Supplementary Information for
Topological quantum computing with a very noisy network and local error rates approaching one percent
Naomi H. Nickerson, Ying Li, Simon C. Benjamin

Supplementary Figure S1: Comparison to the monolithic case

Performance of the monolithic architecture. This Figure is to be compared with the graphs in Fig. 4 of the main paper. In order to make a clear comparison with the threshold that can be achieved with the monolithic architecture, we applied our same superoperator description and numerical simulation to this case. (Of course there is no meaning for $p_n$, the network error rate, in a monolithic architecture). The figure shows the circuit we used for the case of a Z stabilizer. Note that this approach requires initialisation of the single shared auxiliary qubit; we used the value of $p_m$ as the fidelity of this initialisation. One sees that the threshold is between 0.9% and 0.95% and is therefore appreciably higher than that obtained by STRINGENT whilst in the same “ball park”. As noted below, the addition of another ancilla per cell can further close this gap.
Supplementary Figure S2: Sweeping the network error rate

Varying network error rate with fixed local error rates. In the main paper we set the network error rate to 10% and consider a range of local (intra-cell) error rates in order to determine the threshold. We find thresholds in the range 0.6% to 0.82% depending on the details of the protocol. We have also performed a complimentary series of simulations where we fix the local error rate and sweep the network error rate. Supplementary Figure 2 shows the corresponding data for the case of the EXPEDIENT protocol. Fixing the local error rates $p_m = p_g = 0.6$, our sweep of the network error rate reveals a threshold in the 10.0% to 10.1% range, consistent with the calculations in the main paper.

Supplementary Tables

| Level, $L$ | Time steps, $t_L$ | $p_{\text{success}, L}$ | FRL |
|-----------|------------------|------------------------|-----|
| 1         | Round one Bell pair production | 7 | 0.7346 | 1 |
| 2         | Round two Bell pair production | 6 | 0.7506 | 1 |
| 3         | Round one single selection | 4 | 0.8619 | 3 |
| 4         | Round two single selection | 3 | 0.8550 | 3 |
| 5         | Make GHZ | 2 | 0.8651 | 1 |
| 6         | Round one single selection | 4 | 0.8619 | 6 |
| 7         | Round two single selection | 3 | 0.8550 | 7 |
| 8         | Check GHZ | 2 | 0.8654 | 1 |
| 9         | Measure stabilizer | 2 | - | - |

Supplementary Table S1: Listing of probabilities used in modelling the number of steps required for EXPEDIENT. FRL stands for Failure Reset Level. Please see Supplementary Note 2 for further details.
| Level, $L$ | Time steps, $t_L$ | $p_{\text{success},L}$ | FRL |
|-----------|-----------------|-----------------|-----|
| 1         | Bell pair production | 7               | 0.7277 | 1   |
| 2         | Bell pair check 1   | 6               | 0.7429 | 1   |
| 3         | Round one single selection | 4               | 0.8586 | 3   |
| 4         | Round two single selection | 3               | 0.8509 | 3   |
| 5         | Bell pair check 2   | 5               | 0.8019 | 1   |
| 6         | Round one single selection | 4               | 0.8586 | 6   |
| 7         | Round two single selection | 3               | 0.8509 | 6   |
| 8         | Bell pair check 3   | 5               | 0.8043 | 1   |
| 9         | Round one single selection | 4               | 0.8586 | 9   |
| 10        | Round two single selection | 3               | 0.8509 | 9   |
| 11        | Make GHZ           | 5               | 0.6588 | 1   |
| 12        | Round one single selection | 4               | 0.8586 | 12  |
| 13        | Round two single selection | 3               | 0.8509 | 12  |
| 14        | Check GHZ          | 5               | 0.6454 | 1   |
| 15        | Measure stabilizer | 2               | -     | -   |

**Supplementary Table S2:** As Supplementary Table S1, but now for STRINGENT protocol with $p_g = p_m = 0.75\%$, $p_n = 10\%$. Please see Supplementary Note 2 for further details.

