Universiteit
Leiden
The Netherlands

Adaptive weight estimator for quantum error correction in a time-dependent environment

Spitz, S.T.; Tarasinski, B.; Beenakker, C.W.J.; O'Brien, T.E.

Citation
Spitz, S. T., Tarasinski, B., Beenakker, C. W. J., & O'Brien, T. E. (2018). Adaptive weight estimator for quantum error correction in a time-dependent environment. *Advanced Quantum Technologies, 1*(1). doi:10.1002/qute.201800012

Version:	Publisher's Version
License:	Creative Commons CC BY 4.0 license
Downloaded from:	https://hdl.handle.net/1887/3281647

Note: To cite this publication please use the final published version (if applicable).

Adaptive Weight Estimator for Quantum Error Correction in a Time-Dependent Environment

*Stephen T. Spitz, Brian Tarasinski, Carlo W. J. Beenakker, and Thomas E. O'Brien**

Quantum error correction of a surface code or repetition code requires the pairwise matching of error events in a space-time graph of qubit measurements, such that the total weight of the matching is minimized. The input weights follow from a physical model of the error processes that affect the qubits. This approach becomes problematic if the system has sources of error that change over time. Here, it is shown that the weights can be determined from the measured data in the absence of an error model. The resulting adaptive decoder performs well in a time-dependent environment, provided that the characteristic timescale $\tau_{\rm env}$ of the variations is greater than $\delta t/\bar{p}$, with δt the duration of one error-correction cycle and $\bar{p}$ the typical error probability per qubit in one cycle.

1. Introduction

To execute algorithms on a quantum computer, one must prevent the accumulation of errors by monitoring and correcting them in control hardware. The monitoring is made possible by a nonlocal encoding of the quantum information in a redundant set of qubits, allowing for repeated measurements via auxiliary (ancilla) qubits without collapsing and destroying the quantum superposition of the logical degrees of freedom.[1,2] Parity-check measurements produce strings of bits, the so-called error syndrome, that must be decoded to infer the correction which should be applied to the logical qubits.

For an important class of error correcting codes, the syndrome identifies the end points of an error chain in a space-time graph of ancilla measurements (See **Figure 1**). The dimensionality of space can differ; it equals 1 in the repetition code,[3,4] 2 in the surface code,[5–7] and 3 for topological cluster states.[8] The identification is not unique: there is in general no unique way to construct a chain of error events consistent with a given syndrome (the decoding problem). One approach to decoding refers to the optimization problem of minimum-weight perfect matching on a graph, which may be solved by the "blossom" algorithm[9,10] in polynomial time. The blossom decoder is sub-optimal,[11–14] but it performs sufficiently well for current quantum hardware to achieve the fault-tolerance threshold.[15]

The weights that govern the optimization problem can be readily obtained if one has a calibrated model of the sources of error in the system.[16] Such an error model may not be available, and moreover the error rates may vary in time during the quantum computation. This complication has motivated the search for an adaptive decoder, that would infer the weights from the syndrome without requiring updates of the error model.[17–21] Since the syndrome depends nonlinearly on the weights, this inversion problem is nontrivial—a recent approach[21] employs a machine learning algorithm to learn the weights from the measured data.

Here we show that the inversion can be actually carried out by purely algebraic means: The covariance of measurements on pairs of ancillas exactly determines the weight of their matching. We demonstrate the method on the repetition code with time-dependent error rates.

2. Quantum Error Correction and the Repetition Code

To set the stage, we summarize the elements of quantum error correction[2,22] that we need in what follows. The expert reader may skip this section.

A quantum error correcting code stores quantum information nonlocally in an array of physical qubits, such that it is protected from local errors (bit flips or sign flips). The encoded state $|\psi\rangle$ evolves for a cycle time δt, after which a set of "stabilizer" measurements is carried out. The stabilizers project $|\psi\rangle$ onto a state $|\psi'\rangle$ that may differ from $|\psi\rangle$ if an error occurrs during the cycle. The outcome of the stabilizer measurements, called the syndrome, identifies the error and allows for a correction. It is crucial that the stabilizer measurements do not measure the degrees of freedom of $|\psi\rangle$ in which the relevant quantum information is stored; otherwise this information will be lost upon projection.

The simplest example of error correction via stabilizer measurement is a 1D array, in which a logical qubit is encoded into

Dr. S. T. Spitz, Prof. C. W. J. Beenakker, T. E. O'Brien
Instituut-Lorentz
Universiteit Leiden
P.O. Box 9506, 2300 RA Leiden, The Netherlands
E-mail: obrien@lorentz.leidenuniv.nl
Dr. B. Tarasinski
QuTech
Delft University of Technology
P.O. Box 5046, 2600 GA Delft, The Netherlands

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/qute.201800012

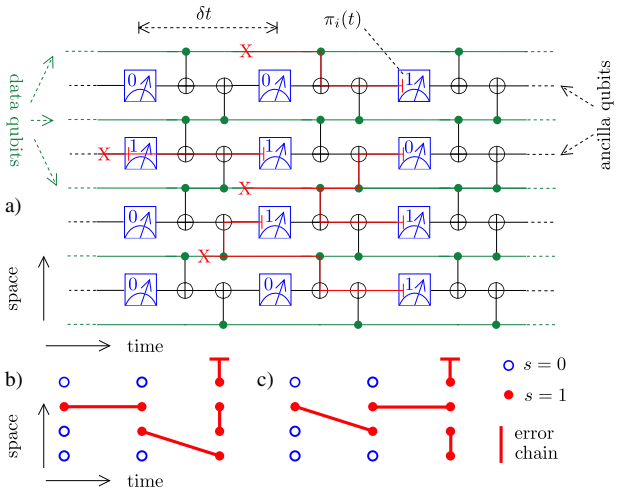

Figure 1. a) Space-time circuit of a $d = 5$ quantum repetition code. Five data qubits (green) are entangled with four ancilla qubits (black) by means of a CNOT gate ($\oplus$). The ancillas are measured at the end of each cycle (blue boxes, spaced by δt, with measurement outcomes $\pi_i(t)$, $i = 1, 2, 3, 4$). A bit flip (red X) produces an error chain (red line) with end points on an ancilla measurement or on the boundary of the array. b) Syndrome $s_i(t) = \pi_i(t) \oplus \pi_i(t - 2\delta t)$ corresponding to the error events in panel (a). The measurements that are connected by an error chain can be separated in space, in time, or in both space and time. c) Alternative matching consistent with the same syndrome. The minimum weight decoder associates a weight to each error chain and finds the matching with the smallest total weight.

