

# Efficient quantum circuit synthesis for SAT-oracle with limited ancillary qubit

Shuai Yang,<sup>1,2</sup> Wei Zi,<sup>1,2</sup> Bujiao Wu,<sup>3</sup> Cheng Guo,<sup>1,2</sup> Jialin Zhang,<sup>1,2</sup> and Xiaoming Sun<sup>1,2</sup>

<sup>1</sup>*Institute of Computing Technology, Chinese Academy of Sciences, 100190 Beijing, China*

<sup>2</sup>*University of Chinese Academy of Sciences, 100049 Beijing, China*

<sup>3</sup>*Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China*

How to implement quantum oracle with limited resources raises concerns these days. We design two ancilla-adjustable and efficient algorithms to synthesize SAT-oracle, the key component in solving SAT problems. The previous work takes  $2m - 1$  ancillary qubits and  $\tilde{O}(m)$  elementary gates to synthesize an  $m$  clauses oracle. The first algorithm reduces the number of ancillary qubits to  $2\sqrt{m}$ , with at most an eightfold increase in circuit size. The number of ancillary qubits can be further reduced to 3 with a quadratic increase in circuit size. The second algorithm aims to reduce the circuit depth. By leveraging of the second algorithm, the circuit depth can be reduced to  $\tilde{O}(\log m)$  with  $m$  ancillary qubits.

## INTRODUCTION

Quantum computation has been extensively studied since Feynman first proposed in the 1980s [1]. Several quantum algorithms have been proposed which are superior to the best classical algorithms, such as Shor's algorithm and Grover's algorithm [2, 3]. As a result, more and more attention is paid to quantum computation [4, 5].

In these quantum algorithms, quantum oracles are used to evaluate the value of Boolean function [6]. Here a Boolean function is a function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$ . The function of a quantum oracle is transforming  $|x\rangle|c\rangle$  into  $|x\rangle|c \oplus f(x)\rangle$  [6]. To implement these quantum algorithms on quantum devices, we have to decompose the oracle into elementary gates. Since we are in a noisy intermediate-scale quantum (NISQ) era, the number of qubits and the fidelity and decoherence time of the elementary gate is still at a low level by far [7, 8]. Although the above quantum oracle can be implemented theoretically, the huge number of quantum resources is unavailable in the NISQ era. Therefore, it is essential to implement a quantum oracle with as few quantum costs as possible.

There are several works for the synthesis of quantum oracle [9–12]. Those algorithms focus on different representations for Boolean functions. However, for Conjunction Normal Form (CNF) Boolean function, those algorithms need exponential running time to synthesize such an oracle. Here a CNF Boolean function is an AND of several clauses. Each clause is an OR of variables or their negations. We denote the quantum oracle of the CNF formula as the SAT-oracle.

The well-known NP-hard problem — satisfiability (SAT) problem determines whether a CNF is satisfiable [13–15]. SAT problems appear in several practical application domains, such as gene regulatory networks, model checking, electronic design automation, etc [16–18]. In classical computation and quantum computation, enormous studies aim to solve the SAT problem [19–23]. Those quantum algorithms use SAT-oracle to evaluate the value of the CNF Boolean function. SAT-oracle can also be used in quantum state preparation [24].

Now, we give the definition of the quantum circuit synthesis problem for SAT-oracle. For a given CNF formula  $f$  over  $n$  variables  $x = (x_1, x_2, \dots, x_n)$ , the task is to construct a

quantum circuit  $\mathcal{C}$  such that  $\mathcal{C}|\psi\rangle|c\rangle = |\psi\rangle|c \oplus f(\psi)\rangle$ . For convenience, we denote  $n$  variables  $m$  clauses  $k$ -CNF (each clause contains at most  $k$  variables) as  $\text{CNF}_{n,m}^k$ . Ancillary qubits are widely used in the quantum circuit synthesis and the optimization of quantum circuits. An idea in [25, 26] is to store the value of clauses in the ancillary qubits and then calculate the AND function with a Toffoli gate [27]. When the ancillary qubits are limited, this algorithm fails.

Inspired by the construction of multi-controlled Toffoli (MCT) in [6], we design an algorithm to synthesize a general quantum AND (OR) circuit for functions rather than variables to conquer the limitation of ancillary qubits. Based on the general AND circuit, we design the size-oriented algorithm to synthesize  $\text{CNF}_{n,m}^k$ . The algorithm costs  $\ell$  ancillary qubits and  $\tilde{O}(km^{1+o(1)})$  elementary gates. The size of the circuit decreases rapidly with the growth of  $\ell$ . Particularly when  $\ell = m^\epsilon$ , the circuit size drops to  $O(km/\epsilon)$ . Then, we introduce depth-oriented algorithm to reduce the depth of the quantum circuit to  $\tilde{O}(\log km)$  with  $km$  ancillary qubits, where the size increases by a logarithm factor. When the ancillary qubits is limited, the circuit depth is roughly  $\tilde{O}(km^{1+o(1)}/\ell)$ . The running time of the two algorithms is both  $O(km)$ . The experimental results show that with a tolerable (a constant ratio) increase in the size of the quantum circuit, the number of ancillary qubits is reduced from  $2m - 1$  to  $2\sqrt{m}$  using the size-oriented algorithm. The depth-oriented algorithm significantly reduces the circuit depth of the SAT-oracle. We also give a resource estimate of solving a meaningful SAT problem using Grover's algorithm and our synthesis algorithm.

## RESULT AND METHOD

Consider a general AND problem: given  $p$  Boolean functions  $g_1(x), g_2(x), \dots, g_p(x)$  and the corresponding quantum oracles  $\mathcal{O}_i|x\rangle|c\rangle = |x\rangle|c \oplus g_i(x)\rangle$ . The goal is to construct a quantum circuit  $p$ -GAND such that  $p$ -GAND  $|x\rangle_n|q\rangle_\ell|c\rangle = |x\rangle_n|q\rangle_\ell|c \oplus (\bigwedge_{i=1}^p g_i(x))\rangle$  with  $\ell$  ancillary qubits. We call this general  $p$ -AND function and denote  $p$ -GAND as the general  $p$ -AND circuit. We can define the general  $p$ -OR circuit similarly.

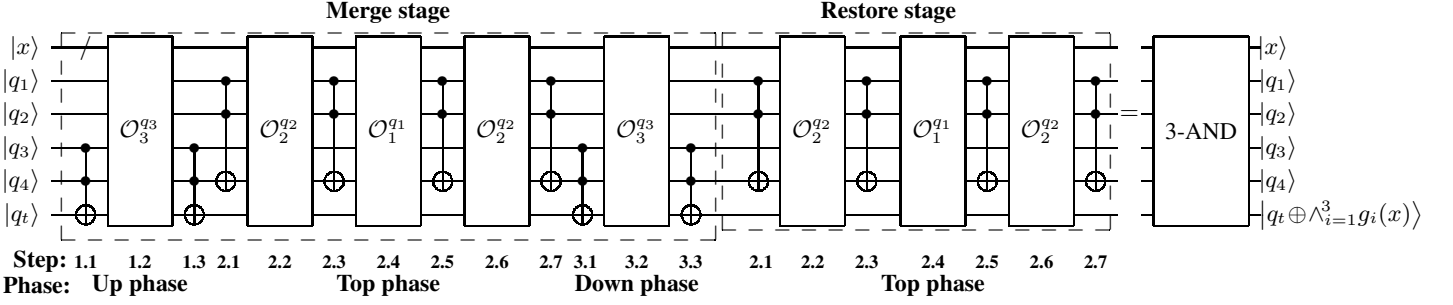


FIG. 1: Construction of 3-GAND circuit.  $|x\rangle$  is input qubits,  $|q_t\rangle$  is target qubit and  $|q_1\rangle, \dots, |q_4\rangle$  are ancillary qubits. Here  $\mathcal{O}_i^q$  is the  $\mathcal{O}_i$  on the target qubit  $q$ .