| Label for $Z$ proj | Label for $X$ proj | Case 1 | Case 2 | Case 3 |
|-------------------|-------------------|--------|--------|--------|
| $A_1$             | $A_3$             | 0.9117 | 0.928  | 0.951  |
| $A_Z$             | $A_X$             | 0.00681| 0.00675| 0.00470|
| $A_X$             | $A_Z$             | 0.00314| 0.00391| 0.00352|
| $A_Y$             | $A_Y$             | 0.00314| 0.00391| 0.00352|
| $A_ZZ$            | $A_{XX}$          | 0.000182| 0.000187| 0.00119|
| $A_{XX}$          | $A_{ZZ}$          | 0.00000758| 0.00001182| 0.0000128|
| $A_{YY}$          | $A_{YY}$          | 0.00000758| 0.00001182| 0.0000128|
| $A_{XZ}$          | $A_{XZ}$          | 0.0000336| 0.0000414| 0.00121|
| $A_{YZ}$          | $A_{YZ}$          | 0.0000336| 0.0000414| 0.00121|
| $A_{XY}$          | $A_{YZ}$          | 0.0000152| 0.0000236| 0.0000256|
| $B_1$             | $A_3$             | 0.0617 | 0.0424 | 0.0178 |
| $B_Z$             | $B_X$             | 0.00674| 0.00665| 0.00470|
| $B_X$             | $B_Z$             | 0.00314| 0.00391| 0.00352|
| $B_Y$             | $B_Y$             | 0.00314| 0.00391| 0.00352|
| $B_{ZZ}$          | $B_{XX}$          | 0.000127| 0.0000849| 0.00119|
| $B_{XX}$          | $B_{ZZ}$          | 0.00000758| 0.00001182| 0.0000128|
| $B_{YY}$          | $B_{YY}$          | 0.00000758| 0.00001182| 0.0000128|
| $B_{XZ}$          | $B_{XZ}$          | 0.0000336| 0.0000414| 0.00121|
| $B_{YZ}$          | $B_{XY}$          | 0.0000336| 0.0000414| 0.00121|
| $B_{XY}$          | $B_{YZ}$          | 0.0000152| 0.0000236| 0.0000256|

**Supplementary Table S3:** Listing of weighting coefficients for three examples superoperators, as explained in the Supplementary Methods.
Supplementary Note 1: Five qubits per cell

The protocols EXPEDIENT and STRINGENT use a total of four qubits per cell of the network (one data qubit, three ancillas). If one adds further ancilla(s) then the performance improves. While it is beyond the scope of the present paper to seek optimal protocols that exploit five (or more) qubits, we can easily modify our approaches to make some use of the additional resource to tolerate more severe network error rates. For example, wherever our original 4-qubit STRINGENT protocol calls for a “raw” Bell pair to be created on a given ancilla pair we can instead insert a small circuit that creates raw pairs using the additional ancillas and purifies them. Similarly one can replace “single selection” purification (two tiers of ancilla yield an improved pair on the higher tier) with “double selection” (three tiers of ancillas yield a significantly improved pair on the highest tier). Adopting these rather naive modifications we immediately find that the network noise is tolerable at the $p_n = 0.2$ level (rather than $p_n = 0.1$) for the same threshold of 0.77% local errors.

Alternatively we can hold the network errors constant and improve our tolerance of intra-cell errors. Data for this case are shown in the third panel of Fig. 4 in the main paper, labelled “STRINGENT+”. We have introduced one further enhancement: We replace the usual Phase 3 of the STRINGENT protocol i.e. the steps where the GHZ resource would simply be coupled to the data qubits and then measured out. Instead we filter the GHZ after coupling it to the data qubits, and if this filter fails we abort the protocol by measuring the GHZ in the Z-basis (this therefore requires the addition of Z-basis measurement to our set of allowed primitive operations). If we have aborted, we then perform a whole new round of stabilization, this time without the filter so that the protocol will certainly complete. This procedure leaves us knowing the stabilizer outcome and some additional classical information about exactly what steps occurred. Specifically there are 3 distinct cases: (a) The filter was successfully passed; this is the best case and results in lower error rates on the data qubits – it is about 92% of cases in for the parameter range we consider. (b) The filter failed, but on measuring the GHZ in the Z-basis it was found to be in a correct state (e.g. 0000 or 1111). This occurs about 4% of the time. It is another “good” case in that there is a low chance of errors having reached the data qubits, so that the second round can perform normally. (c) The filter failed, and on measuring on the GHZ in the Z-basis it was found to be in an incorrect state (e.g. 0001). This occurs about 4% of the time and it is the “bad” case; in this event there is very likely to be an error on the data qubits. Now if we make no use of the classical information and merely “forget it” then the net effect of this protocol is to make things worse versus a simple one-round use of STRINGENT – this is not surprising since the overall risk of an error is not reduced (it is increased slightly due to the extra steps). However if we modify our Edmonds matching algorithm to use the classical information, specifically to favour paths that are consistent with errors occurring where “bad” stabilizers took place, then the threshold improves (this is the case shown in the third panel of Fig. 4). In effect we trade a small amount of increased error risk for a significant amount of classical knowledge: the 4% of “bad” stabilizers account for about half of all errors entering the system.