d data qubits via $|0\rangle_L = |00\cdots0\rangle$, $|1\rangle_L = |11\cdots1\rangle$. In a classical setting, this would correspond to a distance-d repetition code, for which one would compare the value of adjacent bits to identify up to $(d-1)/2$ bit flips. A quantum parity check achieves this goal without collapsing the superposition $|\psi\rangle = a_0|0\rangle_L + a_1|1\rangle_L$ onto the state $|0\rangle_L$ or $|1\rangle_L$. The parity-check measurements are performed on $d-1$ ancilla qubits, which are entangled with pairs of data qubits (see Figure 1). Each ancilla measures the stabilizer operator $Z_i Z_{i+1}$ ($i = 1, 2, \ldots, d-1$, with $Z \equiv \sigma_z$ a Pauli matrix). The stabilizer does not distinguish between the states $|0\rangle_L$ and $|1\rangle_L$ and thus preserves their quantum superposition. A bit flip error of qubit j (X_j error) is detected by the stabilizer measurements $Z_j Z_{j+1}$ and $Z_{j-1} Z_j$, which change their value from $+1$ to -1. A decoder may infer the underlying error from this signature and correct it without needing to measure qubit j itself (which would collapse the state).

As shown in the circuit of Figure 1a, the measurement outcome $\pi_i(t)$ of the ancilla in cycle t is updated in the next cycle by the addition modulo 2 (denoted by $\oplus$) of the adjacent data qubits. For a given stabilizer $Z_i Z_{i+1} = \pm 1$, one has

$$\pi_i(t+1) = \begin{cases} \pi_i(t) & \text{if } Z_i Z_{i+1} = +1, \\ \pi_i(t) \oplus 1 & \text{if } Z_i Z_{i+1} = -1 \end{cases} \quad (1)$$

Hence, if no errors occur, $\pi_i(t)$ is either constant in time, if the data qubits have even parity, or $\pi_i(t)$ toggles between 0 and 1 from one time step to the next, if the data qubits have odd parity. In each case $\pi_i(t) = \pi_i(t - 2\delta t)$. An error event is then signaled by a nonzero element $s_i(t) = 1$ of the error syndrome

$$\vec{s}(t) = \vec{\pi}(t) \oplus \vec{\pi}(t - 2\delta t) \quad (2)$$

A single error event is not sufficient to diagnose an error, as a change of $Z_i Z_{i+1}$ could signal either an X_i error or an X_{i+1} error. To identify which qubit flipped we match pairs of error events. As indicated in Figure 1, the match can be between error events at the end points of an error chain from ancilla i_0 at time t to ancilla j_0 at time $t + n\delta t$. Note that the error events can be separated in both space and time—errors correlated in time are caused by imperfect stabilizer extraction. The error chain may terminate at the boundary of the lattice (corresponding to errors on the boundary data qubits), so some error events may remain unmatched.

This simple description to detect bit flips in a repetition code can be extended to the detection of both bit flips and phase flips (X_i and Z_i errors) and by encoding in 2D and 3D (surface codes and topological cluster states). The generic feature of this class of stabilizer codes is that the decoding entails the pairwise matching of error events in a space-time graph. The method of adaptive quantum error correction presented in the next section applies to this general setting, while for a demonstration we will return to the repetition code.

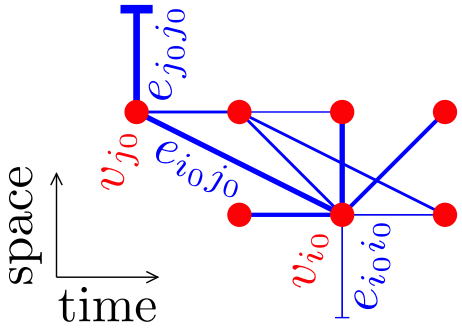

Figure 2. Generic space-time graph showing a pair of vertices v_{i_0}, v_{j_0} connected by an edge $e_{i_0 j_0}$. A few other vertices and connecting edges are also shown, as well as edges that connect a vertex to a boundary ($e_{i_0 i_0}$ and $e_{j_0 j_0}$). The line thickness of an edge e_{ij} is proportional to the probability p_{ij} that a single error affects the ancilla qubit measurements on vertices i and j. We seek to determine these probabilities from measurements of the error syndrome.

3. Weight Inference from Error Syndromes

3.1. Formulation of the Inversion Problem

We collect the binary output of the stabilizer measurements in the error syndrome $\vec{s}(t)$. The discrete time variable t counts the error correction cycle and the elements of the vector $\vec{s}$ identify the ancilla qubits. For N_{ancilla} ancillas and T cycles there are a total of $N_{\text{ancilla}} T$ variables $v_i \in \{0, 1\}$, arranged as vertices in a space-time graph (See **Figure 2**). An error event corresponds to $v_i = 1$, while $v_i = 0$ if the ancilla has not detected an error.

The vertices are pairwise connected by undirected edges $e_{ij} \equiv e_{ji} \in \{0, 1\}$ such that $e_{ij} = 1$ with probability p_{ij}. We allow for $i = j$, the edge e_{ii} connects a vertex to the boundary of the graph.