**Lemma 1** ( $p$ -GAND circuit). *For any natural number  $p$ , general  $p$ -AND circuit can be implemented with  $2p$  dirty ancillary qubits,  $O(p)$  Toffoli gate and 4 calls of each  $\mathcal{O}_i$ .*

Due to the space constraints, we briefly introduce the structure of the circuit implemented by the algorithm [28]. We give an example as illustrated in Figure 1. The construction of  $p$ -GAND is divided into two stages: the merge stage and the restore stage. The merge stage contains  $2p - 3$  steps, and the restore stage is to repeat steps 2 to  $2p - 4$  of the merge stage. Those  $2p - 3$  steps can be further divided into three phases: *Up phase*, *Top phase*, and *Down phase*.

1. *Up phase*: In this phase, we add the information in the ancillary qubits to corresponding qubits, which can help to eliminate unexpected information. Up phase contains the first  $p - 2$  step. Step  $i$  contains 3 sub-steps: 2 Toffoli gates and a call of  $\mathcal{O}_i$ .
2. *Top phase*: In this phase, we implement a circuit that merges two clauses and stores the result in an ancillary qubit. Top phase only contains step  $(p - 1)$ , which has seven sub-steps (4 Toffoli gates, 2 calls of  $\mathcal{O}_2$  and 1 call of  $\mathcal{O}_1$ ).
3. *Down phase*: In this phase, we merge all the clauses in the ancillary qubits. With the help of *Up phase*, the target qubit stores the value of input CNF. There are 3 sub-steps in each step  $i \in \{p, p + 1, \dots, 2p - 3\}$ , which are 2 Toffoli gates and a call of oracle  $\mathcal{O}_{2p-i}$ .

By recursively calling this circuit, we design two efficient algorithms to implement SAT-oracles. The input CNF will be divided into several blocks recursively. In the innermost sub-block, which is numbered 1st level block, we use the general  $k$ -OR circuit to generate a sub-CNF. The circuit to construct  $i$ -th level block is used as an oracle in  $(i + 1)$ -th level block. We give the pseudo-code of size-oriented algorithm and depth-oriented algorithm in the algorithm 1 and 2 respectively.

**Theorem 1.** *By applying algorithm 1, any instance of  $\text{CNF}_{n,m}^k$  can be implemented by an  $O(n(km/n)^{1+\log_{\ell/2+1} 4})$ -size circuit with  $\ell$  ancillary qubits.*

The size-oriented algorithm performs well in the size of the circuit. When the ancillary qubits are clean (the initial state

---

**Algorithm 1: Size-oriented algorithm**

---

**input** : A  $\text{CNF}_{n,m}^k$  instance  $f = \bigwedge_{a=1}^m C_a$  and the number of ancillary qubits  $\ell \geq 3$ .

**output**: A circuit  $\mathcal{C}$  such that

$$\mathcal{C}|x\rangle|0\rangle_\ell|c\rangle = |x\rangle|0\rangle_\ell|c \oplus f(x)\rangle, \forall x \in \{0, 1\}^n.$$

- 1  $d \leftarrow \log_{\ell/2} m, s \leftarrow \frac{2m}{\ell};$
  - 2 **for**  $j$  in 1 to  $\ell/2$ :
  - 3     **Clause**  $((j-1)s+1, js, \text{Ancilla}[j], d-1);$
  - 4     **MCT** (Ancilla, Target);
  - 5 **for**  $j$  in 1 to  $\ell/2$ :
  - 6     **Clause**  $((j-1)s+1, js, \text{Ancilla}[j], d-1);$
  - 7 **Clause**(Sid, Eld, Target, Depth) :
  - 8 **if** (Depth=0):
  - 9     Synthesize clauses on target qubit;
  - 10    **return**;
  - 11  $s \leftarrow (\text{Eld}-\text{Sid})/(\ell/2+1);$
  - 12 Apply  $(2\ell+1)$ -GAND circuit, where  $\mathcal{O}_i$  is synthesized by
  - Clause**  $(\text{Sid}+(i-1)s, \text{Sid}+is-1, \text{Ancilla}[i], \text{Depth}-1);$
  - 13 **return**;
- 

is  $|0\rangle$ ), we design depth-oriented algorithm to further reduce the circuit depth, with a little increase in the circuit size. We partition the ancillary qubits into 3 registers,  $q_{\text{mem}}$ ,  $q_{\text{dirty}}$  and  $q_{\text{clean}}$ . The qubits in  $q_{\text{mem}}$  and  $q_{\text{clean}}$  are clean at the beginning. The synthesis framework is implemented with the  $q_{\text{dirty}}$  register. The difference between the two algorithms is in the innermost recursion, where we use the  $q_{\text{mem}}$  and  $q_{\text{clean}}$  to parallel the quantum circuit.

**Theorem 2.** *By applying algorithm 2, any instance of  $\text{CNF}_{n,m}^k$  can be implemented by an  $O\left(k\left(\frac{mS}{\ell}\right)^{1+c} \log \ell\right)$ -depth circuit, where  $\ell$  is the number of ancillary qubits,  $S = \max\{\frac{k}{\log \ell}, 1\}$  and  $c = \log_{\ell/2} 4$ .*

The depth of the quantum circuit is declined rapidly with the growth of the number of ancillary qubits. Some numerical experiments are designed to show the performance of our algorithm in the following subsection. We also point out the asymptotically lower bound for synthesizing the CNF formula by counting. The classical running time is the same as the number of calls to the clause function, which is polynomial to the input size. Using the counting method, we show the lower

**Algorithm 2:** Depth-oriented algorithm

---

**input :** A  $\text{CNF}_{n,m}^k$  instance  $f = \bigwedge_{a=1}^m C_a$  and the number of ancillary qubits  $\ell \geq 3$ .  
**output:** A circuit  $\mathcal{C}$  such that  $\mathcal{C}|x\rangle|0\rangle_\ell|c\rangle = |x\rangle|0\rangle_\ell|c \oplus f(x)\rangle, \forall x \in \{0, 1\}^n$ .

- 1  $S \leftarrow \max\{k/\log \ell, 1\}, d \leftarrow \log_{\ell/2(S+1)} m, s \leftarrow \frac{2(S+1)m}{\ell}$ ;
- 2 Divide the ancillary qubits into 3 parts  $q_{mem}, q_{dirty}, q_{clean}$ , size of each part is  $(S-1)\ell/(S+1), \ell/(S+1), \ell/(S+1)$ ;
- 3 **for**  $j$  in 1 to  $\frac{\ell}{2(S+1)}$ :
- 4     **Clause**  $((j-1)s+1, js, q_{dirty}[j], d-1)$ ;
- 5     **MCT**  $(q_{dirty}, \text{Target})$ ;
- 6 **for**  $j$  in 1 to  $\frac{\ell}{2(S+1)}$ :
- 7     **Clause**  $((j-1)s+1, js, q_{dirty}[j], d-1)$ ;
- 8 **Clause**  $(SId, Eld, \text{Target}, \text{Depth})$  :
- 9 **if**  $(\text{Depth}=0)$ :
- 10     Copy the input to  $q_{mem}$ ;
- 11     Use the input and  $q_{mem}$  to synthesize  $\ell/(S+1)$  clauses in parallel at  $q_{dirty}$ ;
- 12     Use the  $q_{clean}$  to merge all the clauses in parallel;
- 13     Reset the ancillary qubits;
- 14     **return**;
- 15  $s \leftarrow \frac{2(S+1)(EId-SId)}{\ell}$ ;
- 16 Apply the GAND circuit, where  $O_i$  is synthesized by **Clause**  $(SId+(i-1)s, SId+is-1, q_{dirty}[i], \text{Depth}-1)$ ;
- 17 **return**;

---

bound of the SAT-oracle synthesis problem.

