



HAL
open science

Is error detection helpful on IBM 5Q chips ?

Christophe Vuillot

► **To cite this version:**

Christophe Vuillot. Is error detection helpful on IBM 5Q chips ?. Quantum Information & Computation, 2018, 18 (11&12), pp.0949–0964. 10.26421/QIC18.11-12 . hal-02351420

HAL Id: hal-02351420

<https://hal.science/hal-02351420v1>

Submitted on 6 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IS ERROR DETECTION HELPFUL ON IBM 5Q CHIPS ?

CHRISTOPHE VUILLOT^a

*1 - JARA Institute for Quantum Information, RWTH Aachen University
Physikzentrum Campus Melaten Otto-Blumenthal-Strasse, 52074 Aachen, Germany*

*2 - QuTech, TU Delft
Lorentzweg 1, 2628 CJ Delft, Netherlands*

Received (received date)

Revised (revised date)

This paper reports on experiments realized on several IBM 5Q chips which show evidence for the advantage of using error detection and fault-tolerant design of quantum circuits. We show an average improvement of the task of sampling from states that can be fault-tolerantly prepared in the $[[4, 2, 2]]$ code, when using a fault-tolerant technique well suited to the layout of the chip. By showing that fault-tolerant quantum computation is already within our reach, the author hopes to encourage this approach.

Keywords: quantum, fault-tolerance, error detection, experiment

Communicated by: to be filled by the Editorial

1 Introduction

Quantum systems in laboratories around the world are reaching unprecedented level of control and precision for a variety of devices aiming at implementing reliable qubits. Yet, due to the very nature of those devices, errors are still notably present. In order to be able to execute long quantum algorithms, improvements should be made using quantum error correction and fault-tolerant schemes. So-called threshold theorems [1, 2] prove that there exist error rates below which this approach is guaranteed to improve the device performance. However, between those theorems and real experiments there remains a fog of technical details which clouds the actual practicality of error correcting codes.

Already some experiments which demonstrate the usefulness of quantum error correcting codes to protect a quantum memory have been done, e.g. [3–7], but a demonstration of protected computation is still missing.

Inspired by a recent proposal by D. Gottesman [8], we use an IBM 5Q chip to show that error detection can improve some simple sampling tasks, thus turning a tiny portion of fog into blue sky. Closely related to this work, [6] and [7] present experiments based on the same error detecting code, on trapped-ion qubits and on a similar superconducting chip, respectively. Both those works focus on the preparation of few states and only one of the two encoded qubits is handled fault-tolerantly. In this work an extensive set of fault-tolerant circuits is studied and both logical qubits are protected. Moreover [6] and [7] introduce artificial Pauli errors to study the robustness of their preparation whereas this work takes a higher level

^ac.vuillot@tudelft.nl

approach, treating the experimental set-up as a black box, probing only the intrinsic errors in the system. These considerations make the present work closer to the spirit of [8]. This work is, to the author’s knowledge, the first experimental demonstration of error detection and fault-tolerance improving a quantum computation.

The paper is organized as follows. In section 2 we present the principle of the approach and its specialization to the IBM 5Q chip. In section 3 the experimental results are shown and analyzed.

2 Demonstrating fault-tolerance

The idea behind fault-tolerance is to devise error correcting codes and the corresponding processes of encoding, correcting, computing or measuring with encoded information such that more errors are corrected than introduced [9]. The difficulty in devising such processes is that they are built out of components that are all faulty, so adding some might do more harm than it can help. Devising, proving and simulating fault-tolerant schemes has been pursued for the last 20 years, e.g. [1, 9–15], but improving and finding better schemes is still important and a subject of ongoing research, e.g. [16–20].

Taking one of these schemes and experimentally demonstrating its fault-tolerance, i.e. finding an improvement of performance going from bare to encoded implementation, provides the ultimate validation of the scheme.

2.1 General approach

A description of a general approach to demonstrate fault-tolerant quantum computation is given in [8]; we just briefly recall it here. The idea is to choose a quantum error correcting code \mathcal{C} , admitting fault-tolerant circuits for the preparation of some logical states, denoted as $|s_1\rangle, \dots, |s_m\rangle \in S_{\text{FT}}^{\mathcal{C}}$, as well as for some logical quantum gates, denoted as $U_1, \dots, U_n \in G_{\text{FT}}^{\mathcal{C}}$. Using these as building blocks one can then randomly draw a state preparation from $S_{\text{FT}}^{\mathcal{C}}$, then draw a sequence of gates from $G_{\text{FT}}^{\mathcal{C}}$ to obtain an encoded fault-tolerant circuit.

More precisely, following the formalism of *rectangles* and *extended rectangle* introduced in [10], one interleaves the logical units (gates in $G_{\text{FT}}^{\mathcal{C}}$, or preparation of states in $S_{\text{FT}}^{\mathcal{C}}$) with a fault-tolerant circuit for error correction. *Rectangles* designate circuits comprising one logical unit preceded by a round of error correction, *extended rectangles* designate circuits comprising one logical unit preceded and followed by a round of error correction. The precise conditions to satisfy in order to be fault-tolerant for distance 3 codes are stated in [10]. This ensures that the whole computation is error free if there is at most one fault per extended rectangle.

The error model considered in this paper is that the failure of any component can introduce any Pauli error on the qubits acted on by the component. In our case, we cannot repeat fault-tolerant detection of errors, so our circuits will only tolerate a single fault in total.