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

We say that the edge is *on* or *off* depending on whether $e_{ij} = 1$ or 0. The state of a vertex depends on the edges according to

$$v_i = \tfrac{1}{2}\left[1 - (-1)^{\sum_j e_{ij}}\right] \tag{3}$$

Each edge that is *on* toggles its vertex between the states 0 and 1, so that $v_i = 1$ if an odd number of connecting edges is *on*.

The p_{ij}'s are probabilities of a single-qubit error that correlates ancilla measurements i and j. Correlations of ancilla measurements due to uncorrelated multiple-qubit errors are described by weights w_{ij}. The weight w_{ij} for $i \neq j$ is determined from the p's by following all paths $i \mapsto k_1 \mapsto k_2 \mapsto \cdots \mapsto k_n \mapsto j$ through the graph from vertex i to vertex j via n intermediate vertices $k_1, k_2, \ldots, k_n$:

$$w_{ij} = -\ln\left(p_{ij} + \sum_n {\sum_{k_1,k_2,\ldots,k_n}}' p_{ik_1} p_{k_1k_2} \cdots p_{k_nj}\right) \tag{4}$$

The prime in the sum indicates that the path should not pass through the boundary ($i \neq k_1 \neq k_2 \cdots \neq k_n \neq j$). For a boundary weight w_{ii} the path terminates on the boundary

$$w_{ii} = -\ln\left(p_{ii} + \sum_n {\sum_{k_1,k_2,\ldots,k_n}}' p_{ik_1} p_{k_1k_2} \cdots p_{k_nk_n}\right) \tag{5}$$

These sums over error chains can be carried out in terms of the matrix $A_{ij} = (1 - \delta_{ij})\, p_{ij}$ by matrix inversion[15]

$$e^{-w_{ij}} = \begin{cases} [(1-A)^{-1}]_{ij} & \text{if } i \neq j, \\ \sum_k [(1-A)^{-1}]_{ik}\, p_{kk} & \text{if } i = j \end{cases} \tag{6}$$

see Appendix A.

Given a set of error events $\mathcal{V} = \{v_i | v_i = 1\}$ the minimum-weight perfect matching decoder searches for a subset $\mathcal{M} = \{e_{ij}, w_{ij}\}$ of weighted edges such that each vertex in $\mathcal{V}$ is connected either to one other vertex in $\mathcal{V}$ or to the boundary, at minimal total weight $\sum_{\mathcal{M}} w_{ij}$.

Modern implementations[10] of the blossom algorithm[9] solve this optimization problem efficiently given an error model: A physical model for qubit errors from which the error probabilities p and hence the weights w can be calculated. Here we consider the opposite approach: can we infer the weights from the measured error syndromes, without an underlying error model? This is an inversion problem for Equation (3), where we seek to reconstruct the statistics of the edges e_{ij} from the measured statistics of the vertices v_i. The inversion is possible, in spite of the nonlinearity of Equation (3), as we show in the following two subsections (with a more detailed derivation in Appendix B).

3.2. Solution for Edges Connecting Pairs of Vertices

We first consider a pair of distinct vertices $i_0 \neq j_0$, connected by an edge $e_{i_0j_0}$. (The case of a single vertex connected to the boundary will be dealt with later.) We denote the average by $\langle \cdots \rangle$, with

$$\langle e_{i_0j_0} \rangle = 1 \times p_{i_0j_0} + 0 \times (1 - p_{i_0j_0}) = p_{i_0j_0} \tag{7}$$

Theorem.

$$p_{i_0j_0}(1 - p_{i_0j_0}) = \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle} \tag{8}$$

Proof. We define

$$E_{i_0\setminus j_0} = \tfrac{1}{2}\left[1 + (-1)^{\sum_{j \neq j_0} e_{i_0j}}\right] \tag{9}$$

In words, $E_{i_0\setminus j_0}$ equals 1 or 0 depending on whether the vertex i_0 has an even or an odd number of connecting edges that are *on*—excluding the connection to vertex j_0. Note that the sum over j includes $j = i_0$, it only excludes $j = j_0$.

We then rewrite Equation (3) for vertex i_0 as

$$v_{i_0} = e_{i_0j_0} E_{i_0\setminus j_0} + (1 - e_{i_0j_0})(1 - E_{i_0\setminus j_0}) \tag{10}$$

Similarly, for vertex j_0, one has

$$v_{j_0} = e_{j_0i_0} E_{j_0\setminus i_0} + (1 - e_{j_0i_0})(1 - E_{j_0\setminus i_0}) \tag{11}$$

Since $e_{i_0j_0} = e_{j_0i_0} = e^2_{i_0j_0}$, the product (AND) of v_{i_0} and v_{j_0} equals

$$\begin{aligned} v_{i_0} v_{j_0} = {} & (1 - e_{i_0j_0})(1 - E_{i_0\setminus j_0} - E_{j_0\setminus i_0}) \\ & + E_{i_0\setminus j_0} E_{j_0\setminus i_0} \end{aligned} \tag{12}$$

while the binary sum (XOR) equals

$$\begin{aligned} v_{i_0} \oplus v_{j_0} \equiv {} & v_{i_0} + v_{j_0} \bmod 2 \\ = {} & E_{i_0\setminus j_0} + E_{j_0\setminus i_0} - 2 E_{i_0\setminus j_0} E_{j_0\setminus i_0} \end{aligned} \tag{13}$$

By construction, all three variables $e_{i_0j_0}$, $E_{i_0\setminus j_0}$, and $E_{j_0\setminus i_0}$ are statistically independent. The averages of the e's are given by Equation (7). The averages of the E's are unknown, but they can be eliminated by combining the four equations (10)–(13). We thus arrive at the desired Equation (8).