**Theorem 3.** *There exists an instance of  $\text{CNF}_{n,m}^k$  such that any quantum circuits approximating it with error  $\varepsilon < \frac{\sqrt{2}}{2}$  must have size at least  $\Omega(km)$ .*

Combining the Theorem 1 and Theorem 3, we see that our algorithm is asymptotically optimal when the number of ancillary qubits is  $\Omega((km)^\epsilon)$  for any  $\epsilon > 0$ .

Both two algorithms show significant advantages compare to the previous work. For the size-oriented algorithm, we can reduce the number of ancillary qubits to  $O(\sqrt{m})$  with only a constant ratio in size. For the depth-oriented algorithm, we use  $O(\log km)$  ancillary qubits to construct the same depth circuit generated by qiskit [28].

We use random  $k$ -CNF as the experimental benchmark to test the performance of different algorithms. To sample a  $\text{CNF}_{n,m}^k$ , we first randomly sample  $k$  variables from the input variables, then randomly choose the variables or the negations of the variables. After two steps, a clause of  $\text{CNF}_{n,m}^k$  is generated. Then, repeating the first two steps  $m$  times, we generate a random  $\text{CNF}_{n,m}^k$ . We randomly sample 100 CNFs of different parameters and use the average quantum cost to measure the performance of quantum synthesis algorithms. The width of the CNF  $k$  is 3 and 4, and the number of variables  $n$  is 40, 80, 400, and 800. The number of clauses  $m$  we choose in this manuscript is determined by the number of variables  $n$  and the width of CNF  $k$ . When  $k$  is 3 and 4, there are  $m = \lfloor 4.267n \rfloor$  and  $m = \lfloor 9.931n \rfloor$ , respectively, which is

called SAT phase transition [29]. Our algorithm is suitable for all the input. To evaluate the quantum cost to conquer the most difficult SAT instance, we choose such a specific  $m$  in our experiments. Here, the size and the depth of the quantum circuit are considered appropriate quantum costs in the NISQ era. To verify the relationship between the number of ancillary qubits  $\ell$  and the quantum cost of our two algorithms, we choose several different  $\ell$ . Some are near to the  $n$ , and others are near to  $2m-1$ . The results of different widths (the number of variables in a clause) seem similar. For convenience, we plot the result of 4-CNF in the Figure 2, which is appropriate to show the performance of our algorithms.

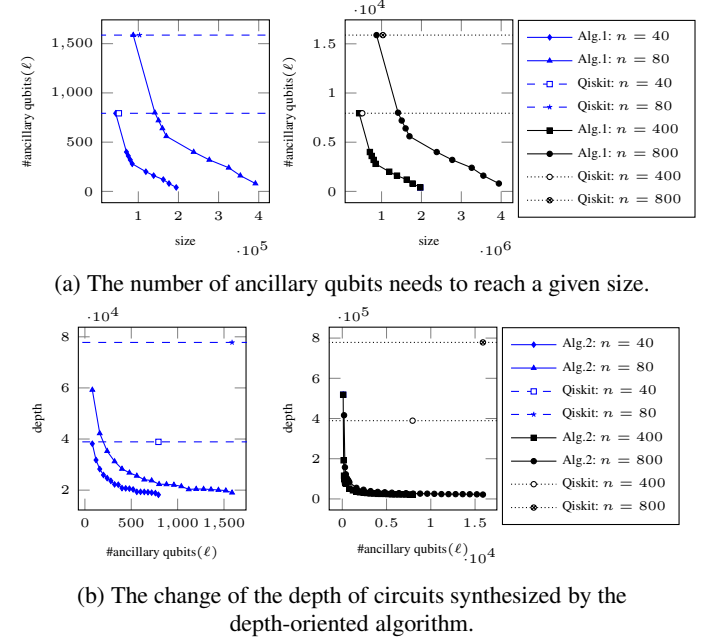


FIG. 2: The performance of algorithms. Choose the random 4-CNF with 40, 80, 400, and 800 variables to compare the performance of our algorithm and qiskit's. The result of qiskit contains only one point.

In Figure 2 (a), we compare the size of the quantum circuit synthesized by our algorithm and the CNF synthesis algorithm used in the qiskit. The circuit synthesis algorithm in qiskit requires  $2m-1$  ancillary qubits, so there is only one point and the corresponding horizontal dashed line in Figure 2 [27]. This Figure shows that when  $2m-1$  ancillary qubits are used, the size-oriented algorithm can generate a quantum circuit with a smaller size than qiskit. If we want to reduce the number of ancillary qubits significantly, the corresponding size will only increase by a constant multiple. For example, for a random 4-CNF with  $n$  equal to 80, qiskit needs 1587 ancillary qubits to synthesize a quantum circuit of size 103205, while the size-oriented algorithm can use 80 ancillary qubits (about 5% of qiskit) to synthesize a circuit of size 391760 (less than 4 times).

In Figure 2 (b), a more notable advantage appears in comparing the circuit depth between the depth-oriented algorithm

and the algorithm used in qiskit. The depth-oriented algorithm requires very few ancillary qubits to synthesize low-depth quantum circuits. When using the same  $2m - 1$  ancillary qubits as qiskit, the circuit depth synthesized by depth-oriented algorithm is significantly lower. For example, for a random 4-CNF with  $n$  equal to 800, qiskit needs 15887 ancillary qubits to synthesize a quantum circuit of depth 778498. In comparison, the depth-oriented algorithm can use 200 ancillary qubits (about 1.2% of qiskit) to synthesize a circuit of depth 416179 (about 53.5% of qiskit). If the depth-oriented algorithm uses 15887 ancillary qubits, the depth of the circuit can be reduced to 21735 (about 2.8% of qiskit). For both two algorithms, the quantum cost of the output circuit is declined with the growth of the number of ancillary qubits. Despite a few single points, our experimental results are in good agreement with the theory.

$k$	$n$	#clause	#ancillae	size (Alg.1)		depth (Alg.2)	
				full round	one round	full round	one round
3	40	170	240	$1.8 \times 10^{10}$	21384	$6.2 \times 10^9$	7523
3	80	341	240	$4.3 \times 10^{16}$	49522	$1.0 \times 10^{16}$	11586
5	40	844	240	$4.3 \times 10^{11}$	$5.2 \times 10^5$	$7.6 \times 10^{10}$	92444
5	80	1689	240	$1.0 \times 10^{18}$	$1.2 \times 10^6$	$1.1 \times 10^{17}$	$1.3 \times 10^5$
7	40	3511	240	$3.4 \times 10^{12}$	$4.1 \times 10^6$	$6.7 \times 10^{11}$	$8.1 \times 10^5$
7	80	7023	240	$7.3 \times 10^{18}$	$8.4 \times 10^6$	$9.3 \times 10^{17}$	$1.1 \times 10^6$

TABLE I: The quantum resource estimation for solving  $k$ -SAT problems using Grover’s algorithm, where  $k = 3, 5, 7$ .  $n$  is the number of variables. The left 4 columns are the parameters of the  $k$ -SAT problem instance, and the rest columns show the quantum cost needed to solve the  $k$ -SAT problem via Grover’s algorithm and our synthesis algorithms. We list the full round cost and the one round cost separately.

Further, we apply our synthesis algorithms to estimate the quantum resources required for solving  $k$ -SAT using Grover’s algorithm. In [26], they also estimate the quantum resources required for solving 14-SAT algorithm. To solve a 65 to 78 variables 14-SAT, the number of ancillary qubits used is  $10^{12}$  to  $10^{14}$ , which is unavailable in NISQ era. Hence, we fix the number of ancillary qubits as 240, which is more available in NISQ era. In Grover’s algorithm there is multiple rounds of Grover iteration. Table I shows the quantum resources needed for both the full round (reaching the highest success probability) and one round execution of Grover iteration. For example, a random 7-SAT with 80 variables, which is among the most challenging instances that a classical computer can solve today [30–34], can be solved by using 240 ancillary qubits and a  $7.3 \times 10^{18}$ -size quantum circuit via Grover’s algorithm.

