

**Low-overhead constructions for the fault-tolerant Toffoli gate**

Cody Jones\*

*Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305-4088, USA*

(Received 9 January 2013; published 20 February 2013)

We present two constructions for the Toffoli gate which substantially reduce resource costs in fault-tolerant quantum computing. The first contribution is a Toffoli gate requiring Clifford operations plus only four  $T = \exp(i\pi\sigma^z/8)$  gates, whereas conventional circuits require seven  $T$  gates. An extension of this result is that adding  $n$  control inputs to a controlled gate requires  $4n$   $T$  gates, whereas the best prior result was  $8n$ . The second contribution is a quantum circuit for the Toffoli gate which can detect a single  $\sigma^z$  error occurring with probability  $p$  in any one of eight  $T$  gates required to produce the Toffoli gate. By postselecting circuits that did not detect an error, the posterior error probability is suppressed to lowest order from  $4p$  (or  $7p$ , without the first contribution) to  $28p^2$  for this enhanced construction. In fault-tolerant quantum computing, this construction can reduce the overhead for producing logical Toffoli gates by an order of magnitude.

DOI: [10.1103/PhysRevA.87.022328](https://doi.org/10.1103/PhysRevA.87.022328)

PACS number(s): 03.67.Pp, 03.67.Lx

**I. INTRODUCTION**

Fault-tolerant quantum computing is the effort to design quantum-information processors which are resilient to sufficiently small (but nonzero) probability of failure in any individual component [1,2]. Enhanced reliability comes at the cost of redundancy, and recent study in this area has focused on minimizing the overhead, or additional resource costs, associated with converting a perfect quantum operation into a form compatible with error correction [3–5]. This work focuses on the Toffoli gate, which appears in both reversible-classical and quantum logic and which may be defined as  $U_{\text{Toffoli}}|x,y,z\rangle = |x,y,z \oplus xy\rangle$ , with  $x$ ,  $y$ , and  $z$  being binary variables. Unlike many quantum gates, the quantum Toffoli gate has a classical analog, so it is favored as a building block for importing more complex classical operations, such as binary arithmetic, into quantum algorithms like Shor’s factoring algorithm [4,6,7] and quantum simulation [4,8,9]. For these reasons, the Toffoli gate is critically important to quantum computing in general, and improvements in the design of the Toffoli gate make the realization of large-scale quantum computation more tractable.

Several researchers have studied circuit constructions for the Toffoli gate. The most often cited implementation is probably the one on page 182 of Ref. [2], which may have been derived from Ref. [10]. As can be seen in Ref. [2], the Toffoli gate is decomposed into smaller quantum gates, each of which can be made fault tolerant by conventional means [1]. The most nettlesome of these is the  $T = \exp(i\pi\sigma^z/8)$  gate, which is much more expensive in both time and space resources to produce [3,4,11–16]; notably, the Toffoli circuit in Ref. [2] uses seven  $T$  gates. In fact, Ref. [10] contains a construction nearly identical to one derived here (we use four  $T$  gates), except for an undesirable ( $-1$ ) phase on one output state (we show how to correct this with modest effort). However, “complete” implementations of the Toffoli gate without a phase error have used seven  $T$  gates in the literature to date. Amy *et al.* [17] studied classical search methods for decomposing gates like the Toffoli gate into fault-tolerant

primitives, and Selinger [18] investigated circuit constructions with particular emphasis on  $T$ -gate count and depth, where the latter metric allows parallel  $T$  gates on different qubits. We use Selinger’s work as our starting point, as we turn his almost-Toffoli gate into a proper Toffoli gate, using four  $T$  gates and some quantum teleportation. Finally, the importance of this topic has attracted the attention of other researchers, and Eastin [19] has independently discovered equivalent results.

This paper presents two important results. First, Sec. II describes how to implement the Toffoli gate with only four  $T$  gates and Clifford-group operations [2,20]. Second, Sec. III introduces a Toffoli construction requiring eight  $T$  gates that can detect an error in any single  $T$  gate. This circuit is an important development for fault-tolerant quantum computing, because it relaxes the requirements on high-fidelity  $T$  gates that are expensive to produce; however, the circuit is probabilistic, and we discuss its proper usage. Section IV presents some analysis of the resource costs and error rates for these circuits. The paper concludes with a brief discussion of the impact these results have on large-scale quantum computing.