Supplementary Note 2: Time costs and memory errors

Since our stabilizer measurement protocols are considerably longer than the equivalent procedures for a monolithic architecture, and indeed our approach is post-selective and therefore of uncertain duration, it is important to assess the potential impact of memory errors. In the main paper we assert that memory errors should have negligible impact, if our cells can employ qubits at least as good as those demonstrated in the Science papers of Maurer et al and Steger et al (Refs. [23] and [24] of the main paper). To substantiate this we need to estimate the duration of the protocols. In the following we assign one “time step” to any elementary operation in our protocol, whether a gate, a raw Bell creation or a measurement. We assume that the operations within a cell must be strictly sequential. Because we are evaluating an entire ‘sheet’ of stabilizers in parallel across a large array of many cells (see Fig. 3 of the main paper), we need to be concerned not merely with the average time that a stabilizer protocol might take, but rather with the the time required for a given target proportion of all stabilizers to succeed.

In Supplementary Table S1 and Supplementary Table S2 we give data for both EXPEDIENT and STRINGENT. In the following analysis we will focus on the former since it is the protocol designed for use when
memory errors are an issue. The process of building a GHZ state can be separated into distinct sections, each of which is terminated by a measurement, which, if it results in the ‘wrong’ outcome, will reset the process to an earlier stage. Each of these measurement outcomes has an associated probability of success, \( p_L \), where the index \( L \) denotes the level. For the EXPEDIENT protocol operating at error rates of \( p_g = p_m = 0.6\% \) and \( p_n = 10\% \), these levels and their probabilities are given in Supplementary Table S1. It should be noted that levels 1 and 2 utilise two cells (and 3, 4, 6, 7) are run twice in parallel across four cells, and consequently the longer of the two times will determine when the process may proceed to the next level.

The GHZ production process is probabilistic, with a minimum duration of \( \sum L t_L = 33 \), which occurs with a probability \( \prod L p_L = 0.2242 \). Using the parameters in Supplementary Table S1 the general process was simulated; 100,000 samples were generated and used to estimate the parameters of the resulting distribution. This found the expected completion time for a given stabilizer measurement to be to be 68.2 time steps, while 50\% of operations were completed after 57 time steps, 95\% after 138 time steps, 99\% after 195 time steps and 99.9\% after 278 time steps.

Our topological code simulations indicate that if we wait for 99\% of stabilizers to evaluate, and abandon the remaining 1\%, then there is negligible impact on the threshold. (A stabilizer measurement that fails to report is less damaging than a stabilizer measurement that performs a ‘wrong’ projection; moreover such stabilizers will not introduce errors since the ancilla GHZ state is never created and coupled to the data qubits.) Therefore we simply take the expected time for 99\% of stabilizer protocols to complete as the characteristic time for a ‘sheet’ of stabilizers to be evaluated. Obviously there is scope for far more sophisticated approaches which minimise waiting time, but this naive strategy suffices to give us a bound. The 195 steps required for 99\% completion are of course a mix of different operations: remote entanglement, local gates and measurement. From the physical timescales summarised in Supplementary Note 3 below, a picture emerges where 10\(\mu\text{s}\) may be a reasonable average time for an operation (assuming that there is progress on the crucial technical issues of photon loss, so that the more fundamental limits can be approached). We therefore estimate that 195 steps will take approximately 2 milliseconds.