## DISCUSSION

In this manuscript we design two synthesis algorithms for the different quantum costs. We first construct a general  $p$ -AND circuit. We design the size-oriented algorithm, which recursively use  $p$ -AND module to construct the circuit for

SAT-oracle. Notice that the ancillary qubits used here could be dirty, which means this algorithm can use the temporarily idle qubits. The size of the circuit generated with this algorithm is  $O(n(km/n)^{1+c})$ , where  $c = o(1)$  is determined by  $\ell$ . Specially we can reduce the number of ancillary qubits to  $O(\sqrt{m})$  with a constant ratio increase in the size of the circuit. We also prove a matched lower bound of this problem using counting method. Further, we propose the depth-oriented algorithm to reduce the depth of the quantum circuit. The depth of the circuit generated by depth-oriented algorithm is  $O(k \log \ell (mS/\ell)^{1+c'})$ , where  $S = \max\{k/\log \ell, 1\}$  and  $c' = o(1)$  is determined by  $\ell$ . We design several experiments to evaluate the performance of our algorithm. Finally, we apply our synthesis algorithms to give an estimate of quantum costs to solve the SAT problem.

Some interesting open problems left. Can we use the dirty ancillary qubits to replace the clean ancillary qubits in the synthesis algorithm? Is there some essential difference between the clean and dirty ancillary qubits in general circuit? Are there some efficient algorithms that can generate size-optimal or depth-optimal circuit for any given  $\text{CNF}_{n,m}^k$  instance?

- 
- [1] RP Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6), 1982.
  - [2] Peter W. Shor. Polynomial time algorithms for discrete logarithms and factoring on a quantum computer. In *Algorithmic Number Theory, First International Symposium, ANTS-I, Ithaca, NY, USA, May 6-9, 1994, Proceedings*, 1994.
  - [3] Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996.
  - [4] Salman Beigi and Leila Taghavi. Quantum speedup based on classical decision trees. *Quantum*, 4:241, 2020.
  - [5] Frédéric Magniez, Miklos Santha, and Mario Szegedy. Quantum algorithms for the triangle problem. *SIAM Journal on Computing*, 37(2):413–424, 2007.
  - [6] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.
  - [7] John Preskill. Quantum computing in the nisy era and beyond. *Quantum*, 2:79, 2018.
  - [8] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.
  - [9] Vivek V Shende, Aditya K Prasad, Igor L Markov, and John P Hayes. Synthesis of reversible logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(6):710–722, 2003.
  - [10] D Michael Miller, Dmitri Maslov, and Gerhard W Dueck. A transformation based algorithm for reversible logic synthesis. In *Proceedings of the 40th annual Design Automation Conference*, pages 318–323, 2003.
  - [11] Robert Wille and Rolf Drechsler. BDD-based synthesis of reversible logic for large functions. In *Proceedings of the 46th Annual Design Automation Conference*, pages 270–275, 2009.



- [12] Kenneth Fazel, Mitchell A Thornton, and JE Rice. ESOP-based toffoli gate cascade generation. In *2007 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pages 206–209. IEEE, 2007.
- [13] CNF is an AND of several clauses, each clause is an OR of the variables or their negations. The number of variables is usually denoted by  $n$  and the number of clauses is denoted by  $m$ . The number of variables used in a single clause is called the width of the clauses, and the width of CNF is defined by the max width of the clauses, which is denoted by  $k$ .
- [14] Stephen A Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158, 1971.
- [15] Leonid Anatolevich Levin. Universal sequential search problems. *Problemy peredachi informatsii*, 9(3):115–116, 1973.
- [16] Fabien Corblin, Lucas Bordeaux, Youssef Hamadi, Eric Fanchon, and Laurent Trilling. A sat-based approach to decipher gene regulatory networks. *Integrative Post-Genomics, RIAMS, Lyon*, 2007.
- [17] Kenneth L McMillan. Interpolation and sat-based model checking. In *International Conference on Computer Aided Verification*, pages 1–13. Springer, 2003.
- [18] Wolfgang Kunz and Dominik Stoffel. *Reasoning in Boolean Networks: logic synthesis and verification using testing techniques*, volume 9. Springer Science & Business Media, 1997.
- [19] T Schoning. A probabilistic algorithm for k-sat and constraint satisfaction problems. In *40th Annual Symposium on Foundations of Computer Science (Cat. No. 99CB37039)*, pages 410–414. IEEE, 1999.
- [20] Ramamohan Paturi, Pavel Pudlák, Michael E Saks, and Francis Zane. An improved exponential-time algorithm for k-sat. *Journal of the ACM (JACM)*, 52(3):337–364, 2005.
- [21] Thomas Dueholm Hansen, Haim Kaplan, Or Zamir, and Uri Zwick. Faster k-sat algorithms using biased-ppsz. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 578–589, 2019.
- [22] Vedran Dunjko, Yimin Ge, and J Ignacio Cirac. Computational speedups using small quantum devices. *Physical Review Letters*, 121(25):250501, 2018.
- [23] Alberto Leporati and Sara Felloni. Three “quantum” algorithms to solve 3-sat. *Theoretical Computer Science*, 372(2-3):218–241, 2007.
- [24] Gregory Rosenthal. Query and depth upper bounds for quantum unitaries via grover search. *arXiv preprint arXiv:2111.07992*, 2021.
- [25] IBM Q. <https://www.research.ibm.com/ibm-q/>.
- [26] Earl Campbell, Ankur Khurana, and Ashley Montanaro. Applying quantum algorithms to constraint satisfaction problems. *Quantum*, 3:167, 2019.
- [27] The synthesis algorithm used in qiskit has been changed. When we wrote this manuscript, the algorithm used in qiskit is first generating all clauses in the clean ancillary qubits and then merged them up. The new algorithm used in qiskit is a heuristic algorithm, which can not work when the input  $n$  is large. Hence, we still compare the result with the old synthesis algorithm in qiskit.
- [28] More detail can be found in supplemental material at (link).
- [29] David Mitchell, Bart Selman, Hector Levesque, et al. Hard and easy distributions of sat problems. In *AAAI*, volume 92, pages 459–465. Citeseer, 1992.
- [30] SAT 2013 Competition. <http://www.satcompetition.org/2013/>.
- [31] SAT 2014 Competition. <http://www.satcompetition.org/2014/>.
- [32] SAT 2016 Competition. <https://baldur.iti.kit.edu/sat-competition-2016/>.
- [33] SAT 2017 Competition. <https://baldur.iti.kit.edu/sat-competition-2017/>.
- [34] SAT 2018 Competition. <https://satcompetition.github.io/2018/>.

# Supplementary materials for “Efficient quantum circuit synthesis for SAT-oracle with limited ancillary qubit”

Shuai Yang,<sup>1,2</sup> Wei Zi,<sup>1,2</sup> Bujiao Wu,<sup>3</sup> Cheng Guo,<sup>1,2</sup> Jialin Zhang,<sup>1,2</sup> and Xiaoming Sun<sup>1,2</sup>

<sup>1</sup>*Institute of Computing Technology, Chinese Academy of Sciences, 100190 Beijing, China*

<sup>2</sup>*University of Chinese Academy of Sciences, 100049 Beijing, China*

<sup>3</sup>*Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China*

## NOTATION AND PRELIMINARY

A Boolean formula consists of the variables and logical operations AND, OR, and NOT. In specific, a Conjunction Normal Form (Disjunctive Normal Form) formula is the AND (OR) of OR's (AND's) of variables or their negations. To synthesis a CNF (DNF) formula  $f$  is to find a quantum circuit  $C_f$  such that for any input  $x$ , we have  $C_f |x\rangle |y\rangle \rightarrow |x\rangle |y \oplus f(x)\rangle$ . X gate, Controlled-NOT gate (CNOT gate), and Toffoli gate form a universal gate set for Boolean functions oracle. A layer in a quantum circuit is a set of consecutive disjoint quantum gates. We use the number of the elementary gates (CNOT gate and single-qubit gate) and the number of layers to measure the cost of a quantum circuit.