**II. TOFFOLI GATE USING JUST FOUR  $T$  GATES**

In most fault-tolerant quantum computing architectures, the most difficult quantum gates to produce are non-Clifford gates [12]. The Hadamard gate  $H = (1/\sqrt{2})(\sigma^x + \sigma^z)$ , the phase gate  $S = \exp(i\pi\sigma^z/4)$ , and the CNOT gate are generators for the Clifford group, as any gate in this group can be produced by combinations thereof, up to a global phase that we ignore. However, at least one gate outside the Clifford group is required for universal quantum computing [2]. The  $T$  gate is often selected because it is the easiest to produce; however, as we explain below, “easy” is relative, and this gate is still quite expensive in computing resources.

The choice of  $T$  gate as the non-Clifford operation that enables universal quantum computing is motivated by the set structure  $\{C_n\}$  studied by Gottesman and Chuang [20].  $C_1$  is the Pauli group,  $C_2$  is Clifford group, and  $C_n \equiv \{U|UC_1U^\dagger \subseteq C_{n-1}\}$ . For example, this means that a frame transformation by some gate  $U$  in  $C_3$  maps Pauli operators to the Clifford group. As shown in Ref. [20], elements of  $C_3$  can be implemented

\*ncodyjones@gmail.com

using teleportation, with conditional by-product operators that are in the Clifford group. Thus, when Clifford gates are natively available in error correction, elements of  $C_3$  are the most convenient choice of non-Clifford gates to enable universality. The  $T$  gate is in  $C_3$  (as are other rotations by angle  $\pi/4$  on the Bloch sphere), which explains why it is often selected to complete universality [3,4,11–16].

In most quantum codes, including the surface code [21], non-Clifford gates are produced using an ancilla state that is injected into the circuit [2,20]. As this ancilla is produced in a faulty manner, it must be purified through magic-state distillation [12,14]. The handful of rounds of state distillation required to distill input states, where each has error probability  $p$  in the range  $10^{-2}$ – $10^{-4}$ , to reach the roughly  $10^{-12}$  error rates required for quantum algorithms [3,4,22] are considerably expensive, such that a single  $T$  gate requires about  $50\times$  the circuit volume (product of qubits and time steps) of a CNOT gate or an  $H$  gate [4,5], making its production the dominant cost among fault-tolerant gate primitives. This poses an issue for quantum computing, as very many  $T$  gates in the form of Toffoli gates are required for typical quantum algorithms like integer factoring or quantum simulation. The first Toffoli gate construction we present uses four  $T$  gates instead of seven, thereby reducing the overhead due to state distillation.

Let us denote the Toffoli\* gate as the operation in Fig. 1(a), which requires four  $T$  gates and was introduced by Selinger [18]. Toffoli and Toffoli\* gates differ only by a controlled- $S^\dagger$  gate between the control qubits  $x$  and  $y$ . Beginning with the Toffoli\* gate, we need only an ancilla qubit, a phase gate  $S$ , and teleportation to implement the exact Toffoli gate, as shown in Fig. 1(b). We first apply the Toffoli\* gate using the same controls as the desired Toffoli gate but with an ancilla  $|0\rangle$  as the target. The erroneous controlled- $S^\dagger$  is corrected by a simple  $S$  gate applied to the ancilla. Afterwards, the CNOT and measurement teleport the doubly conditional NOT operation encoded in the ancilla to the target qubit of the desired Toffoli gate. The measurement result determines whether a corrective gate of controlled- $\sigma^z$ , which is in the Clifford group, is required to correct a  $(-1)$  phase resulting from measurement back-action.