The encoded circuits then have to be compared to a bare implementation on the physical qubits. Sampling from the final state produced by the circuit in the computational basis and comparing the probability distribution obtained with the expected one gives a simple comparison metric. Note that this supposes that the circuits are small enough or simple enough so that one can classically simulate sampling from the final state efficiently. A successful demonstration of fault-tolerant quantum computation happens if the encoded circuits show better performance than the bare circuits. To be the most convincing, one needs to use the

best of the physical qubits as well as the most efficient implementation for the bare circuits.

2.2 The IBM 5Q chip and $[[4,2,2]]$

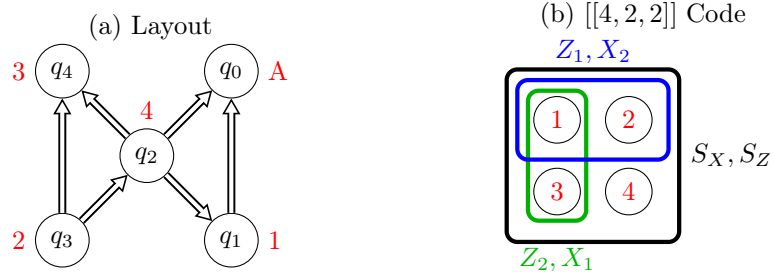


Fig. 1. (a) The connectivity between the qubits in the IBM 5Q Raven chip. Arrows indicate CNOT capabilities where the arrow points towards the target of the CNOT. Red labels indicate our choice of qubit numbering. A is an ancillary qubit used in verification of state preparation. (b) The usual representation layout for the $[[4,2,2]]$ code. The numbered circles represent the qubits and the rectangular boxes represent the support of different Pauli operators: logical Pauli operators Z_1, X_2 are highlighted in blue, logical Pauli operators Z_2, X_1 in green and stabilizers S_X, S_Z in black.

In 2016 IBM released a quantum chip with fixed-frequency superconducting transmon qubits, named IBM 5Q. They provide worldwide cloud access to the chip under an initiative called the ‘‘IBM Quantum Experience’’ [21]. The current iteration on which the experiments were done is nicknamed *Raven*. The appendix also presents and compares preliminary results obtained with a previous iteration, *Sparrow*. The chip has five qubits, natively named q_0 to q_4 . It features single qubit Clifford gates, the T gate ($\text{diag}(1, e^{i\pi/4})$) as well as some two-qubits CNOTs with a certain layout represented in Fig. 1a. This many qubits is the right number to use for a demonstration using the $[[4,2,2]]$ code as discussed in [8].

The $[[4,2,2]]$ code encodes two qubits into four physical qubits. Its code space is stabilized by the all- X and all- Z Pauli operators ($S_X = X \otimes X \otimes X \otimes X$, $S_Z = Z \otimes Z \otimes Z \otimes Z$), together with the logical Pauli operators, they are represented in Fig. 1b. The logical code states are

$$|00\rangle_L = \frac{1}{\sqrt{2}} (|0000\rangle + |1111\rangle) \quad (1)$$

$$|01\rangle_L = \frac{1}{\sqrt{2}} (|1100\rangle + |0011\rangle) \quad (2)$$

$$|10\rangle_L = \frac{1}{\sqrt{2}} (|1010\rangle + |0101\rangle) \quad (3)$$

$$|11\rangle_L = \frac{1}{\sqrt{2}} (|1001\rangle + |0110\rangle). \quad (4)$$

The code also admits transversal implementations of two Clifford gates, namely $H \otimes H \cdot \text{SWAP}$, where H is the Hadamard gate, see Fig. 2a, and the controlled phase or CZ gate, see Fig. 2b. Moreover, if the five qubits have the right connectivity then there are fault-tolerant circuits preparing the logical states $|00\rangle_L$, $|0+\rangle_L$ and $|00\rangle_L + |11\rangle_L$. For the cat state $|00\rangle_L$, see Equation (1), there exists fault-tolerant preparations including one ancillary qubit to verify the preparation. One post-selects on a successful preparation. The logical states $|0+\rangle_L$ and $|00\rangle_L + |11\rangle_L$ are both two Bell pairs but with different pairings. They both also have fault-tolerant preparation circuits. In [8], some circuits are proposed, they can be adapted to the layout of the chip. We tested another preparation for $|00\rangle_L$ inspired by [16], which is substantially shorter so more relevant to this layout.

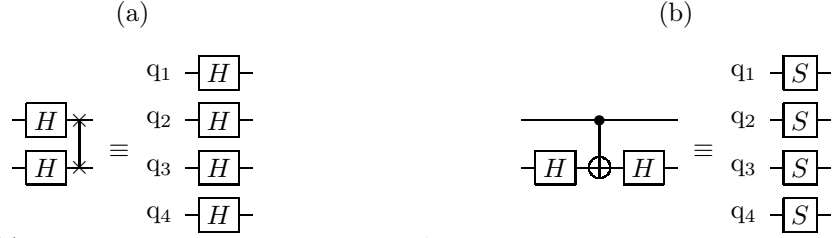


Fig. 2. (a) Bare and encoded versions of $H \otimes H \cdot \text{SWAP}$, where H is the Hadamard gate. In practice the SWAP gate is not done physically on the bare version but just in software, by renumbering the two qubits. (b) Bare and encoded versions of the CZ gate, where $S = \text{diag}(1, i)$.