For convenience, we use the notation  $\text{CNF}_{n,m}^k$  to denote an instance of  $n$  variables  $m$  clauses  $k$ -CNF and  $\text{Size}_\ell(\text{CNF}_{n,m}^k)$  to denote the size of the circuit generated by the first algorithm for  $\text{CNF}_{n,m}^k$  with  $\ell$  ancillary qubits. We may omit the rounding notation ‘ $\lceil \cdot \rceil$ ’ for convenience in the manuscript.

A multi-controlled Toffoli gate in quantum circuit can calculate AND/OR of variables. Denote  $\text{Toff}_{q_1, q_2, \dots, q_k}^{q_t}$  as the  $k$ -controlled Toffoli gate, where the control qubits are  $\{q_1, q_2, \dots, q_k\}$  and the target qubit is  $q_t$ . For examples:

$$\begin{aligned} \text{Toff}_{q_1, q_2, \dots, q_k}^{q_t} |q\rangle |c\rangle &\rightarrow |q\rangle |c \oplus \bigwedge_{j=1}^k q_j\rangle, \\ (\bigotimes_{j=1}^k X_j) X_{q_t} (\text{Toff}_{q_1, q_2, \dots, q_k}^{q_t}) (\bigotimes_{j=1}^k X_{q_j}) |q\rangle |c\rangle \\ &\rightarrow |q\rangle |c \oplus \bigvee_{j=1}^k q_j\rangle. \end{aligned}$$

## THE DETAIL OF GENERAL AND (OR) CIRCUIT

**Lemma 1** (general  $p$ -AND circuit). *For any natural number  $p$ , general  $p$ -AND circuit can be implemented with  $2p-2$  dirty ancillary qubits,  $O(p)$  Toffoli gate and 4 calls of each  $\mathcal{O}_i$ .*

For convenience, let us introduce some notations.

$$q'_i = \begin{cases} q_i \oplus (q_{i-1} \wedge g_{i-p+1}(x)) & 2p-2 \geq i \geq p+2 \\ q_i \oplus (q_1 \wedge g_2(x)) & i = p+1 \\ q_i \oplus g_i(x) & p \geq i \geq 1 \end{cases}$$

and  $q''_i = q_i \oplus (\bigwedge_{k=1}^{i-p+1} g_k(x))$ ,  $i \in \{p+1, p+2, \dots, 2p-2\}$ . Let  $q_t$  be the target qubit,  $q_t^a = q_t \oplus (\bigwedge_{k=1}^a g_k(x))$ , and  $q_t' = q_t \oplus (q_p \wedge g_p(x))$ . Let  $Q_{a,b} = \bigotimes_{k=a}^b |q_k\rangle$ ,  $Q'_{a,b} = \bigotimes_{k=a}^b |q'_k\rangle$ , and  $Q''_{a,b} = \bigotimes_{k=a}^b |q''_k\rangle$ . The number of dirty ancillary qubits  $\ell = 2p-2$ .

*Proof.* Using Toffoli gate can easily merge the information stored in the qubits. However if the information is stored in the oracle, we need to apply the oracle in some qubits at first. The information stored in the qubits will influence the result. Hence, we add additional operation to eliminate the dirty information. We firstly divide the  $\mathcal{C}$  into 2 sub-circuit: merge stage  $\mathcal{C}_2$  and restore stage  $\mathcal{C}_1$ . The construct of quantum circuit  $\mathcal{C}_1$  contains  $2p-3$  steps. These steps can be divided into 3 phases: *Up phase*, *Top phase* and *Down phase*. The *Up phase* first store the information about  $q_{p+1-i}$  and oracle  $\mathcal{O}_{p+1-i}$  at qubit  $q_{2p-1-i}$ . The *Top phase* then merge the information about  $\mathcal{O}_1$  and  $\mathcal{O}_2$ . In the *Down phase*, we merge the all the information store in the  $\mathcal{O}_i$ . Notice the dirty information is added twice under  $\mathbb{F}_2$ . Finally in  $\mathcal{C}_2$  we repeat step 2 to step  $2p-4$  to restore the ancillary qubits.

1. *Up phase*: In this phase, we add the information in the ancillary qubits to corresponding qubits, which can help to eliminate unexpected information. There are 3 parts in step  $i \in [p-2]$ . The details of each part are shown as follows.

- (a) In step  $i.1$ , as well as step  $i.3$ , corresponds to a Toffoli gate. The control qubits are  $q_{p+1-i}$  and  $q_{2p-1-i}$ . When  $i = 1$ , the target qubit is  $q_t$ , otherwise the target is  $q_{2p-i}$ .
- (b) In step  $i.2$ , we call the  $\mathcal{O}_{p+1-i}$  at  $q_{p+1-i}$ .

At the begining, the state is at the

$$Q_{1,\ell} |q_t\rangle$$

After step 1.1, we apply an Toffoli gate to store the dirty information stored in the dirty ancillary qubits  $q_{p+1}$  and  $q_{2p}$ . Then the state transfer to

$$Q_{1,\ell} |q_t \oplus (q_p \wedge q_{2p-2})\rangle.$$

Then, we apply an  $\mathcal{O}_{p+1}$  at the qubit  $q_{p+1}$ . This step add the information of the oracle into the qubits. The state after step 1.2 becomes

$$Q_{1,p-1} Q'_{p,p} Q_{p+1,2p-2} |q_t \oplus (q_p \wedge q_{2p-2})\rangle.$$

Finally in step 1.3, we use a Toffoli gate to add the information to the target qubit and some dirty information has been eliminated. The state after step 1 is

$$Q_{1,p-1} Q'_{p,p} Q_{p+1,2p-2} |q_t'\rangle.$$

Same to the analyze in the step 1, the state before the step  $i.1$  is

$$Q_{1,p-i+1} Q'_{p-i+2,p} Q_{p+1,2p-i} Q'_{2p-i+1,2p-2} |q'_t\rangle.$$

After step  $i.3$ , the state transfer to

$$Q_{1,p-i} Q'_{p-i+1,p} Q_{p+1,2p-i-1} Q'_{2p-i,2p-2} |q'_t\rangle.$$

2. *Top phase*: In this phase, we implement a circuit that merges two clauses and stores the result in an ancillary qubit. There are seven parts in step  $p-1$ . The details of each part are shown as follows.

- (a) In steps  $(p-1).j$ ,  $j \in \{1, 3, 5, 7\}$ , the operation are the same. Each step corresponds to a Toffoli gate. The control qubits of Toffoli gate is  $q_1$  and  $q_2$ , and the target qubit is  $q_{p+1}$ .
- (b) In steps  $(p-1).2$ , as well as  $p.6$ , we call the  $\mathcal{O}_2$  at  $q_2$ .
- (c) In steps  $(p-1).4$ , we call the  $\mathcal{O}_1$  at  $q_1$ .

The step from  $(p-1).1$  to  $(p-1).3$  are similar to the step in *Up phase*. The state after step  $(p-1).3$  is

$$|q_1\rangle Q'_{2,2p-2} |q'_t\rangle$$

The step  $(p-1).4$  just apply an Oracle  $\mathcal{O}_1$ . So the state is  $Q'_{1,2p-2} |q'_t\rangle$ .

