap given by one minus the value of the last error (see Eq. (58) for more detail). After the gaps have been summed the syndrome value is multiplied by $\pm_1$ to take into account the parity of the state at the start of interval. An analogous procedure can be be carried out for $\bar{S}_i^{(2)}$ using $\pm_2$ and the gaps between errors on the second and third qubits. Once the syndrome values have been sampled, the values of $M_i^{(1)}$ and $M_i^{(2)}$ are generated by simply adding Gaussian noise with mean zero and variance $\frac{k}{T}$ to $\bar{S}_i^{(1)}$ and $\bar{S}_i^{(2)}$.

The approach to trajectory simulation described here is more involved than most methods for simulating continuous measurements, in that it treats the syndrome mean as a random variable on the interval [-1, 1] that must be conditionally sampled based on the number of errors that occurred in the interval. This is necessary in order to accurately test the performance of a filter on different values of T. Other simulation techniques are usually only concerned with the limiting behavior as $T \to 0$, and therefore just sample syndrome values of ± 1 based on the state of the system.