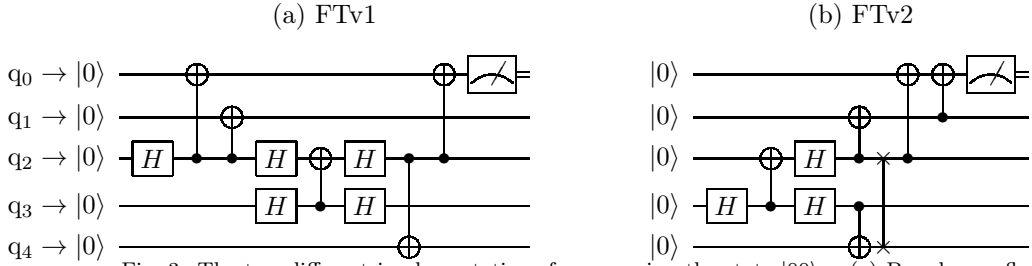


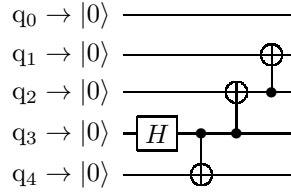
Fig. 3. The two different implementations for preparing the state $|00\rangle_L$: (a) Based on a flagging technique with 5 CNOTs, (b) based on the circular connectivity circuit with 8 CNOTs. In both circuits if the ancillary qubit q_0 is measured as 1, we reject the preparation.

The flagging technique in [16] for the preparation of $|00\rangle_L$ can be implemented with the given CNOT connectivity, see Fig. 3a. We call this the fault-tolerant version one (FTv1). The technique based on a circular layout proposed by [8] cannot be implemented directly due to the connectivity. A clever introduction of one SWAP gate permits to implement it still fault-tolerantly at the cost of 8 CNOT gates in total, see Fig. 3b. We call this the fault-tolerant version two (FTv2). We present below, the results for FTv1, which is more appropriate for the layout. For both those circuits one can check that every possible single fault leads to a detectable error, or an error that stabilizes the prepared state. We also tried a non-fault-tolerant one that has the advantage of being short, involving only 3 CNOTs, see Figure 4. We call this the non-fault-tolerant version (NFT). For the preparation of $|0+\rangle_L$, we want to create a Bell pair between q_1 and q_3 and between q_2 and q_4 . There is a missing connection between q_1 and q_3 , but we can do a similar SWAP trick, see the resulting circuit in Fig. 6a. This circuit is not fault-tolerant as an undetected logical $Z_1 \otimes Z_2 = Z \otimes \mathbf{1} \otimes \mathbf{1} \otimes Z$ can occur if the SWAP gate fails. This is the only possible harmful logical error for this circuit. We'll see below that we only use this circuit in cases were the possible undetected $Z_1 \otimes Z_2$ error doesn't change the outcome of the sampling test. The preparation for $|00\rangle_L + |11\rangle_L$ can itself be straightforwardly implemented fault-tolerantly, see Fig. 6b. To summarize, our sets of initial states and gates are

$$S_{\text{FT}} = \left\{ |00\rangle, |0+\rangle, \frac{|00\rangle + |11\rangle}{\sqrt{2}} \right\},$$

$$G_{\text{FT}} = \{X_1, X_2, Z_1, Z_2, H \otimes H \cdot \text{SWAP}, \text{CZ}\}.$$

The $[[4, 2, 2]]$ code is only an error detecting code, that is, it can detect one Pauli error


 Fig. 4. A short but non-fault-tolerant circuit to prepare $|00\rangle_L$.

but cannot correct it. This means that in place of error correction we have to rely on post-selection to remove errors. In other words, we throw away runs where we detect that an error occurred, either from the ancilla measurement checking state preparation, or from the final measurement which indicates, when the parity of the outcomes is odd, that S_Z has value -1 .

As mentioned before, we cannot interleave rounds of error detection between state preparation and the gates. That means that the final circuit can only be tolerant to a single fault during the whole computation, except for the specific undetectable $Z_1 \otimes Z_2$ failure mentioned above and when we use the non-fault-tolerant version to prepare $|00\rangle_L$.

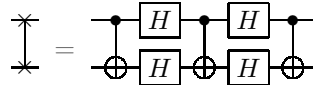
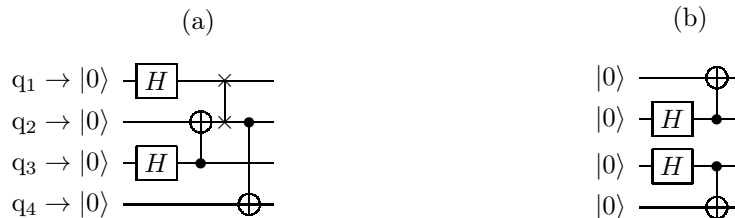


Fig. 5. Implementation of the SWAP gate.


 Fig. 6. (a) Preparation circuits of the logical state $|0+\rangle_L$, with 5 CNOTs. (b) Preparation circuits of the logical Bell state $(|00\rangle_L + |11\rangle_L) / \sqrt{2}$ with 2 CNOTs.

2.3 Comments on the tested circuits

Since we work with such a small system, we can exhaustively find and try all the logically equivalent circuits and optimize the bare version for each one. Essentially, we are making it most likely for the bare version of the task to prevail. From the set of states S_{FT} , and the set of gates G_{FT} , one can obtain 20 different stabilizer states. All the states with their most efficient bare preparation circuit, using the native set of gates provided by IBM, are listed in Figure 7. A brute force approach was used to find them.

Note that adding more gates in the sequence would test states that are already tested with shorter circuits. Therefore those would certainly give worse results as we cannot interleave error detections between gates. Hence we kept only these 20 different preparations. This still extensively probes the computational capabilities of the $[[4, 2, 2]]$ code.