As a final step, we note that the left-hand-side of Equation (8) is symmetric under the exchange $p_{i_0j_0} \leftrightarrow 1 - p_{i_0j_0}$. We may safely assume that the error probabilities are $< 1/2$, resulting in the probability

$$p_{i_0j_0} = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle}} \tag{14}$$

This is an exact relation between the probability of an edge and correlators of the pair of connected vertices. These correlators are measurable from the error syndrome, without any prior knowledge of the error model. □

3.3. Solution for Boundary Edges

Theorem.

$$p_{i_0i_0} = \frac{1}{2} + \frac{\langle v_{i_0} \rangle - 1/2}{\prod_{j \neq i_0} (1 - 2p_{i_0j})} \tag{15}$$

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Proof. The probability $p_{i_0 i_0}$ of an edge $e_{i_0 i_0}$ connecting the vertex v_{i_0} to the boundary cannot be determined by a correlator, since there is nothing to correlate with. We do have access to the average

$$\langle v_{i_0}\rangle = 1 - p_{i_0 i_0} - (1 - 2p_{i_0 i_0})\langle E_{i_0\setminus i_0}\rangle \tag{16}$$

$$E_{i_0\setminus i_0} = \frac{1}{2}\left[1 + (-1)^{\sum_{j\neq i_0} e_{i_0 j}}\right] = \frac{1}{2} + \frac{1}{2}\prod_{j\neq i_0}(1 - 2e_{i_0 j}) \tag{17}$$

Using again the independence of the variables, we obtain Equation (15). So once the probabilities $p_{i_0 j}$ for non-boundary edges are determined from Equation (14), we can use Equation (15) to obtain the probability of a boundary edge. □

3.4. Computational Complexity

For local sources of error the matrix p_{ij} of error probabilities is sparse: Error events on ancilla qubits i and j are only correlated if they are close together (typically nearest neighbor or next-nearest neighbor), so the number of p_{ij}'s unequal to zero at any given time t is of the order of the number d^D of ancilla qubits (for a code of depth d in D spatial dimensions)—not to the square of that number.

The adaptive decoder uses Equations (14) and (15) to determine the error probabilities p_{ij} at time t from the averages of a stochastic variable $X \in \{v_i,\ v_i v_j,\ v_i \oplus v_j\}$. We collect ancilla measurements for N time steps to determine these averages with sufficient accuracy. As time proceeds, the averages may be stored and incremented via the recursion relation

$$\langle X(t)\rangle \equiv \frac{1}{N}\sum_{i=0}^{N-1} X(t-i) = \frac{1}{N}\sum_{i=0}^{N-1} X(t-1-i) + \frac{X(t)}{N} - \frac{X(t-N)}{N} = \langle X(t-1)\rangle + N^{-1}[X(t) - X(t-N)] \tag{18}$$

The number of operations required to increment one single p_{ij} is order 1 and there are of order d^D nonzero p_{ij}'s to determine, so the total number of time steps required for the calculation of the error probabilities is of order d^D.

Once we have the error probabilities, we can obtain the weights w_{ij} of the edges in the space-time graph by means of Equation (6), which requires inversion of the matrix $A_{ij} = (1 - \delta_{ij})p_{ij}$. Pseudocode for these operations is given in algorithm 1.

The inversion of A is time consuming, because this matrix is not sparse. It extends over $O(d^D)$ qubits in space and $O(d)$ rounds in time, and inversion of a $d^{D+1} \times d^{D+1}$ matrix requires $O(d^{2.4(D+1)})$ operations. To speed up the calculation, we used the iterative inversion method described in ref. [15], which reduces the complexity to $O(d^{2D})$.

Parallelization, which we did not implement, offers a further reduction of computational complexity. The different p_{ij}'s can all be calculated in parallel, reducing the time requirement from $O(d^D)$ to $O(1)$. On-demand parallelization of the weight calculations as described in ref. [10] reduces the average calculation of individual w_{ij}'s to $O(1)$ complexity, so that the optimal complexity of the adaptive decoder per error-correcting round is $O(1)$, independent of d or D.

Algorithm 1: Flow chart of the adaptive decoder

```
function WeightGen(errorSignals)
  for each non-zero p_ij do
    Calculate v_i v_j and v_i ⊕ v_j for this round and store.
    Increment ⟨v_i v_j⟩ and ⟨v_i ⊕ v_j⟩ via Equation (18)
    Calculate p_ij via Equation (14).
  end for
  for each p_i do
    Store v_i for this round.
    increment ⟨v_i⟩
    Calculate p_i via Equation (15)
  end for
    Construct matrix A, extending back (d + 1)/2 rounds.
    Invert matrix 1 − A.
    Extract weights w_ij following Equation (6).
    Select and return required weights.
end function
```

4. Implementation of the Adaptive Decoder

4.1. Convergence in the Large-Time Limit

We test the adaptive decoder on the repetition code of Figure 1, for a bit-flip error model. Between successive `CNOT` gates and before each measurement, each qubit i is flipped independently with probability γ_i. Though this is a fairly simple model, we stress that our adaptive technique is not limited to this, and converges to optimal weights for minimum-weight perfect matching regardless of the error model.

We implement the blossom decoder without any prior knowledge of the error probabilities, using Equations (14) and (15) to determine them from the measured syndrome data. We assume local sources of error and set $p_{ij} \equiv 0$ for ancilla measurements i and j that are not connected by any local error. In a nonlocal situation, for example, because of non-negligible crosstalk, a proliferation of negligibly small error probabilities can be avoided by setting $p_{ij} \equiv 0$ when the deviation from zero is statistically insignificant.