How severe will memory errors be in such a period? Experimentally reported memory qubit lifetimes have dramatically improved over the last couple of years, and one can hope that this may continue. As a note of caution we should however remember that our protocol will require many manipulations of other, nearby qubits while our memory qubits remain passive; memory performance in such a complicated environment has not been established. Maurer et al (Ref. [23] of the main paper) reported that their best NV memory qubit had lifetimes of about 2 seconds (note the silicon impurity qubits studied by Steger et al (Ref. [24]) survived far longer, \(\sim 3\) mins). Taking the shorter lifetime, if a proportion \(1/e\) survive 2 seconds we infer an error rate of about 40\% per second, or a rate of 0.1\% over our 2ms protocol. This is nearly an order of magnitude below the error rates from the active processes, i.e. the gates and measurements, considered in the main paper. Obviously using the far longer lived silicon memory qubits would lead to an absolutely negligible error rate.

In comparison, the STRINGENT protocol takes about five times longer to achieve the same 99\% of complete stabilizers (see Supplementary Table S2). The minimum duration is 63 time steps, which occurs with a probability 0.0422. Simulation of the distribution found the mean duration of the stabilizer to be 278 time steps, while 50\% of operations were completed after 211 time steps, 95\% after 718 time steps, 99\% after 1067 time steps and 99.9\% after 1537 time steps. These higher costs in STRINGENT are the price paid for the increased error threshold as noted in the main paper. Whether EXPEDIENT or STRINGENT is the better protocol to adopt therefore depends on the relative severity of ‘active’ errors associated with stabilizer measurement versus ‘passive’ memory errors.

**Supplementary Note 3: Physical timescales**

Operations such as measurements, long range entanglement, and local qubit-qubit gates have timescales that vary considerably from one class of physical system to another. Moreover in some cases there is substantial potential for technological improvement, whereas other timescale are already close to fundamental limits.
So called single shot optical measurement has now been accomplished in several classes of system (see e.g. PRL 100 200502; Science 329 599; Nature 467 687). Typically there is a key period during which a driving laser gives rise to photon emission from the qubit system if, and only if, it is in a given state. Photon detector(s) monitor the qubit to observe such photons. The duration needs to be long enough that it is very likely at least one photon will be detected, if indeed the qubit is in the optically active state. Typical periods are of the order of 100µs, although notably a 10µs measurement achieving a fidelity beyond our requirements has been reported for trapped Ca ions (PRL 100 200502). Also a 5.5µs window has been employed in an NV system, albeit with limited fidelity (Nature 477 574). In any case there is good scope to shorten the detection timescale in all systems suffering high rates of photon loss (a ubiquitous issue in current experiments) – given more efficient detectors and less lossy optical interfaces, one may expect to see the measurement time fall below 10µs.

The timescale for achieving remote entanglement can be very long in experimental demonstrations to date. This is true of both ion trap systems (Nature 449, 68) and the very recent NV centre experiments (arXiv:1212.6136). The time requirements, which may extended to minutes, are again essentially due to the photon loss that occurs at various points in the system. Typically successful entanglement requires two specific photons to be detected, so that heralded failure occurs if either is lost. Therefore many such failures occur before eventual success. Assuming that technical advances can largely remedy these losses, we can focus on a given instance of the underlying entanglement protocol. The time for a single entanglement attempt can be dominated by the initialisation of the two qubits involved. In the work of Bernien et al (arXiv:1212.6136) this initialisation is composed of a pump phase followed by a measurement verification; the latter is affected by the photon loss issues described above, the former may be the more fundamental and was of 10µs duration. In the atomic experiment of Moehring et al (Nature 449 68) the time for a complete entanglement attempt was faster, around 2µs.

Local conditional gates performed within a given structure typically rely on exploiting interactions between the two qubits involved, such as the hyperfine interaction between an electron and a nuclear spin. The magnitude of this interaction strength sets a limit on the speed with which a conditional evolution can take place. In practice other constraints such as the intensity of applied field may limit the speed further. Timescales reported in the experimental literature can vary from a few microseconds (Science 320, 1326) to a few tens of microseconds (Nature Physics 208 464).