Initial state	Unitary	Final state	# Instructions
$ 00\rangle$	$\mathbb{1} \otimes \mathbb{1}$	$ 00\rangle$	5
$ 0+\rangle$	$\mathbb{1} \otimes \mathbb{1}$	$\frac{ 00\rangle+ 01\rangle}{\sqrt{2}}$	6
$ 00\rangle$	$\mathbb{1} \otimes X$	$ 01\rangle$	6
$ 00\rangle$	$X \otimes \mathbb{1}$	$ 10\rangle$	6
$\frac{ 00\rangle+ 11\rangle}{\sqrt{2}}$	$\mathbb{1} \otimes \mathbb{1}$	$\frac{ 00\rangle+ 11\rangle}{\sqrt{2}}$	7
$ 0+\rangle$	$\mathbb{1} \otimes Z$	$\frac{ 00\rangle- 01\rangle}{\sqrt{2}}$	7
$ 00\rangle$	$H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle+ 01\rangle+ 10\rangle+ 11\rangle}{2}$	7
$ 0+\rangle$	$X \otimes \mathbb{1}$	$\frac{ 10\rangle+ 11\rangle}{\sqrt{2}}$	7
$ 00\rangle$	$X \otimes X$	$ 11\rangle$	7
$\frac{ 00\rangle+ 11\rangle}{\sqrt{2}}$	$\mathbb{1} \otimes Z$	$\frac{ 00\rangle- 11\rangle}{\sqrt{2}}$	8
$\frac{ 00\rangle+ 11\rangle}{\sqrt{2}}$	$X \otimes \mathbb{1}$	$\frac{ 10\rangle+ 01\rangle}{\sqrt{2}}$	8
$ 00\rangle$	$\mathbb{1} \otimes Z \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle- 01\rangle+ 10\rangle- 11\rangle}{2}$	8
$ 00\rangle$	$Z \otimes \mathbb{1} \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle+ 01\rangle- 10\rangle- 11\rangle}{2}$	8
$ 0+\rangle$	$X \otimes Z$	$\frac{ 10\rangle- 11\rangle}{\sqrt{2}}$	8
$\frac{ 00\rangle+ 11\rangle}{\sqrt{2}}$	$\mathbb{1} \otimes ZX$	$\frac{ 10\rangle- 01\rangle}{\sqrt{2}}$	9
$ 00\rangle$	$Z \otimes Z \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle- 01\rangle- 10\rangle+ 11\rangle}{2}$	9
$ 00\rangle$	$CZ \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle+ 01\rangle+ 10\rangle- 11\rangle}{2}$	10
$ 00\rangle$	$CZ \cdot \mathbb{1} \otimes Z \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle- 01\rangle+ 10\rangle+ 11\rangle}{2}$	11
$ 00\rangle$	$CZ \cdot Z \otimes \mathbb{1} \cdot H \otimes H \cdot \text{SWAP}$	$\frac{ 00\rangle+ 01\rangle- 10\rangle+ 11\rangle}{2}$	11
$ 00\rangle$	$\mathbb{1} \otimes X \cdot CZ \cdot H \otimes H \cdot \text{SWAP} \cdot X \otimes \mathbb{1}$	$\frac{ 00\rangle- 01\rangle- 10\rangle- 11\rangle}{2}$	12

Fig. 7. The list of initial states and unitary circuits and the number of QASM instructions in the bare circuit (including final measurements). The final states are written exclusively in the computational basis since that is how they are measured.

3 Experimental results

The code and data can be found on GitHub [22]: it uses the Python SDK that can be found at [23].

3.1 Parameters and runs

Using the IBM chip one can run circuits in batches. For each individual run, 8192 shots are done right one after another in a short time, where each shot outputs 5 bits as measurement outcomes. We consider that the chip is exactly in the same conditions during each run. We consider all the runs to be independent of one another but each to be done in different conditions, so sampling from different output distributions.

For each run, IBM also provides calibration data describing the state of the chip such as: gate error rates, readout error rates, T_1 , T_2 and fridge temperature. We give the average values that we observed for our runs in the appendix.

3.2 Performance metric

For each circuit run we want to compare the observed outcome distribution with the ideal one. The ideal distribution is 8192 independent samples from a distribution with four possible outcomes occurring with probabilities p_{00} , p_{01} , p_{10} and p_{11} . The values for p_{ij} can be read from Figure 7. Since we assume that the conditions stay identical during one run and that the 8192 shots are independent, we observe independent samples from a four-outcome distribution with some different probabilities \tilde{p}_{00} , \tilde{p}_{01} , \tilde{p}_{10} and \tilde{p}_{11} .