The adaptive decoder needs sufficient syndrome data in the training stage to estimate the probabilities. Since p_{ij} is the mean of a Bernoulli random variable with variance $\sigma_{ij}^2 = p_{ij}(1 - p_{ij})$, the statistical uncertainty δp_{ij} in the estimation after $N = t/\delta t$ error cycles is of order

$$\delta p_{ij} = N^{-1/2}\sqrt{p_{ij}(1 - p_{ij})} \tag{19}$$

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

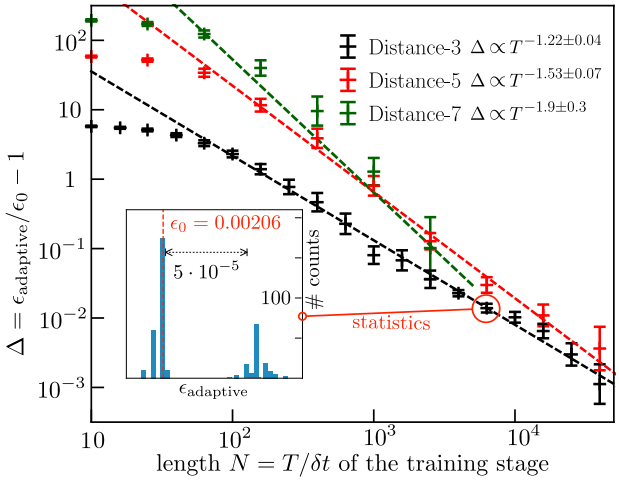

Figure 3. Convergence of the adaptive decoder towards the ideal blossom decoder, as determined by the relative decoder error Δ as a function of the number of cycles N used to estimate the error probabilities in the training stage. Each data point with error bars results from the repetition of N training stages, with $N = 400$ for distance-3, $N = 200$ for distance-5, and $N = 100$ for distance-7. The inset shows the statistics for one particular data point. The dashed lines through the data points are fits to the observed power-law decay (coefficients reported in labels).

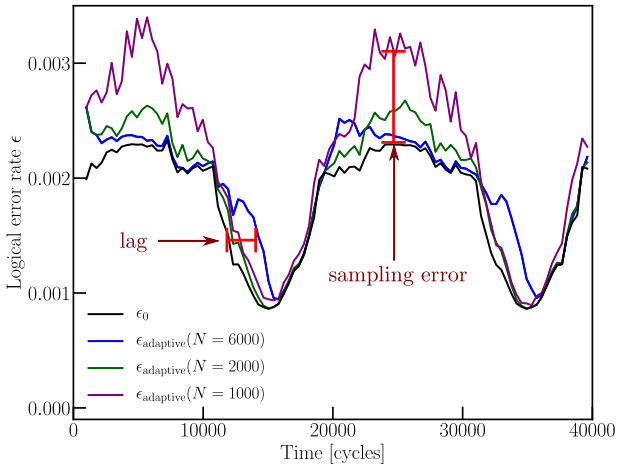

Figure 4. Performance of the adaptive decoder in the presence of a fluctuating noise ($d = 3$, $\gamma_i = 0.005$ for data qubits, $\gamma_i = 0.005 + 0.005 \sin(\pi t/10^4 \delta t)$ for ancilla qubits) using three different time windows $T = N\delta t$ for the error estimation. The average over 200 training stages is compared to a blossom decoder (black) with optimally chosen weights at every point in time. Small time windows suffer from sampling error, but adapt quickly to changing error rates, while a decoder with a larger time window lags behind. The optimal time window that balances the two effects is around $T = 2000\,\delta t$ in this case.

The requirement that $\delta p_{ij} \ll p_{ij} \ll 1$ implies that a minimum of

$$N_{\min} \simeq 1/\bar{p} \tag{20}$$

measurements are needed for a reliable estimation of error probabilities of typical magnitude $\bar{p}$. After the training stage the probabilities are inserted into Equation (6) to determine the weights which are passed to the blossom decoder for error correction.

As a figure of merit, we introduce a testing stage after the training stage in which we calculate the probability $\epsilon_{\text{adaptive}}(N)$ of a logical error per cycle using the adaptive decoder trained on N cycles of data. The error rates are calculated following the method of ref. [15], measuring the average logical qubit fidelity over 100 cycles. The combination of training and testing is repeated a few hundred times to obtain an accurate value of $\epsilon_{\text{adaptive}}(N)$. We compare this with the probability ϵ_0 that would follow from a blossom decoder with pre-determined weights calculated from the error model. The relative error

$$\Delta = \epsilon_{\text{adaptive}}/\epsilon_0 - 1 \tag{21}$$

measures how well the adaptive decoder has converged to the ideal blossom decoder.

Results are shown in **Figure 3**, for distance $d = 3$, 5, and 7 repetition codes with uniform single-qubit error rate $\gamma_i = 5 \cdot 10^{-3}$. We observe a power law convergence $\Delta \propto N^{-\alpha}$ with α increasing from 1.2 to 1.9 as the distance increases to $d = 7$. (We do not have an analytical result for this exponent.) The data indicates a convergence to $\Delta \lesssim 10^{-2}$ within $N = 10^4$ cycles.

4.2. Performance in a Time-Dependent Environment

The adaptive decoder can be readily applied to sources of noise that vary in time, by recalibration of the weights as time proceeds. We implement this by estimating the error probabilities at time t from the syndrome data in the time interval $(t - T, t)$. The optimal time window $T = N\delta t$ should not be too short in view of the statistical error (19), and it should not be too large in view of the variation $\omega T p_{ij}$ of the probabilities in the time-dependent environment (with characteristic frequency ω). The sum of these sources of error is minimized for

$$N_{\text{opt}} \simeq (p_{ij}\omega^2 \delta t^2)^{-1/3} \Rightarrow \delta p_{ij}^{\text{opt}} \simeq p_{ij}^{2/3} (\omega \delta t)^{1/3} \tag{22}$$

The adaptive decoder fails if the noise fluctuates too rapidly to acquire sufficient data for the probability estimation. The condition $\delta p_{ij}^{\text{opt}} \ll p_{ij}$ implies an upper bound

$$\omega_c \simeq \bar{p}/\delta t \tag{23}$$

on the frequency of the noise variations that is adaptable for a typical error probability $\bar{p}$.