Supplementary Methods

Twirling

The protocols given in the main paper lead to slight irregularities between the weights associated with given errors occurring on different specific data qubits; so for example, in our superoperator the weight associated with ZZZZ might be slightly higher or lower than the weight associated with 1ZZZ1 . This is because the protocol performs $A - B$, $C - D$ pairings in Phase 1, and then $A - C$, $B - D$ pairing in Phase 2, thus the errors that survive this purification process will not be equally distributed under rotation of cell labels. While these irregularities have no significance for our threshold calculations, they do cause the weightings to have a spurious complexity. Therefore in our analysis we append onto the protocols of Fig. 2 a additional twirling operation which randomly applies swap operations between cells so as to ‘smooth’ the weightings. Physically this is equivalent to the rather perverse act of programming one’s system with a range of possible protocols, identical except for permutation of the cell labels A to D, and then applying one at random without retaining a record of the choice. But we emphasise that this is merely a theoretical convenience and there would be no need, or reason, to perform the process physically.

Example superoperators

In the main paper we wrote the following to represent the overall effect of the stabilizer protocol on four data qubits initially in state $\rho$.

$$P^M_{\text{real}}(\rho) = \sum_e a^M_e E_e P^M_{\text{ideal}}(\rho) E_e^\dagger + b^M_e E_e P^M_{\text{ideal}}(\rho) E_e^\dagger$$

(S1)
with \( E_e = (ABCD)_e \) and \( \{A, B, C, D\} \in \{1, \sigma_x, \sigma_y, \sigma_z\} \).

Here \( M \) stands for the reported outcome, “odd” or “even”, and the \( P^M_{\text{ideal}}(\cdot) \) represents the perfect parity projector in either the \( Z \) or \( X \) basis depending on which class of stabilizer one is performing. The four operators making up \( E \) are understood to act on data qubits 1 to 4 respectively, and index \( e \) runs over all their combinations. The symbol \( \bar{M} \) represents the compliment of \( M \), i.e. “odd” for \( M = \text{“even”} \) and vice versa.

We noted that this real projector is made up of a mix of the ‘correct’ and ‘incorrect’ ideal projectors together with possible Pauli errors; the various weights \( a \) and \( b \) capture their relative significance. Here we give some examples of those weightings. The full list has 256 terms but they rapidly fall in magnitude so that the latter half are extremely small; here we list the terms corresponding to all single- and two-qubit errors. For the examples we will give, in fact the superscript \( M \) can be omitted from the \( a \) and \( b \) weights, i.e. the same set of weights \( a, b \) apply in \( P^\text{even}_{\text{real}} \) as in \( P^\text{odd}_{\text{real}} \). In other words the weight of the ‘correct’ projector \( P^M_{\text{real}} \) is the same in each, as are the weights of all the various error combinations.

In Supplementary Table S3 the weights \( A \) are associated with the ‘correct’ projector \( P^M_{\text{ideal}} \) and the \( B \) weights are associated with the ‘incorrect’ projector \( P^M_{\bar{\text{ideal}}} \). The subscripts refer to the Pauli errors in the corresponding term, so that the subscript \( 1 \) indicates no errors, while subscript \( X \) means “a \( \sigma_X \) on one of the four data qubits”, and \( YZ \) means “a \( \sigma_Y \) on one of the qubits and a \( \sigma_Z \) on another”. These capital letters \( A \) and \( B \) are therefore sum of one or more of the weights \( a, b \) in Eqn. S1; for example \( A_X = a_{111}X + a_{11}X1 + a_{1}X11 + a_{X}111 \). As noted in a previous section, the twirling technique means that the individual \( a \) or \( b \) terms in such a sum in fact all have the same value.

Case 1 is the EXPEDIT protocol operating with \( p_n = 0.1 \) and \( p_m = p_g = 0.006 \). Note that many of the corresponding \( A \) and \( B \) terms are identical due to symmetries in the algorithm (including the twirling). We see that there is a 91% chance of a pure ‘correct’ projection, a 6.2% chance of a pure ‘wrong’ projection, a 0.68% chance of a correct projection followed by a single \( \sigma_Z \) error, and so on.

Case 2 is STRINGENT operating with \( p_n = 0.1 \) and \( p_m = p_g = 0.0075 \). Finally for the comparison Case 3 is the monolithic circuit where there is no network noise to consider, and we set \( p_m = p_g = 0.09 \).