We then use the statistical distance as a metric:

$$D = \frac{1}{2} \left(|p_{00} - \tilde{p}_{00}| + |p_{01} - \tilde{p}_{01}| + |p_{10} - \tilde{p}_{10}| + |p_{11} - \tilde{p}_{11}| \right).$$

This quantity is estimated for each run by

$$\hat{D} = \frac{1}{2} \left(\left| p_{00} - \frac{n_{00}}{n_{\text{valid}}} \right| + \left| p_{01} - \frac{n_{01}}{n_{\text{valid}}} \right| + \left| p_{10} - \frac{n_{10}}{n_{\text{valid}}} \right| + \left| p_{11} - \frac{n_{11}}{n_{\text{valid}}} \right| \right),$$

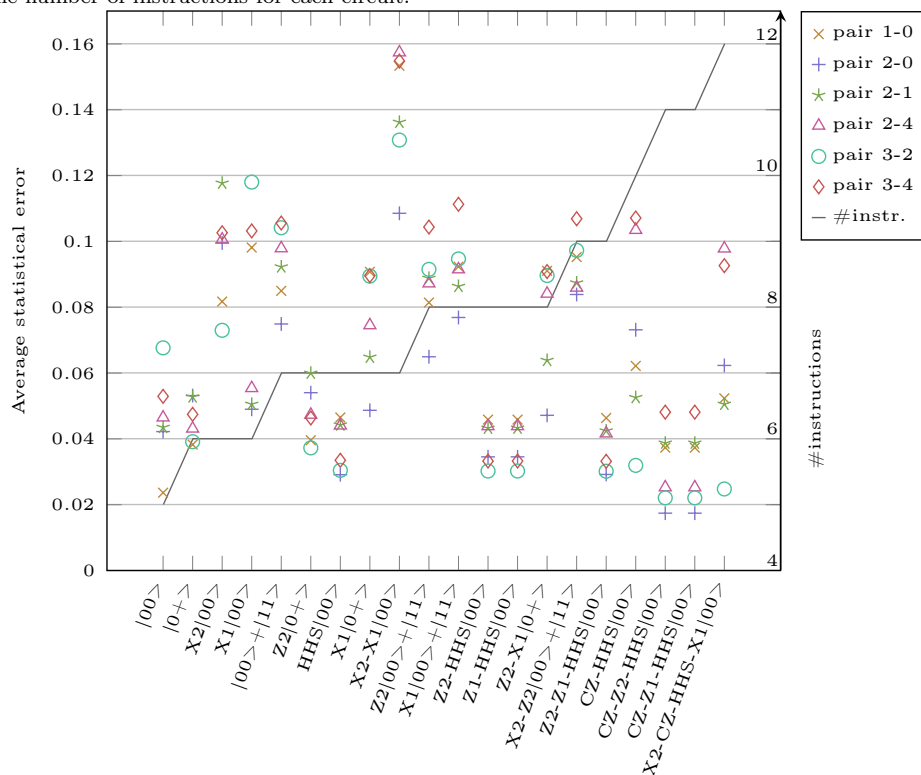
where n_{ij} is the number of observation of outcome ij after post-selection and n_{valid} is the number of shots kept after post-selection. This estimator for D is slightly biased except for the case where only one of the p_{ij} is non-zero (because in this case it becomes linear). We use this estimator to keep the analysis simple.

Each of the runs has some different \tilde{p}_{ij} and we have no information about how the \tilde{p}_{ij} s vary. Therefore we will assume that there are fluctuations around the mean of \hat{D} following some unknown normal distribution and use this model to compute confidence intervals [24]. This means that the final data points and their confidence intervals don't exactly reflect knowledge of D but only of \hat{D} which we believe is still a valid quantity to characterize the performance of the circuits.

3.3 Comparisons

We first need to decide on what pair of bare qubits is the best to be compared to the encoded qubits. It is not immediately clear how to choose this given only the calibration numbers. Thus we tried out the six different connected pairs and we found the performance shown in Figure 8. There is no clear "best pair" but one can see that the pairs $[2-0]$ or $[3-2]$ seem to

Fig. 8. Comparison of the different pairs of qubits to implement the bare circuits together with the number of instructions for each circuit.