We test the adaptive decoder in the presence of time dependent errors by taking $\gamma_i = \gamma_0$ for the data qubits and $\gamma_i = \gamma_0(1 + \sin \omega t)$ for the ancilla qubits (with $\gamma_0 = 5 \cdot 10^{-3}$ and $2\pi/\omega = 2 \cdot 10^4\, \delta t$). The predicted optimal time window at this frequency, for $\bar{p} = 5 \cdot 10^{-3}$, is $N_{\text{opt}} \approx 1265$. As shown in **Figure 4**, when a larger window $N \gg N_{\text{opt}}$ is used, the decoder experiences a time lag in determining optimal weights; for a smaller window $N < N_{\text{opt}}$ the weight estimation is degraded by sampling errors.

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

5. Conclusion

We have demonstrated that it is possible to analytically calculate the underlying error probabilities from measured error syndromes in a broad class of stabilizer codes. As this requires inverting a set of non-linear equations, it is surprising that it should be possible at all, let alone with such small overhead. Because the inversion is exact, the convergence of our adaptive decoder to the ideal blossom decoder should be optimal in the absence of additional information about the error rates. This implies that fluctuations faster than a critical frequency ω_c are uncorrectable; we have estimated $\omega_c \simeq \bar{p}/\delta t$, with $\bar{p}$ the single-qubit error probability and δt the duration of one error-correction cycle. Such rapid fluctuations will contribute relatively more to the logical error rate of a quantum error correcting code than slow fluctuations to which the decoder can adapt.

It would be interesting for future work to test the adaptive decoder on more complex noise models, where the optimal window must be chosen for an entire noise frequency spectrum, instead of for a single frequency. We expect white noise to be significantly worse for quantum error correction than $1/f$ noise, due to the much larger contributions from high frequencies. Future work could also extend our results to simulations of the surface code or topological cluster states.

Appendix A: Derivation of Equation (6)

Starting from the definition of the matrix $A_{ij} = (1-\delta_{ij})\,p_{ij}$, we evaluate the inverse $(1-A)^{-1}$ by series expansion

$$\begin{aligned}[(1-A)^{-1}]_{ij} &= \delta_{ij} + (1-\delta_{ij})\,p_{ij} + \sum_{k_1}(1-\delta_{ik_1})(1-\delta_{k_1 j})\,p_{ik_1}p_{k_1 j} \\ &\quad + \sum_{k_1,k_2}(1-\delta_{ik_1})(1-\delta_{k_1k_2})(1-\delta_{k_2 j})\,p_{ik_1}p_{k_1k_2}p_{k_2 j} + \cdots \end{aligned} \tag{A1}$$

For $i \neq j$ this reduces to

$$\begin{aligned}[(1-A)^{-1}]_{ij} &= p_{ij} + \sum_{k_1\neq i,j} p_{ik_1}p_{k_1 j} \\ &\quad + \sum_{k_1\neq i,k_2}\sum_{k_2\neq k_1,j} p_{ik_1}p_{k_1k_2}p_{k_2 j} + \cdots \end{aligned} \tag{A2a}$$

$$= e^{-w_{ij}} \tag{A2b}$$

with weight w_{ij}, $i \neq j$ defined in Equation (4).

For the boundary weight we evaluate

$$\begin{aligned}\sum_k [(1-A)^{-1}]_{ik}\,p_{kk} &= \sum_k \delta_{ik}\,p_{kk} + \sum_k (1-\delta_{ik})\,p_{ik}\,p_{kk} \\ &\quad + \sum_{k,k_1}(1-\delta_{ik_1})(1-\delta_{k_1k})\,p_{ik_1}\,p_{k_1k}\,p_{kk} \\ &\quad + \sum_{k,k_1,k_2}(1-\delta_{ik_1})(1-\delta_{k_1k_2})(1-\delta_{k_2k})\,p_{ik_1}\,p_{k_1k_2}\,p_{k_2k}\,p_{kk} + \cdots \end{aligned} \tag{A3a}$$

$$= p_{ii} + \sum_{k\neq i} p_{ik}\,p_{kk} + \sum_{k\neq i}\sum_{k_1\neq i,k} p_{ik_1}\,p_{k_1k}\,p_{kk} + \sum_{k\neq k_2}\sum_{k_1\neq i,k_2}\sum_{k_2\neq k_1,k} p_{ik_1}\,p_{k_1k_2}\,p_{k_2k}\,p_{kk} + \cdots \tag{A3b}$$

$$= e^{-w_{ii}} \tag{A3c}$$

with weight w_{ii} defined in Equation (5).

Taken together, these results constitute Equation (6).

Appendix B: Intermediate Steps for the Derivation of the Inversion Formulas in Section 3

B.1. Derivation of Equation (14)

We start from the definition

$$v_i = \tfrac{1}{2}\left[1-(-1)^{\sum_j e_{ij}}\right] \tag{B1}$$

and rewrite this using the fact that $e_{ij} \in \{0,1\}$:

$$v_{i_0} \equiv \tfrac{1}{2}\left[1-(-1)^{\sum_j e_{i_0 j}}\right] \tag{B2a}$$

$$= \begin{cases} 0 & \text{if } \sum_j e_{i_0 j} \text{ even} \\ 1 & \text{if } \sum_j e_{i_0 j} \text{ odd} \end{cases} \tag{B2b}$$

$$= \begin{cases} e_{i_0 j_0} & \text{if } \sum_{j\neq j_0} e_{i_0 j} \text{ even} \\ 1-e_{i_0 j_0} & \text{if } \sum_{j\neq j_0} e_{i_0 j} \text{ odd} \end{cases} \tag{B2c}$$

$$\begin{aligned} &= e_{i_0 j_0}\tfrac{1}{2}\left[1+(-1)^{\sum_{j\neq j_0} e_{i_0 j}}\right] \\ &\quad +(1-e_{i_0 j_0})\tfrac{1}{2}\left[1-(-1)^{\sum_{j\neq j_0} e_{i_0 j}}\right] \end{aligned} \tag{B2d}$$