The step from  $(p-1).5$  to  $(p-1).7$  restore the qubit 2 and store the  $g_1 \wedge g_2(x)$  without dirty information in qubit  $p+1$ . The equation below only focus on the qubits  $q_1, q_2$  and  $q_{p+2}$ .

$$\begin{aligned} & |q'_1\rangle |q'_2\rangle |q'_{p+1}\rangle \\ \xrightarrow{(p-1).5} & |q'_1\rangle |q'_2\rangle |q'_{p+1} \oplus ((q_1 \oplus g_1(x)) \wedge (q_2 \oplus g_2(x)))\rangle \\ \xrightarrow{(p-1).6} & |q'_1\rangle |q_2\rangle |q'_{p+1} \oplus ((q_1 \oplus g_1(x)) \wedge (q_2 \oplus g_2(x)))\rangle \\ \xrightarrow{(p-1).7} & |q'_1\rangle |q_2\rangle |q'_{p+1} \oplus ((q_1 \oplus g_1(x)) \wedge g_2(x))\rangle \\ & = |q'_1\rangle |q_2\rangle |q_{p+1} \oplus (g_1(x) \wedge g_2(x))\rangle \end{aligned}$$

3. *Down phase*: In this phase, we merge all the clauses in the ancillary qubits. With the help of *Up phase*, the target qubit stores the value of input CNF. There are 3 parts in step  $i \in \{p, p+2, \dots, 2p-3\}$ .

- (a) Step  $i.1$ , as well as step  $i.3$ , corresponds to a Toffoli gate. The control qubits are  $q_{i-p+3}$  and  $q_{i+1}$ . When  $i = 2p-3$ , the target qubit is  $q_t$ , otherwise the target is  $q_{i+2}$ .
- (b) In step  $i.2$ , we call the  $\mathcal{O}_{i-p+3}$  at  $q_{i-p+3}$ .

The analyse is similar to the step  $(p-1).5$  to  $(p-1).7$ . In each step, we restore a qubit  $i-p+3$  and store the correct information in qubit  $q_{i+3}$ .

After the  $\mathcal{C}_1$ , the state is

$$|q'_1(x)\rangle Q_{2,p} Q''_{p+1,2p-2} |q_t \oplus f(x)\rangle.$$

What we need to do is repeating the steps from 2 to  $2p-4$  to restore the ancillary qubits.

Totally, we use at most  $8p-12$  Toffoli gates and call each  $\mathcal{O}_i$  at most four times. In the merge stage, each step except step  $p$  contain 2 Toffoli gates. The total number of Toffoli gates in merge stage is  $2(2p-3-1)+4=4p-4$ . Similarly, we use  $4p-8$  Toffoli gates in restore stage. So we use at most  $8p-12$  Toffoli gates in  $p$ -AND circuit. In each stage, each oracle  $\mathcal{O}_i$  is called at most 2 times. So each  $\mathcal{O}_i$  is called at most 4 times in the  $p$ -AND circuit.  $\square$

**Corollary 1.** *By adding  $X$  gates when we call the  $\mathcal{O}$  and adding  $X$  gates at target qubit, we can construct a circuit  $\mathcal{C}$  for function  $f(x) = \bigvee g_i(x)$ .*

### SIZE-ORIENTED SYNTHESIS ALGORITHM

This section will introduce how to use the general  $p$ -AND circuit to construct the circuit for SAT-Oracle.

**Lemma 2.** *Any  $n$  variables  $m$  clauses  $k$ -CNF  $f \in \text{CNF}_{n,m}^k$  can be implemented by  $O(km^{1+\log_{\ell/2+1} 4})$ -size quantum circuits  $\mathcal{C}$  with  $\ell$  ancillary qubits.*

*Proof.* We divide the clauses in CNF into  $p = \lfloor \ell/2 \rfloor + 1$  sub-blocks recursively until in each blocks only one clause or less. An oracle  $\mathcal{O}$  for a single clause can be realized by a  $k$ -controlled Toffoli gate and several  $X$  gate, where  $k$  is the width of the clause, which means  $\text{Size}_0(\text{CNF}_{n,1}^k) = O(k)$ .

Lemma 1 shows that we can construct a circuit to calculate the AND(OR) of some sub-function. By recursively using the general  $p$ -AND/OR circuit, we can finally construct the circuit for any given CNF. In each recursion we need  $O(\ell)$  elementary gates and call  $4p$  sub-blocks oracle.

The total quantum cost is  $\text{Size}_{\ell}(\text{CNF}_{n,m}^k) = 4p\text{Size}_{\ell}(\text{CNF}_{n,m/p}^k) + 4\ell$ . Notice that  $\text{Size}_{\ell}(\text{CNF}_{n,1}^k) \leq \text{Size}_0(\text{CNF}_{n,1}^k) = O(k)$ . Solving this recursion formula, we have that  $\text{Size}_{\ell}(\text{CNF}_{n,m}^k) = O(km^{1+\log_{\ell/2+1} 4})$ .  $\square$

Ancillary qubits used in this section are dirty ancillary qubits, which means we can use the unused input qubits as ancillary qubits. In a  $t$  ( $t \leq \frac{n-n^{\epsilon}}{k}$ ) clauses CNF formula, at most  $kt$  variables is used in this CNF formula. To generate such a CNF formula, we can regard other input qubits as ancillary qubits, which means  $\text{Size}_0(\text{CNF}_{n, \frac{n-n^{\epsilon}}{k}}^k) = \text{Size}_{n^{\epsilon}}(\text{CNF}_{n, \frac{n-n^{\epsilon}}{k}}^k) = O(n)$ . This idea can improve the upper bound that  $\text{Size}_{\ell}(\text{CNF}_{n,m}^k) = O\left(n \left(\frac{km}{n-n^{\epsilon}}\right)^{1+\log_{\ell/2+1} 4}\right)$ . Then the upper bound is proved.

If the ancillary qubits are clean, which means the initial state of ancillary qubits are  $|0\rangle$  at the beginning, the outermost recursive circuit can be improved. We can merge  $\ell$  terms

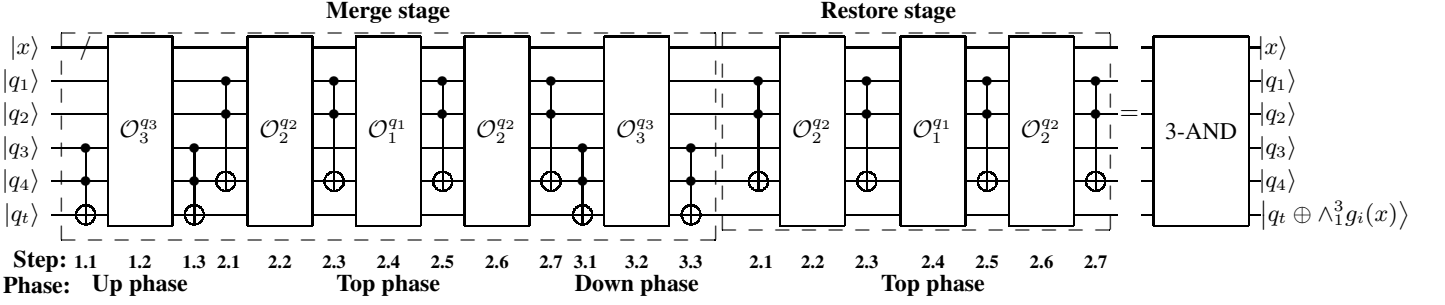


FIG. 1: A example of general 3-AND circuit.

into one in the outer-most recursion. This optimization will reduce the quantum cost by about half.

Approximate Toffoli gate consists of 3 CNOT gates and four single-qubit gates. Like the Toffoli gate, approximate Toffoli gate can implement AND/OR gate in the quantum circuit, with an additional change over control qubit. Using approximate Toffoli to synthesize Toffoli gate can reduce the quantum cost of  $m$ -control Toffoli gate.

When the number of ancillary qubits is small, we can use the circuit shown in Figure 2 to replace the circuit in Figure 1. The recursive circuit shown in Figure 2 can merge  $\ell$  functions, rather than  $\ell/2$  functions shown in the previous discussion, in one recursion.

Choosing these optimizations can help us synthesize CNF with less quantum cost.

### DEPTH-ORIENTED SYNTHESIS ALGORITHM

To further reduce the depth of the circuit, we need to parallelize our construct algorithm. In the size-oriented synthesis algorithm, each sub-block is implemented one by one. So we try to use some of the ancillary qubits, which are clean, to parallelize the circuit.