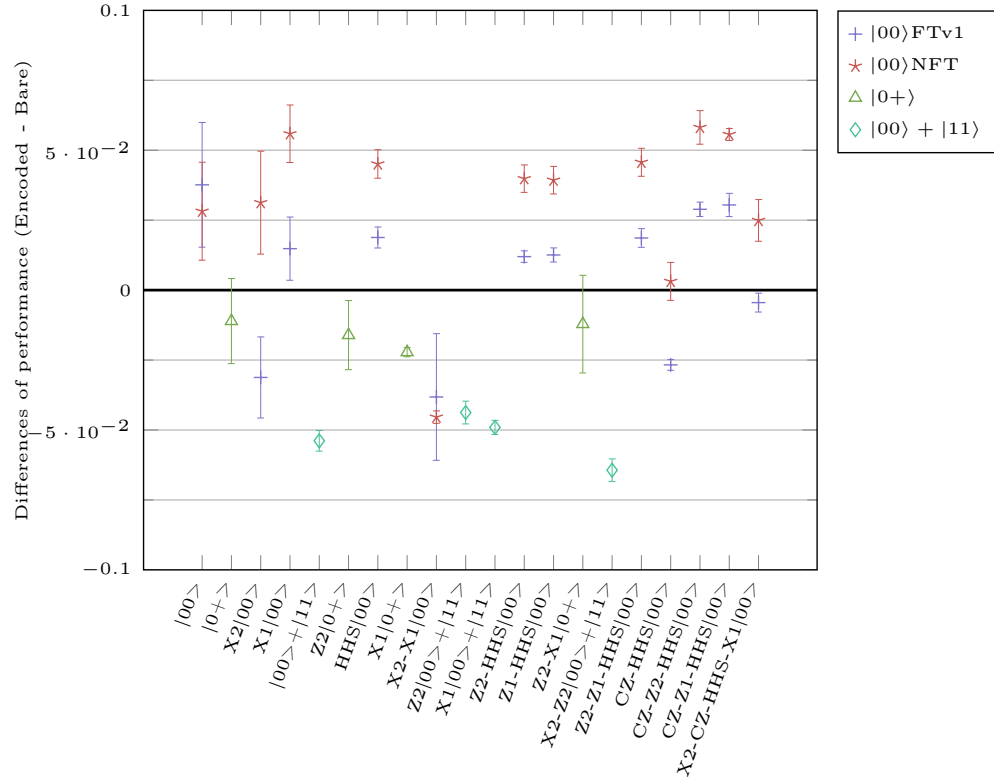


be slightly better. Since we don't have a systematic method for predicting the best pair given the calibration data and the circuit, we will look at average performance for each pair over all the circuits. When averaging, we find the pair $[2 - 0]$ to be the best, see Table 1. The second observation that we can make based on Figure 8 is that the number of instructions does not explain the performance. It seems that the type of state sampled from is more important. Roughly it is easier to sample from equal superposition of the four computational basis states, than from equal superposition of two, than from just one. This can be understood with the fact that with more states in equal superposition there are less Pauli errors or readout errors that can affect the outcome distribution.

We then go on, comparing the encoded versions of the circuits to the elected best pair in Figure 9. The encoded circuits using the preparations $|0+\rangle_L$ and $|00\rangle_L + |11\rangle_L$ are short and fault-tolerant and indeed show better performance than the bare ones. For the different preparations for $|00\rangle_L$, the fault-tolerant version FTv1 is, except in a few cases, better than the non-fault-tolerant one despite being substantially longer. This shows that fault-tolerant design of circuits can be useful. Although they both compare unfavourably to the bare version, except for 4 circuits for FTv1 and only 1 circuit for NFT.

We cannot make an absolute statement as for some circuits the bare version is always better. We also don't want to cherry pick the best version for each circuit since we don't have a systematic way of predicting the best version. Average performance when fixing a

Fig. 9. Comparing encoded performance with bare qubit pair [2–0]. The performance is defined as the statistical distance to the ideal outcome distribution. The figure shows the difference between encoded and bare performance. The error bars show confidence intervals at 99%.



preparation for $|00\rangle_L$ are shown in Table 1. One can see that FTv1, with 4.51×10^{-2} , is better than the best bare implementation with 5.5×10^{-2} , whereas NFT (6.05×10^{-2}) is worse.

4 Conclusion

In conclusion, we have shown that already on the IBM 5Q chip one can improve some quantum computation task, namely sampling from a class of states, by using error detection and fault-tolerant design of circuits. This improvement is only on average over 20 different states that the $[[4, 2, 2]]$ code can fault-tolerantly prepare. We also can see that shorter but non fault-tolerant circuits can be bested by longer but fault-tolerant circuits, showing the usefulness of these. As better and better hardware is developed, with more physical qubits, more connectivity and smaller error rates, demonstrations of fault-tolerance will become easier to produce and become more convincing. The set of gates shown to be fault-tolerant in this paper is very restricted. For the $[[4, 2, 2]]$ code a few more qubits and connections would be needed to realize fully fault-tolerant circuits with error detection in between logical units. Being able to demonstrate the fault-tolerance of larger gate sets, for example the whole Clifford group for several logical qubits would be an important milestone towards harnessing universal

Implementation	Avg.Perf.($\times 10^{-2}$)	Post-selection ratio
Bare[1-0]	6.72	1.00
Bare[2-0]	5.50	1.00
Bare[2-1]	6.49	1.00
Bare[2-4]	6.98	1.00
Bare[3-2]	6.27	1.00
Bare[3-4]	7.73	1.00
FTv1	4.51	0.65
NFT	6.05	0.71

Table 1. Average performance for the different bare and encoded possible implementations with the post-selection ratio (ratio of data kept). The encoded implementations only differ in how they prepare the state $|00\rangle_L$.

quantum computation.

Acknowledgements

I would like to thank the IBM Quantum Experience team for providing extensive access to the chip and support on how to interface with it. I also would like to thank Barbara Terhal for valuable comments and feedback. CV acknowledge support through the EU via the ERC GRANT EQEC and support by the Excellence Initiative of DFG. The views expressed are those of the author and do not reflect the official policy or position of IBM or the IBM Quantum Experience team.

References

- [1] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. In *Proceedings of the Twenty-ninth Annual ACM Symposium on Theory of Computing, STOC '97*, pages 176–188, New York, NY, USA, 1997. ACM. ISBN 0-89791-888-6.
- [2] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation: error models and thresholds. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 454(1969):365–384, 1998. ISSN 1364-5021.
- [3] E. Knill, R. Laflamme, R. Martinez, and C. Negrevergne. Benchmarking quantum computers: The five-qubit error correcting code. *Phys. Rev. Lett.*, 86:5811–5814, Jun 2001.
- [4] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt. Quantum computations on a topologically encoded qubit. *Science*, 345(6194):302–305, 2014. ISSN 0036-8075.
- [5] A. D. Córcoles, Easwar Magesan, Srikanth J. Srinivasan, Andrew W. Cross, M. Steffen, Jay M. Gambetta, and Jerry M. Chow. Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. *Nature Communications*, 6(6979), April 2015.