With the definition

$$E_{i_0\setminus j_0} = \tfrac{1}{2}\left[1+(-1)^{\sum_{j\neq j_0} e_{i_0 j}}\right] \tag{B3}$$

we thus obtain

$$v_{i_0} = e_{i_0 j_0}\,E_{i_0\setminus j_0} + (1-e_{i_0 j_0})(1-E_{i_0\setminus j_0}) \tag{B4}$$

The same algebra with the indices i_0 and j_0 interchanged gives

$$v_{j_0} = e_{j_0 i_0}\,E_{j_0\setminus i_0} + (1-e_{j_0 i_0})(1-E_{j_0\setminus i_0}) \tag{B5}$$

Now we multiply v_{i_0} and v_{j_0}, using $e_{i_0 j_0} = e_{j_0 i_0}$,

$$\begin{aligned} v_{i_0}v_{j_0} &= \left[e_{i_0 j_0}\,E_{i_0\setminus j_0} + (1-e_{i_0 j_0})(1-E_{i_0\setminus j_0})\right] \\ &\quad \times \left[e_{j_0 i_0}\,E_{j_0\setminus i_0} + (1-e_{j_0 i_0})(1-E_{j_0\setminus i_0})\right] \end{aligned} \tag{B6a}$$

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

$$= [1 - e_{i_0 j_0} + (2e_{i_0 j_0} - 1)\, E_{i_0 \setminus j_0}][1 - e_{j_0 i_0} + (2e_{j_0 i_0} - 1)\, E_{j_0 \setminus i_0}] \tag{B6b}$$

$$= (1 - e_{i_0 j_0})^2 + (2e_{i_0 j_0} - 1)^2\, E_{i_0 \setminus j_0}\, E_{j_0 \setminus i_0} + (1 - e_{i_0 j_0})(2e_{i_0 j_0} - 1)(E_{i_0 \setminus j_0} + E_{j_0 \setminus i_0}) \tag{B6c}$$

Since $e_{i_0 j_0}^2 = e_{i_0 j_0}$, this simplifies to

$$v_{i_0} v_{j_0} = (1 - e_{i_0 j_0})(1 - E_{i_0 \setminus j_0} - E_{j_0 \setminus i_0}) + E_{i_0 \setminus j_0}\, E_{j_0 \setminus i_0} \tag{B7}$$

In the same way we evaluate the sum of v_{i_0} and v_{j_0},

$$v_{i_0} + v_{j_0} = 1 - e_{i_0 j_0} + (2e_{i_0 j_0} - 1)\, E_{i_0 \setminus j_0} + 1 - e_{j_0 i_0} + (2e_{j_0 i_0} - 1)\, E_{j_0 \setminus i_0} \tag{B8a}$$

$$= 2(1 + e_{i_0 j_0} E_{i_0 \setminus j_0} + e_{j_0 i_0} E_{j_0 \setminus i_0}) - e_{i_0 j_0} - e_{j_0 i_0} - E_{i_0 \setminus j_0} - E_{j_0 \setminus i_0} \tag{B8b}$$

$$= 2(1 + e_{i_0 j_0} E_{i_0 \setminus j_0} + e_{i_0 j_0} E_{j_0 \setminus i_0} - e_{i_0 j_0}) - E_{i_0 \setminus j_0} - E_{j_0 \setminus i_0} \tag{B8c}$$

Modulo 2, we are left with

$$v_{i_0} + v_{j_0} \bmod 2 = -E_{i_0 \setminus j_0} - E_{j_0 \setminus i_0} \bmod 2 = E_{i_0 \setminus j_0} + E_{j_0 \setminus i_0} - 2E_{i_0 \setminus j_0}\, E_{j_0 \setminus i_0} \bmod 2 \tag{B9}$$

The final expression can only take the values 0 or 1, so the modulo 2 indication is superfluous. We have thus arrived at

$$v_{i_0} \oplus v_{j_0} = E_{i_0 \setminus j_0} + E_{j_0 \setminus i_0} - 2E_{i_0 \setminus j_0}\, E_{j_0 \setminus i_0} \tag{B10}$$

At this point we restrict ourselves to $i_0 \neq j_0$. (The case $i_0 = j_0$ is addressed in the next subsection.) We can then use the statistical independence of $e_{i_0 j_0}$, $E_{i_0 \setminus j_0}$, and $E_{j_0 \setminus i_0}$ to break up the averages of products into products of averages. (This decoupling explains why we excluded the $e_{i_0 j_0}$ term from the definition of E.) We denote the averages by

$$\langle e_{i_0 j_0} \rangle = p_{i_0 j_0} = \langle e_{j_0 i_0} \rangle,$$
$$\langle E_{i_0 \setminus j_0} \rangle = U_{i_0 j_0}, \quad \langle E_{j_0 \setminus i_0} \rangle = U_{j_0 i_0} \tag{B11}$$

The averages of Equations (B4), (B5), (B7), and (B10)a give the four equations

$$\langle v_{i_0} \rangle = p_{i_0 j_0} U_{i_0 j_0} + (1 - p_{i_0 j_0})(1 - U_{i_0 j_0}) \tag{B12}$$

$$\langle v_{j_0} \rangle = p_{i_0 j_0} U_{j_0 i_0} + (1 - p_{i_0 j_0})(1 - U_{j_0 i_0}) \tag{B13}$$

$$\langle v_{i_0} v_{j_0} \rangle = (1 - p_{i_0 j_0})(1 - U_{i_0 j_0} - U_{j_0 i_0}) + U_{i_0 j_0} U_{j_0 i_0} \tag{B14}$$

$$\langle v_{i_0} \oplus v_{j_0} \rangle = U_{i_0 j_0} + U_{j_0 i_0} - 2U_{i_0 j_0} U_{j_0 i_0} \tag{B15}$$