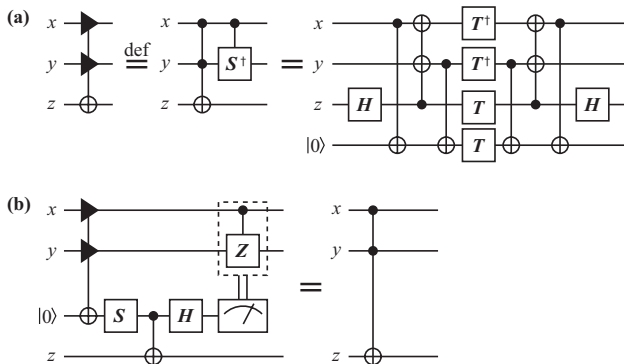


FIG. 1. A circuit construction for a Toffoli gate using four  $T$  gates. (a) The Toffoli\* circuit by Selinger [18] that is almost a Toffoli gate, with the difference being the controlled- $S^\dagger$  operation. (b) Our circuit combines the Toffoli\* circuit with a phase correction and teleportation to produce an exact Toffoli gate. The measurement is in the  $\sigma^z$  basis, and the double vertical lines indicate that the controlled- $\sigma^z$  correction is conditioned on the measurement result being  $|1\rangle$ .

One can readily verify that only four  $T$  gates are required in this procedure [2,20]. Note that the inverse gate  $T^\dagger$  requires the same ancilla-based teleportation circuit as  $T$ , so these gates are equivalent in state-distillation cost and construction.

The construction in Fig. 1(b) can also be used to add control-qubit inputs to an existing controlled- $G$  gate, where  $G$  is any unitary. Replace the CNOT in Fig. 1(b) with controlled- $G$  (targeting however many qubits  $G$  acts on), and the result is controlled-controlled- $G$ . By iterating this procedure, one can add  $n$  controls to controlled- $G$  using  $4n$   $T$  gates. The best prior result required  $8n$   $T$  gates [18].

III. ERROR-DETECTING TOFFOLI CIRCUIT

Whereas the previous section reduced the number of  $T$  gates needed to make a Toffoli circuit, this section addresses the resource-cost problem differently by making each  $T$  gate less expensive. The cost of a  $T$  gate scales inversely with the probability  $p$  of it having an undetected error, with a relationship where the circuit volume (qubits  $\times$  gates) is  $O(\text{poly}[\log(1/p)])$ . We introduce a new Toffoli gate that can detect an error in any one of eight  $T$  gates. As a result, the effective error probability of the Toffoli gate is  $28p^2$  instead of  $4p$  (we only consider the lowest nonvanishing order throughout this paper since  $p \ll 1$ ). Even though twice as many  $T$  gates are needed, they can tolerate larger error rates, so they are substantially less expensive to produce than would otherwise be necessary.

The error-detecting Toffoli circuit is rather simple to derive. It consists of two Toffoli\* gates acting on a target qubit which is in a bit-flip code [2], as shown in Fig. 2. The gate with reversed triangles is the inverse operation (Toffoli\*) $^\dagger$ . Importantly, the controlled- $S$  and controlled- $S^\dagger$  gates acting on the same qubits  $x$  and  $y$  are inverse operations, so they cancel. A logically equivalent decomposition into  $T$  gates is shown in Fig. 3; this circuit is convenient for analyzing how errors propagate. The correspondence is achieved by placing  $T$  gates on ancilla qubits with the aid of CNOT gates [18]. We assume that  $|0\rangle$  preparation,  $H$ , CNOT, and measurement operations are perfect, because fault-tolerant error correction for these processes is economical compared to  $T$  gates. A single  $\sigma^z$  error in any of the  $T$  gates will necessarily propagate to the syndrome measurement for this bit-flip code, as indicated by the red dashed lines. Upon such an event, all of the qubits are discarded. Note that  $\sigma^x$  errors, if present, do not propagate anywhere since they commute with the CNOT gates; they have no effect on the Toffoli gate.