- [6] Norbert M. Linke, Mauricio Gutierrez, Kevin A. Landsman, Caroline Figgatt, Shantanu Debnath, Kenneth R. Brown, and Christopher Monroe. Fault-tolerant quantum error detection. *Science Advances*, 3(10), 2017, <http://advances.sciencemag.org/content/3/10/e1701074.full.pdf>.
- [7] Maika Takita, Andrew W. Cross, A. D. Córcoles, Jerry M. Chow, and Jay M. Gambetta. Experimental demonstration of fault-tolerant state preparation with superconducting qubits. *Phys. Rev. Lett.*, 119:180501, Oct 2017.
- [8] Daniel Gottesman. Quantum fault tolerance in small experiments. *ArXiv e-prints*, October 2016, 1610.03507. URL <https://arxiv.org/abs/1610.03507>.
- [9] P. W. Shor. Fault-tolerant quantum computation. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, FOCS '96, pages 56–, Washington, DC, USA, 1996. IEEE Computer Society.
- [10] Panos Aliferis, Daniel Gottesman, and John Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Info. Comput.*, 6(2):97–165, March 2006. ISSN 1533-7146.
- [11] Panos Aliferis, Daniel Gottesman, and John Preskill. Accuracy threshold for postselected quantum computation. *Quantum Info. Comput.*, 8(3):181–244, March 2008. ISSN 1533-7146.
- [12] Andrew W. Cross, David P. DiVincenzo, and Barbara M. Terhal. A comparative code study for quantum fault tolerance. *Quantum Info. Comput.*, 9(7):541–572, July 2009. ISSN 1533-7146.
- [13] Adam Paetznick and Ben W. Reichardt. Fault-tolerant ancilla preparation and noise threshold lower bounds for the 23-qubit golay code. *Quantum Information & Computation*, 12:1034–1080, 2012, 1106.2190v2.
- [14] Yu Tomita and Krysta M. Svore. Low-distance surface codes under realistic quantum noise. *Phys. Rev. A*, 90:062320, Dec 2014.
- [15] Christopher Chamberland, Tomas Jochym-O’Connor, and Raymond Laflamme. Thresholds for universal concatenated quantum codes. *Phys. Rev. Lett.*, 117:010501, June 2016.
- [16] Rui Chao and Ben W. Reichardt. Quantum error correction with only two extra qubits. *ArXiv e-prints*, May 2017, 1705.02329v1. URL <https://arxiv.org/abs/1705.02329>.
- [17] Rui Chao and Ben W. Reichardt. Fault-tolerant quantum computation with few qubits. *ArXiv e-prints*, May 2017, 1705.05365. URL <https://arxiv.org/abs/1705.05365>.
- [18] Ryuji Takagi, Theodore J. Yoder, and Isaac L. Chuang. Error rates and resource overheads of encoded three-qubit gates. *Phys. Rev. A*, 96:042302, Oct 2017.
- [19] Christopher Chamberland, Tomas Jochym-O’Connor, and Raymond Laflamme. Overhead analysis of universal concatenated quantum codes. *Phys. Rev. A*, 95:022313, Feb 2017.

- [20] Christopher Chamberland and Michael E. Beverland. Flag fault-tolerant error correction with arbitrary distance codes. *ArXiv e-prints*, August 2017, 1708.02246v2. URL <https://arxiv.org/abs/1708.02246>.
- [21] IBM Quantum Experience. URL <https://quantumexperience.ng.bluemix.net>.
- [22] Code and data of the experiments. URL <https://github.com/ChristopheVuillot/>.
- [23] Python sdk. URL <https://github.com/IBM/qiskit-sdk-py>.
- [24] Ronald E. Walpole and Raymond H. Myers. *Probability and Statistics for Engineers and Scientists*, chapter 8.6. Macmillan USA, 1985. ISBN 9780024241702.

Appendix A

This appendix presents the calibration data for the experiments, averaged over all the different runs as well as the observed standard deviation. It also shows previous runs on a previous iteration of the chip, called Sparrow, which was showing better performance but has been taken down by IBM. The current iteration is named Raven: the more extensive runs presented in this paper were done on this chip. Note that for the runs on Sparrow, there were only 4 different calibrations and 36 runs for each circuit. For Raven we have more than 100 runs for each circuit and almost a new calibration per run.

chip	Temperature (mK)
Raven	21
Sparrow	19

Table A.1. Average fridge temperatures for the two chips at the time of the runs in mK.

(a) Raven					(b) Sparrow				
qubit	T_1 (μ s)	$\sigma[T_1]$	T_2 (μ s)	$\sigma[T_2]$	qubit	T_1 (μ s)	$\sigma[T_1]$	T_2 (μ s)	$\sigma[T_2]$
Q0	46	6.5	40	7.3	Q0	49	8.9	39	4.5
Q1	57	6.8	45	6.3	Q1	46	11.0	39	13.3
Q2	39	5.1	40	3.2	Q2	56	1.7	91	9.9
Q3	40	4.8	57	12.5	Q3	47	4.9	62	12.0
Q4	54	9.1	21	1.3	Q4	55	6.0	81	13.2

Table A.2. Average T_1 and T_2 parameters for the runs together with their standard deviation all in μ s.

The two chips are similar, the gate and readout error rates have been slightly improved with Raven but the fridge temperature as well as T_1 and T_2 times became a bit worse. The layout differs only in the orientation of the CNOTs, see Figure A.1. The circuits run on Sparrow are showed in Figure A.2 and Figure A.3. Comparison of the bare and encoded performance between Raven and Sparrow are presented in Figure A.4 and Figure A.5. One can see that the previous iteration was performing better.

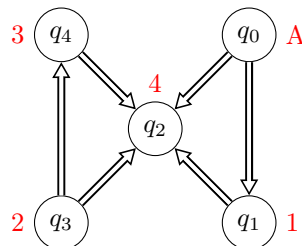


Fig. A.1. Layout of the Sparrow chip.