The framework of the depth-oriented synthesis algorithm is similar to the algorithm described in the size-oriented synthesis algorithm. Different from the size-oriented synthesis algorithm, we divide the ancillary qubits into 3 registers:  $q_{mem}$ ,  $q_{dirty}$ ,  $q_{clean}$ . The size of these 3 registers are  $\frac{(S-1)\ell}{S+1}$ ,  $\frac{\ell}{S+1}$ ,  $\frac{\ell}{S+1}$ , where  $S = \max\{\frac{k}{\log \ell}, 1\}$ . We run the recursive procedure by using  $q_{dirty}$  as ancillary qubits. The only difference lies in the inner-most recursion of our algorithm. We use all the ancillary qubits to synthesize  $\frac{\ell}{S+1}$  clauses in parallel.

The inner-most recursion of our circuit is shown in Figure 3. There are four stages in the inner-most recursion: Copy stage, Clause stage, Merge stage, and Reset stage.

For convenience, let  $\mathcal{CO}$ ,  $\mathcal{CL}$ ,  $\mathcal{ME}$ ,  $\mathcal{RE}$  to denote the Copy stage, Clause stage, Merge stage and Reset stage, respectively. After these 4 stages, we synthesize a  $f = \bigwedge_{j=1}^{\lfloor \ell/S \rfloor} C_j \in \text{CNF}_{n, \lfloor \ell/S \rfloor}^k$  to the target qubit. Without loss of generality, let  $\lfloor \ell/S \rfloor$  is a even.

In the Copy stage  $\mathcal{CO}$ , we copy the information of input

qubits to the  $q_{mem}$  register.

$$\mathcal{CO} |x\rangle |0\rangle |q\rangle |0\rangle |q_t\rangle \rightarrow |x\rangle (\otimes_i |x_i\rangle^{\otimes t_i}) |q\rangle |0\rangle |q_t\rangle, \quad (1)$$

where the number  $t_i$  is determined by the input Boolean function. The depth of Copy stage is  $\log_2(\max_i t_i) = O(\log \ell)$ .

In the Clause stage, we use the information in the  $q_{mem}$  register and input qubits to synthesize clauses in parallel. The result is stored in the first half of  $q_{clean}$  register.

$$\begin{aligned} \mathcal{CL} |x\rangle (\otimes_i |x_i\rangle^{\otimes t_i}) |q\rangle |0\rangle |q_t\rangle \\ \rightarrow |x\rangle (\otimes_i |x_i\rangle^{\otimes t_i}) |q\rangle \left( \otimes_{i=1}^{\lfloor \ell/2S \rfloor} |C_{2i-1} \wedge C_{2i}\rangle \right) |q_t\rangle. \end{aligned}$$

We can synthesize  $O(\ell/k)$  terms with  $O(k)$ -depth circuit. The total depth of Clause stage is  $O(k \log \ell)$ .

In the merge stage, we merge all the clauses stored in the  $q_{clean}$  to the target qubit.

$$\begin{aligned} \mathcal{ME} \left( \otimes_{i=1}^{\lfloor \ell/2S \rfloor} |C_{2i-1} \wedge C_{2i}\rangle \right) |q_t\rangle \\ \rightarrow \left( \otimes_{i=1}^{\lfloor \ell/2S \rfloor} |C_{2i-1} \wedge C_{2i}\rangle \right) |q_t \oplus f(x)\rangle. \end{aligned}$$

A Toffoli gate can merge two CNF formulae on a clean ancillary qubit. To merge  $\ell/2(S+1)$  CNF formulae, the depth of merge stage is  $O(\log(\ell/2(S+1))) = O(\log \ell)$ .

We repeat the Copy stage and the Clause stage to reset all the ancillary qubits in the reset stage.

$$\begin{aligned} \mathcal{RE} |x\rangle (\otimes_i |x_i\rangle^{\otimes t_i}) |q\rangle \left( \otimes_{i=1}^{\lfloor \ell/2S \rfloor} |C_{2i-1} \wedge C_{2i}\rangle \right) |q_t \oplus f(x)\rangle \\ \rightarrow |x\rangle |0\rangle |q\rangle |0\rangle |q_t \oplus f(x)\rangle. \end{aligned}$$

We repeat the first two stages, and the depth of the reset stage is  $O(k \log \ell)$ .

The depth of inner-most recursion circuit is  $O(k \log \ell)$ . In the inner-most recursion,  $\ell/S$  clauses can be synthesized in parallel. We use  $\text{Depth}'_\ell(f)$  to denote the depth of the circuit, obtained by the algorithm described in this section, to synthesize  $f$  with  $\ell$  ancillary qubits.  $\text{Depth}'_\ell(\text{CNF}_{n, \ell/S}^k) = O(k \log \ell)$ . Combine with the recurrence formula in the previous section:  $\text{Depth}'_\ell(\text{CNF}_{n, m}^k) = \frac{2\ell}{(S+1)} \text{Depth}'_\ell(\text{CNF}_{n, 2m(S+1)/\ell}^k)$ . Then we have:

$$\text{Depth}'_\ell(\text{CNF}_{n, m}^k) = O \left( k \log \ell \left( \frac{mS}{\ell} \right)^{1+\log_{\ell/S} 4} \right).$$



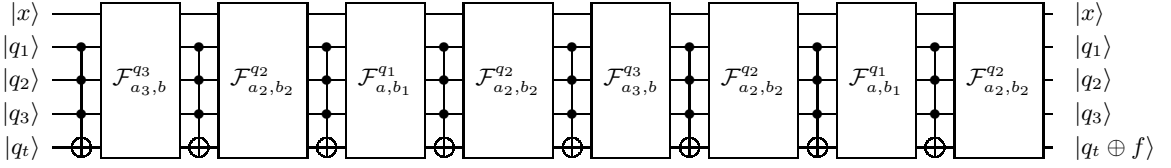


FIG. 2: When the number of ancillary qubits is small, another synthesis circuit for the CNF formula. Here  $f \in \text{CNF}_{n,(b-a)}^k$ .

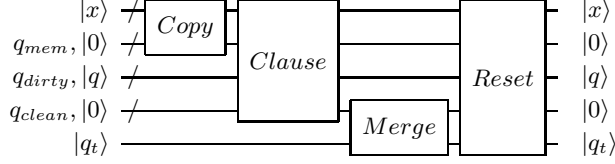


FIG. 3: The inner-most recursion of algorithm. The ancillary qubits are divide into 3 registers:  $q_{mem}$ ,  $q_{dirty}$ ,  $q_{clean}$ , where the size of these registers are  $\frac{(S-1)\ell}{S+1}$ ,  $\frac{\ell}{S+1}$ ,  $\frac{\ell}{S+1}$ , respectively.

### Circuit lower bound for CNF synthesis

We will prove that there exists a  $k$ -CNF with  $m$  clauses which need  $\Omega(km)$  size of quantum circuits to approximate it with any error  $\varepsilon < \frac{\sqrt{2}}{2}$ , as depicted in Theorem 1. To obtain this lower bound, let us first give a lower bound for the number of different  $\text{CNF}_{n,m}^k$ .