The circuit in Fig. 3 must be discarded upon a detected error event, which happens with probability  $8p$ . If this circuit

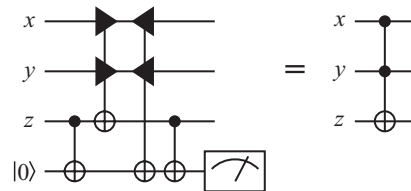


FIG. 2. An error-detecting Toffoli gate. The measurement is in the  $\sigma^z$  basis, and obtaining result  $|1\rangle$  indicates an error was detected, so the qubits should be discarded.

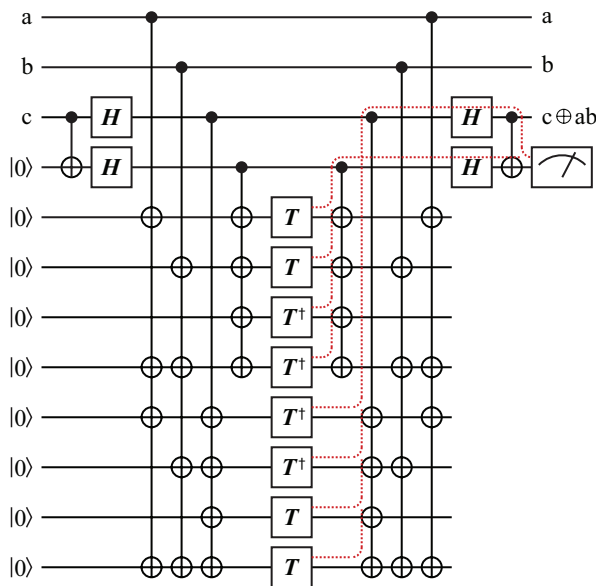


FIG. 3. (Color online) An error-detecting Toffoli gate. The red dashed lines indicate how any single  $\sigma^z$  error will propagate to the readout qubit. The measurement is in the  $\sigma^z$  basis, and obtaining result  $|1\rangle$  indicates an error was detected, so the qubits should be discarded. As long as the ancilla qubits are initialized perfectly to  $|0\rangle$  and the CNOT and  $H$  gates have no errors, then only  $\sigma^z$  errors in the  $T$  gates matter, as  $\sigma^x$  errors cannot propagate to data qubits. If the probability of a  $\sigma^z$  error in each  $T$  gate is an independent, identically distributed (i.i.d.) Bernoulli( $p$ ), then the success probability is  $1 - 8p$  and the *a posteriori* error probability is  $28p^2$ , to lowest order in  $p$ .

is connected by entanglement to other qubits in a quantum algorithm, all qubits must be discarded, and the algorithm fails. To avoid this scenario, one can produce a Toffoli ancilla [2]. If the circuit fails because of a detected error, then the qubits are discarded, but no far-reaching damage occurs since this

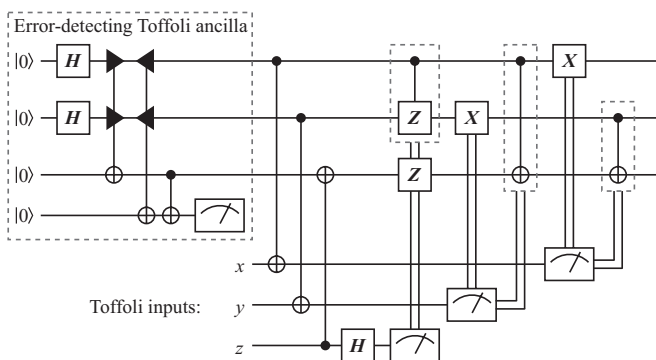


FIG. 4. Proper use of an error-detecting Toffoli ancilla. All measurements are in the  $\sigma^z$  basis. The ancilla-production circuit in the box in the upper left corner is probabilistic. A measurement result of  $|1\rangle$  indicates the circuit failed, in which case the qubits are discarded. Because the Toffoli ancilla is not coupled to any other part of the quantum computation, its production can be repeated until success. The subsequent CNOT gates and measurements teleport data qubits through the Toffoli gate encoded in the ancilla (cf. p. 488 of Ref. [2]). Clifford-group gates are applied conditional on measurements showing outcome  $|1\rangle$ , as indicated by the parallel double lines.

faulty circuit is not entangled to any data qubits. Conditional on the circuit succeeding, the ancilla is teleported into data qubits to enact a Toffoli gate, using only Clifford gates and measurement, as shown in Fig. 4. Using a representative value for the  $T$ -gate error such as  $p = 10^{-8}$  (we consider such a scenario in Sec. IV), the failure probability for preparing the Toffoli ancilla is a modest  $8 \times 10^{-8}$ , which negligibly increases the number of times such preparation circuits must be repeated.