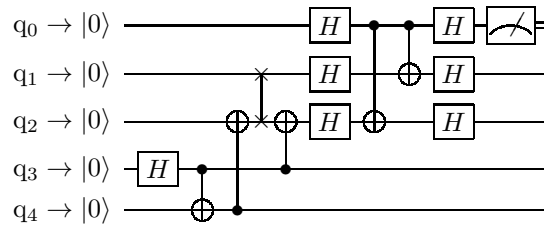
(a) Raven				
qubit	gate error (%)	σ [gate error]	readout error (%)	σ [readout error]
Q0	0.1	0.021	4.1	0.63
Q1	0.08	0.035	4.7	0.96
Q2	0.15	0.024	2.7	0.36
Q3	0.15	0.016	6.2	1.11
Q4	0.19	0.026	5.1	1.1

(b) Sparrow				
qubit	gate error (%)	σ [gate error]	readout error (%)	σ [readout error]
Q0	0.23	0.009	1.9	0.22
Q1	0.30	0.099	8.3	2.19
Q2	0.42	0.085	1.5	0.29
Q3	0.52	0.213	12.3	3.60
Q4	0.33	0.027	6.6	0.83

Table A.3. Average single-qubit gate and readout error rates for the runs together with their standard deviation, all in percent.

(a) Raven			(b) Sparrow		
pair	gate error (%)	σ [gate error]	pair	CNOT error (%)	σ [CNOT error]
1-0	2.2	0.32	1-0	4.0	0.74
2-0	2.3	0.25	2-0	3.0	0.24
2-1	2.7	0.36	2-1	4.6	0.86
2-4	3.9	0.78	2-4	3.7	0.34
3-2	2.2	0.30	3-2	6.2	0.90
3-4	2.7	0.38	3-4	5.8	2.05

Table A.4. Average CNOT gate error rates and their standard deviation for the runs, all in percent.

Fig. A.2. The circuit previously implemented on the IBM 5Q Sparrow chip preparing the logical state $|00\rangle_L$. The SWAP gate is implemented via the circuit in Figure 5. If the SWAP gate fails, it can introduce Pauli X errors on q_1 and q_2 , which would not be detected and which constitute a logical $X_1 \otimes X_2$ error.

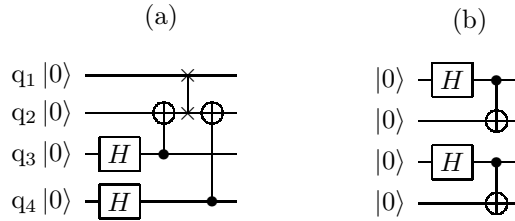


Fig. A.3. (a) Preparation circuits of the logical state $|0+\rangle_L$, with 5 CNOTs. If the SWAP gate fails, it can introduce Pauli X errors on q_1 and q_2 , which would not be detected and which constitute a logical $X_1 \otimes X_2$ error. (b) Preparation circuits of the logical Bell state $(|00\rangle_L + |11\rangle_L) / \sqrt{2}$ with 2 CNOTs.

Fig. A.4. Comparison of the performances of the bare circuits between the Raven and Sparrow chips. Except for the four circuits involving a two-qubit gate, Sparrow was showing better performance. This is probably due to improvement of the two-qubit gates but shorter T_1 and T_2 times on Raven, see Table A.2 and Table A.4.

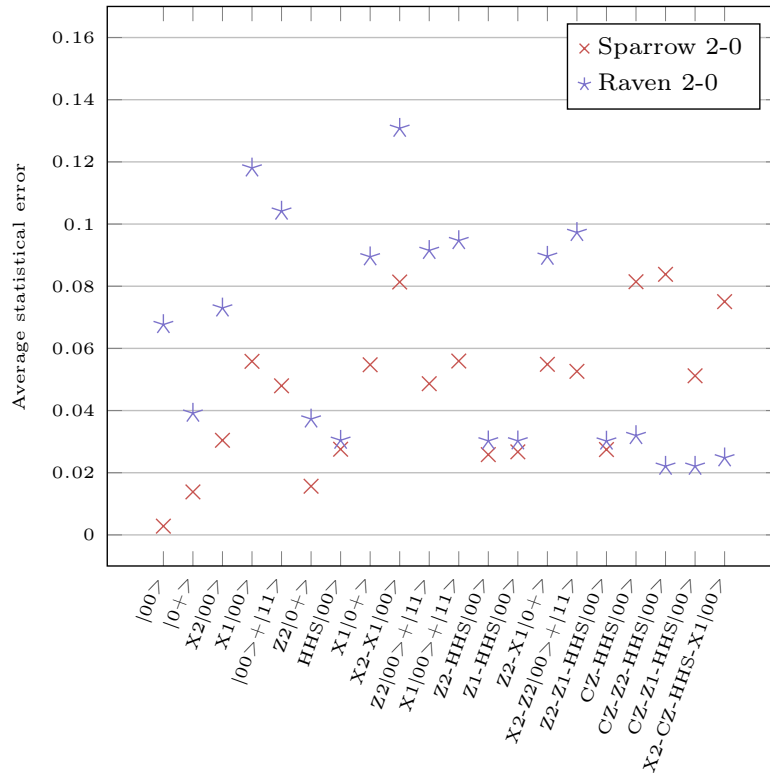


Fig. A.5. Comparing encoded versions on Raven and Sparrow. The $|00\rangle_L$ and $|0+\rangle_L$ preparations were suffering from possible X_1X_2 logical errors.