**Lemma 3.** *There are  $\Omega\left(\binom{n}{m}\right)$  different instances for  $\text{CNF}_{n,m}^k$ .*

*Proof.* Note that a CNF formula  $\phi : \{T, F\}^n \rightarrow \{T, F\}$  can be uniquely represented as a Boolean function  $f_\phi : \{0, 1\}^n \rightarrow \{0, 1\}$ . With a little abuse of symbols, we use the same symbol to represent the input of CNF formula and the corresponding Boolean functions. Let  $\phi = (v_1 \vee \dots \vee v_k) \wedge \dots \wedge (v_{(k-1)m+1} \vee \dots \vee v_{km})$  be a  $k$ -CNF formula, where  $v_i \in \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n\}$ , then it can be represented as a Boolean function

$$f_\phi(x) = \left(1 - \prod_{j=1}^k \bar{v}_j\right) \cdots \left(1 - \prod_{j=1}^k \bar{v}_{(k-1)m+j}\right), \quad (2)$$

where  $\bar{v}_i = 0$  iff  $v_i = T$  and  $\bar{v}_i = 1$  iff  $v_i = F$ , and the input  $x_i \in \{0, 1\}$  of  $f_\phi$  is associated with  $x_i \in \{T, F\}$  of  $\phi$ . Let  $f(x), g(x)$  be two functions (formulas), we say  $f \equiv g$  if  $f(x_1, \dots, x_n) = g(x_1, \dots, x_n)$  for any legal input  $(x_1, \dots, x_n)$ . Let

$$\mathcal{L} := \{v_1 \vee \dots \vee v_k | v_i \in \{\neg x_1, \dots, \neg x_n\}, v_i \neq v_j \text{ for } i, j \in [k]\}$$

be the set of all clauses with  $k$ -variables, where  $x_j \in \{T, F\}$ , and  $\neg x_j$  are the negations of  $x_j$ . Let a set of  $k$ -CNF formulas be

$$\mathcal{A} := \{\phi | \phi = l_1 \wedge \dots \wedge l_m, l_i \in \mathcal{L}, l_i \neq l_j, \text{ for } i \neq j \in [m]\}.$$

We would like to show the size of  $\mathcal{A}$  is  $\text{size}(\mathcal{A}) = \binom{n}{m}$ . i.e., any two different formulas  $\phi_1 = l_1 \wedge \dots \wedge l_m, \phi_2 = l'_1 \wedge \dots \wedge l'_m$

such that there exists  $j \in [m], l_j \neq l'_j$ , we have  $\phi_1 \neq \phi_2$ . By contradiction, suppose  $\phi_1 \equiv \phi_2$ , then  $f_{\phi_1}(x) \equiv f_{\phi_2}(x)$ . Let  $g_\phi(x)$  satisfies  $\deg(g) = k$  be the summations of all degree- $k$  terms of  $f_\phi(x)$ . By the definition of  $f_\phi(x)$  in Equation (2) and the fact that  $\bar{v}_j^2 = \bar{v}_j$ ,

$$g_\phi(x) = - \sum_{j=0}^{m-1} \bar{v}_{jk+1} \cdots \bar{v}_{j(k+1)}.$$

Since  $f_{\phi_1}(x) \equiv f_{\phi_2}(x)$ , then  $\deg(f_{\phi_1} - f_{\phi_2}) = 0$ , i.e.,  $g_{\phi_1}(x) \equiv g_{\phi_2}(x)$ . Let  $\bar{v}_1 \cdots \bar{v}_k$  be one term of  $g_{\phi_1}(x)$ . For a given input  $x = (x_1, \dots, x_n)$  such that  $x_i = 1$  when  $x_i \in \{\bar{v}_1, \dots, \bar{v}_k\}$  and  $x_i = 0$  otherwise. It is easy to check  $g_{\phi_2}(x) = 1$  iff  $x_1 \cdots x_k$  is one term of the function  $g_\phi(x)$ . Hence  $\bar{v}_1 \cdots \bar{v}_k$  is also a term of  $g_{\phi_2}(x)$ . Without loss of generality, each term  $\bar{l}_j = \bar{v}_{jk+1} \cdots \bar{v}_{j(k+1)} \in g_{\phi_1}, \bar{l}_j \in g_{\phi_2}$  at the same time. Therefore  $\phi_1 = \phi_2$ , contrary with the fact that  $\phi_1, \phi_2$  are two different formulas in  $\mathcal{A}$ .  $\square$

**Theorem 1.** *There exists a  $\text{CNF}_{n,m}^k$ , any quantum circuits approximating it with error  $\varepsilon < \frac{\sqrt{2}}{2}$  needs size  $\Omega(km)$ .*

*Proof.* Let  $U \in \mathbb{C}^{4 \times 4}$  be a two qubit gate, and the  $\delta$ -discretization of the  $(j, k)$ -th element  $U_{jk}$  be  $U_{jk}^\delta = \delta[a/\delta] + i\delta[b/\delta]$ , where  $U_{jk} = a + ib$ . Then we have  $\|U - U^\delta\|_2 < 2\delta$ . There are at most  $\left(\frac{2}{\delta}\right)^{32}$  different  $\delta$ -discretizations  $U^\delta$  for the infinite continuous  $U$  in the space by its definition.

In the following, we prove that any two different instances in  $\mathcal{A}$  do not share any common  $\delta$ -discretization. Hence, we can use the counting method to give a lower bound of the circuit size.

Let  $A_G, A_H$  be the quantum circuit representations of two different instances in  $\mathcal{A}$ . Let  $s$  be the maximum size of all the unitaries related to  $A_G$  and  $A_H$ . By the fact that the unitary  $U \in \mathbb{C}^{2^n \times 2^n}$  has a  $s$ -size quantum circuit, the following

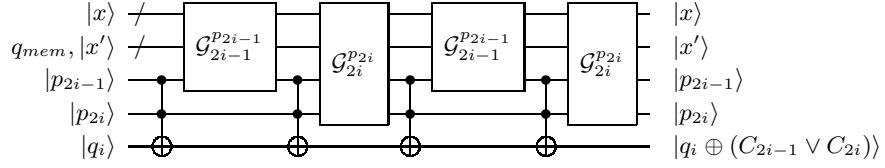


FIG. 4: The Clause stage in the inner-most recursion of algorithm. The  $q_{dirty}$  and  $q_{clean}$  registers are divided into  $\ell/2S$  parts. Here we show the  $i$ -th part of Clause stage, where we synthesize a two clauses function. Different part of Clause is running in parallel. Operator  $\mathcal{G}_j^{q_t}$  generate a clause  $C_j$  on the target qubit  $q_t$ ,  $\mathcal{G}_j^{q_t} |x\rangle |q_t\rangle \rightarrow |x\rangle |q_t \oplus C_j(x)\rangle$ .

inequalities

$$\begin{aligned} \|A_G - A_G^\delta\| &< 2s\delta \leq \varepsilon, \\ \|A_H - A_H^\delta\| &< 2s\delta \leq \varepsilon, \end{aligned}$$

hold when  $\delta = \frac{\varepsilon}{2s}$ , and  $\varepsilon < \frac{\sqrt{2}}{2}$ . Combined with the fact that  $\|A_G - A_H\| = \sqrt{2}$ , we have  $A_G^\delta \neq A_H^\delta$ .

Hence, any two different instances in  $\mathcal{A}$  have different  $\delta$ -discretization. There are  $\Omega\left(\left(\frac{n}{m}\right)\right)$  different instances for  $k$ -CNF with  $m$ -clause by Lemma 3 in main file. By the fact that the number of different instances is upper bounded by the number of  $\delta$ -discretization of quantum circuits,

$$\left(\frac{n}{m}\right) \leq \left(\left(\frac{2}{\delta}\right)^{32} \cdot n\right)^s. \quad (3)$$

Since  $\left(\frac{n}{k}\right) = \Omega((n/k)^k)$  for any  $k$ . Then

$$\left(\frac{n}{m}\right) = \Omega\left(\left(\frac{n}{m}\right)^m\right) = \Omega\left(\left(\frac{n^k}{k^k m}\right)^m\right).$$

By inequality 3, we have  $s = \Omega(km)$  when  $k = o(n)$ . When  $k = cn$  for constant  $c < 1$ , by Stirling's formula,  $\left(\frac{n}{k}\right) = \Omega(2^{an})$  for some constant  $a < 1$ . Hence,

$$\left(\frac{n}{m}\right) = \Omega\left(\left(\frac{2^{an}}{m}\right)^m\right),$$

combined with inequality (3) give us  $s = \Omega(km)$  when  $k = cn$  for constant  $c < 1$ . This implies the lower bound also holds for any  $k < n$  for general  $k$ -CNF.  $\square$