IV. RESOURCE ANALYSIS

Comparing resource costs between the naive Toffoli gate using seven  $T$  gates and our construction using four gates is straightforward. The latter requires about half the resources of the former, under our assumption that  $T$  gates are the dominant cost. There is also a modest improvement in the Toffoli error rate ( $7p$  becomes  $4p$ ). However, in fault-tolerant quantum computing, this result is likely overshadowed by the error-detecting construction.

Doubling the number of  $T$  gates from four to eight to achieve  $O(p^2)$  Toffoli error rate is usually the correct decision. The reason is that this approach is more economical than increasing the accuracy of the  $T$  gates through further magic-state distillation (or other fault-tolerant procedures [2]). Bravyi and Haah [15] present a conjecture in the context of magic-state distillation, stating that to produce one magic state with error  $O(p^2)$  requires at least two input states with error  $p$ ; hence, increasing  $T$ -gate accuracy to  $O(p^2)$  will require at least twice as many circuit resources, and in all practical cases known to this author, the overhead factor is larger than 2 (an example case is considered below). Moreover, there is no known protocol which saturates this bound. Multilevel distillation comes arbitrarily close as  $p \rightarrow 0$ , but this limiting case is not always relevant for finite  $p$ , and multilevel protocols require large and complex circuits [16].

Under conditions relevant to quantum computing, the error-detecting Toffoli gate in Fig. 3 can reach the low error rates required for quantum algorithms with one less round of state distillation, leading to as much as an order-of-magnitude reduction in the resources required to produce a fault-tolerant Toffoli gate. For example, suppose that we wish to produce a Toffoli gate with an error probability below  $10^{-12}$ . We presume the “raw”  $T$ -gate ancilla has a  $\sigma^z$  error probability of  $10^{-2}$ . Using the results in Ref. [14], the simple Toffoli gate would require four  $T$  gates distilled to  $p = 10^{-15}$  using a hybrid scheme of one round of Bravyi-Kitaev (BK) distillation and two rounds of Meier-Eastin-Knill (MEK) distillation, at an average total cost of 1744.8 raw states. Conversely, the error-detecting Toffoli gate would require just one round each of BK and MEK distillation circuits for each of the eight  $T$  gates distilled to  $p = 10^{-8}$ , for a total average cost of 697.6 raw states. The resource savings factor is  $2.5\times$ , just in terms of the number of undistilled states needed for distillation. In practice, the resource savings is amplified by another factor of about  $2\times$  because one less round of distillation is needed (fewer gates means smaller circuit volume). Reference [22] estimates that each additional round of BK distillation increases total circuit volume by a factor of  $3\times$ , although this specific factor may be lower for other protocols like MEK that have a different

“scaling exponent” [15,16]. Additionally, the state-distillation subcircuits for the error-detecting Toffoli gate can use weaker error correction (i.e., lower code distance) than the same preparation circuits for the simple Toffoli, which translates to fewer qubits and gates at the hardware level [22]. Relative to the Toffoli circuit using seven  $T$  gates, there is an additional savings factor of  $7/4$ . Therefore, the error-detecting circuit reduces total overhead for non-Clifford gates by up to an order of magnitude in this representative example.

It is also noteworthy that if “raw”  $T$  gates can be produced with an error rate of  $p = 10^{-4}$ , then the error-detecting Toffoli has a posterior error probability of approximately  $3 \times 10^{-7}$ . This would enable modest quantum computations using about  $10^6$  Toffoli gates, such as the multiplication of two 1000-bit numbers, without the need for resource-intensive magic-state distillation.