Subtraction of the product of the first two equations from the third equation gives

$$\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle = p_{i_0 j_0}(1 - p_{i_0 j_0})(1 - 2U_{i_0 j_0} - 2U_{j_0 i_0} + 4U_{i_0 j_0} U_{j_0 i_0}) \tag{B16}$$

and substitution of the fourth equation results in

$$\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle = p_{i_0 j_0}(1 - p_{i_0 j_0})(1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle) \tag{B17}$$

We solve the quadratic equation for $p_{i_0 j_0}$

$$p_{i_0 j_0} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle}} \tag{B18}$$

There are two solutions, related by $p_{i_0 j_0} \leftrightarrow 1 - p_{i_0 j_0}$. Under the assumption that $p_{i_0 j_0} < 1/2$ we select the smallest of the two solutions

$$p_{i_0 j_0} = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\langle v_{i_0} v_{j_0} \rangle - \langle v_{i_0} \rangle \langle v_{j_0} \rangle}{1 - 2\langle v_{i_0} \oplus v_{j_0} \rangle}} \tag{B19}$$

This is Equation (14) from the main text.

B.2. Derivation of Equation (15)

Turning now to the case $i_0 = j_0$, we start from

$$E_{i_0 \setminus i_0} \equiv \tfrac{1}{2}\left[1 + (-1)^{\sum_{j \neq i_0} e_{i_0 j}}\right] = \tfrac{1}{2} + \tfrac{1}{2} \prod_{j \neq i_0} (1 - 2e_{i_0 j}) \tag{B20}$$

and average, using again the independence of the variables e_{ij}:

$$\langle E_{i_0 \setminus i_0} \rangle \equiv U_{i_0 i_0} = \tfrac{1}{2} + \tfrac{1}{2} \prod_{j \neq i_0} (1 - 2p_{i_0 j}) \tag{B21}$$

This can be substituted into the average of

$$v_{i_0} = e_{i_0 i_0} E_{i_0 \setminus i_0} + (1 - e_{i_0 i_0})(1 - E_{i_0 \setminus i_0}) \tag{B22}$$

to arrive at

$$\langle v_{i_0} \rangle = p_{i_0 i_0} U_{i_0 i_0} + (1 - p_{i_0 i_0})(1 - U_{i_0 i_0}) \tag{B23}$$

$$= 1 - p_{i_0 i_0} - (1 - 2p_{i_0 i_0}) U_{i_0 i_0} \tag{B24}$$

$$= \tfrac{1}{2} - \tfrac{1}{2}(1 - 2p_{i_0 i_0}) \prod_{j \neq i_0} (1 - 2p_{i_0 j}) \tag{B25}$$

Solving for $p_{i 0 i_0}$ we find

$$p_{i_0 i_0} = \frac{1}{2} + \frac{\langle v_{i_0} \rangle - 1/2}{\prod_{j \neq i_0} (1 - 2p_{i_0 j})} \tag{B26}$$

which is Equation (15) from the main text.

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO/OCW), an ERC Synergy Grant, and by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via the U.S. Army Research Office grant W911NF-16-1-0071. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

quantum computing, quantum error correction, surface code

Received: February 6, 2018
Revised: June 6, 2018
Published online: July 30, 2018

[1] *Quantum Error Correction* (Eds: D. A. Lidar, T. A. Brun), Cambridge University Press, Cambridge **2013**.
[2] B. M. Terhal, *Rev. Mod. Phys.* **2015**, *87*, 307.
[3] P. W. Shor, *Phys. Rev. A* **1995**, *52*, R2493.
[4] A. Steane, *Proc. Roy. Soc. Lond. A* **1996**, *452*, 2551.
[5] S. B. Bravyi, A. Yu. Kitaev, arXiv:quant-ph/9811052, **1998**.
[6] A. G. Fowler, A. C. Whiteside, L. C. L. Hollenberg, *Phys. Rev. Lett.* **2012**, *108*, 180501.
[7] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, J. M. Martinis, *Nature* **2015**, *519*, 66.
[8] R. Raussendorf, J. Harrington, K. Goyal, *New. J. Phys.* **2007**, *9*, 199.
[9] J. Edmonds, *Canad. J. Math.* **1965**, *17*, 449.
[10] A. G. Fowler, *Quantum Inf. Comp.* **2015**, *15*, 0145.
[11] A. G. Fowler, arXiv:1310.0863, **2013**.
[12] N. Delfosse, J.-P. Tillich, *IEEE Int. Symp. on Information Theory*, Hawaii **2014**, pp. 1071–1075.
[13] B. Heim, K. M. Svore, M. B. Hastings, arXiv:1609.06373, **2016**.
[14] P. Baireuther, T. E. O'Brien, B. Tarasinski, C. W. J. Beenakker, arXiv:1705.07855, **2018**.
[15] T. E. O'Brien, B. Tarasinski, L. DiCarlo, arXiv:1703.04136, **2017**.
[16] A. G. Fowler, A. C. Whiteside, A. L. McInnes, A. Rabbani, *Phys. Rev. X* **2012**, *2*, 041003.
[17] A. G. Fowler, D. Sank, J. Kelly, R. Barends, J. M. Martinis, arXiv:1405.1454, **2014**.
[18] J. Combes, C. Ferrie, C. Cesare, M. Tiersch, G. J. Milburn, H. J. Briegel, C. M. Caves, arXiv:1405.5656, **2014**.
[19] Y. Fujiwara, arXiv:1405.6267, **2014**.
[20] D. Orsucci, M. Tiersch, H. J. Briegel, *Phys. Rev. A* **2016**, *93*, 042303.
[21] M.-X. Huo, Y. Li, arXiv:1710.03636, **2017**.
[22] S. J. Devitt, K. Nemoto, W. J. Munro, *Rep. Prog. Phys.* **2013**, *76*, 076001.

© 2018 The Authors. Published by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.