## V. CONCLUSIONS

The Toffoli gate is a ubiquitous operation in quantum computing, as it plays a key role in many quantum algorithms.

However, quantum computers that realize these algorithms are still out of reach. In the meantime, engineering a system capable of large-scale, fault-tolerant quantum computation demands that quantum computer architects minimize computing resource costs in terms of execution time and machine size. The constructions in this paper substantially reduce the circuit volume for the fault-tolerant Toffoli gate when one considers how expensive each non-Clifford gate  $T$  is to produce. In the case of the error-detecting Toffoli gate, the resource savings is an order of magnitude in a representative example with a  $T$ -gate error of  $p = 0.01$ . The improved fault-tolerant Toffoli gate brings large-scale quantum computing closer to realization.

## ACKNOWLEDGMENTS

This work was supported by the University of Tokyo Special Coordination Funds for Promoting Science and Technology, NICT, and the Japan Society for the Promotion of Science (JSPS) through its “Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program).”

- 
- [1] John Preskill, *Proc. R. Soc. London A* **454**, 385 (1998).
  - [2] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information*, 1st ed. (Cambridge University Press, Cambridge, UK, 2000).
  - [3] N. Isailovic, M. Whitney, Y. Patel, and J. Kubiawicz, in *Proceedings of the 35th International Symposium on Computer Architecture, 2008 (ISCA'08), 21–25 June 2008, Beijing* (IEEE, New York, 2008), pp. 177–188.
  - [4] N. Cody Jones, Rodney Van Meter, Austin G. Fowler, Peter L. McMahon, Jungsang Kim, Thaddeus D. Ladd, and Yoshihisa Yamamoto, *Phys. Rev. X* **2**, 031007 (2012).
  - [5] Austin G. Fowler and Simon J. Devitt, [arXiv:1209.0510](https://arxiv.org/abs/1209.0510).
  - [6] Peter W Shor, *SIAM J. Comput.* **26**, 1484 (1997).
  - [7] Rodney Van Meter and Kohei M. Itoh, *Phys. Rev. A* **71**, 052320 (2005).
  - [8] C. R. Clark, T. S. Metodi, S. D. Gasster, and K. R. Brown, *Phys. Rev. A* **79**, 062314 (2009).
  - [9] N Cody Jones, James D Whitfield, Peter L McMahon, Man-Hong Yung, Rodney Van Meter, Alán Aspuru-Guzik, and Yoshihisa Yamamoto, *New J. Phys.* **14**, 115023 (2012).
  - [10] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter, *Phys. Rev. A* **52**, 3457 (1995).
  - [11] Xinlan Zhou, Debbie W. Leung, and Isaac L. Chuang, *Phys. Rev. A* **62**, 052316 (2000).
  - [12] Sergey Bravyi and Alexei Kitaev, *Phys. Rev. A* **71**, 022316 (2005).
  - [13] R. Raussendorf, J. Harrington, and K. Goyal, *New J. Phys.* **9**, 199 (2007).
  - [14] Adam M. Meier, Bryan Eastin, and Emanuel Knill, [arXiv:1204.4221](https://arxiv.org/abs/1204.4221).
  - [15] Sergey Bravyi and Jeongwan Haah, *Phys. Rev. A* **86**, 052329 (2012).
  - [16] Cody Jones, [arXiv:1210.3388](https://arxiv.org/abs/1210.3388).
  - [17] Matthew Amy, Dmitri Maslov, Michele Mosca, and Martin Roetteler, [arXiv:1206.0758](https://arxiv.org/abs/1206.0758) [quant-ph].
  - [18] Peter Selinger, [arXiv:1210.0974](https://arxiv.org/abs/1210.0974).
  - [19] Bryan Eastin, [arXiv:1212.4872](https://arxiv.org/abs/1212.4872).
  - [20] Daniel Gottesman and Isaac L. Chuang, *Nature (London)* **402**, 390 (1999).
  - [21] Austin G. Fowler, Ashley M. Stephens, and Peter Groszkowski, *Phys. Rev. A* **80**, 052312 (2009).
  - [22] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, *Phys. Rev. A* **86**, 032324 (2012).