Estimating the cost of generic quantum pre-image attacks on SHA-2 and SHA-3

Matthew Amy,[1,2] Olivia Di Matteo,[1,3] Vlad Gheorghiu,[1,4,*] Michele Mosca,[1,4,5,6] Alex Parent,[1,3] and John Schanck[1,4]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*
[2]*David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*
[3]*Department of Physics & Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*
[4]*Department of Combinatorics & Optimization, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*
[5]*Perimeter Institute for Theoretical Physics, Waterloo, ON, N2L 6B9, Canada*
[6]*Canadian Institute for Advanced Research, Toronto, ON, M5G 1Z8, Canada*

We investigate the cost of Grover's quantum search algorithm when used in the context of pre-image attacks on the SHA-2 and SHA-3 families of hash functions. Our cost model assumes that the attack is run on a surface code based fault-tolerant quantum computer. Our estimates rely on a time-area metric that costs the number of logical qubits times the depth of the circuit in units of surface code cycles. As a surface code cycle involves a significant classical processing stage, our cost estimates allow for crude, but direct, comparisons of classical and quantum algorithms.

We exhibit a circuit for a pre-image attack on SHA-256 that is approximately 2^{149} surface code cycles deep and requires approximately 2^{13} logical qubits. This yields an overall cost of 2^{162} logical-qubit-cycles. Likewise we exhibit a SHA3-256 circuit that is approximately 2^{146} surface code cycles deep and requires approximately 2^{16} logical qubits for a total cost of, again, 2^{162} logical-qubit-cycles. Both attacks require on the order of 2^{128} queries in a quantum black-box model, hence our results suggest that executing these attacks may be as much as 17 billion times more expensive than one would expect from the simple query analysis. See arXiv:1603.09383 [quant-ph] for a full version of this work.

I. INTRODUCTION

A number of quantum algorithms promise significant asymptotic speedups compared with their classical counterparts [1–3]. While most fields of research will be unaffected by these algorithms until large quantum computers are built, cryptography is affected by the possibility of these algorithms being run at any time in the future. The hardness assumptions underlying the public key cryptosystems currently in use – those related to factoring and variants of the discrete logarithm problem – are violated by quantum adversaries. Quantum Fourier sampling techniques break these cryptosystems in polynomial time [1, 4]. As a result these cryptosystems can no longer be considered secure, and ultimately they will have to be replaced. Some standards bodies have already begun discussions about transitioning to new public key cryptographic primitives [5, 6].

Symmetric primitives, by contrast, are weakened but not necessarily broken by quantum algorithms. The best generic attacks on symmetric primitives apply Grover's quantum search algorithm and achieve a corresponding quadratic improvement over exhaustive search in a black-box query model [2, 7, 8]. Such attacks are not formally efficient, but they do require a re-evaluation of the concrete security of symmetric primitives.

A conservative defense against attacks based on Grover's algorithm is to compensate for the potential square root loss in security by doubling the security parameter. This may mean doubling the key size for a cipher, or doubling the output length for a hash function. This is a suitable response for the cryptographer who wants to make worst case assumptions about the potential power of quantum computers. Others, however, may want to know either 1) the exact cost of an attack based on Grover's algorithm for a particular parameterization of a cryptosystem, or 2) the minimal security parameter that provides "adequate protection" in the sense of [9–11].

Estimating either of these quantities requires close analysis of the cost of a realistic implementation of Grover's algorithm. Overhead is introduced at the logical level by the reversibility constraint on quantum computations and by the structure of the Grover iteration itself. Additional overhead may be introduced by fault-tolerance mechanisms required by a particular model of quantum computation.

To better understand these issues, we present an estimate of the cost of performing pre-image attacks on the SHA-2 and SHA-3 families of hash functions. A similar analysis has been performed recently for AES [12]. We execute the following procedure for each hash function. First, we implement the function as a reversible circuit. We then use a quantum circuit optimization tool, T-par [13], to minimize the circuit's T-count and T-depth. This is necessary because T gates are expensive in our chosen model of quantum computation. With the optimized circuit in hand we estimate the cost of executing Grover's algorithm on a surface code based quantum computer. A similar analysis was performed by Fowler et al. [14] to estimate the physical resources required for part of Shor's factoring algorithm. Our resource estimates focus on the number of logical qubits in the fault-tolerant circuit and the overall depth of the circuit in units of surface code cycles. Each surface code cycle involves the execution of a classical syndrome decoding routine for every logical qubit. Thus in estimating these quantities we obtain the cost of a pre-image attack purely in terms of classical computing resources. Separately, we obtain an estimate for the number of physical qubits required for the circuit, and an estimate for the wall-clock time of the computation.

* Electronic address: vlad.gheorghiu@uwaterloo.ca

Our resource estimation methodology takes into account several of the layers between the high level description of an algorithm and the physical hardware required for its execution. Our approach is modular should assumptions about any of these layers change, and hence it allows one to calculate the impact of improvements in any particular layer. We illustrate our method schematically in Fig. 1 and Fig. 2.

II. INTRINSIC COST OF GROVER SEARCHING

Suppose there is polynomial overhead per Grover iteration, i.e. $\Theta(2^{k/2})$ Grover iterations cost $\approx k^a 2^{k/2}$ logical qubit cycles for some real a independent of k. Then an adversary who is willing to execute an algorithm of cost 2^C can use Grover's algorithm to search a space of k bits provided that

$$k/2 + a\log_2(k) \leq C. \tag{1}$$

We define the *overhead* of the circuit as a and the *advantage* of the circuit as k/C. Note that if we view k as a function of a and C then for any fixed a we have

$$\lim_{C\to\infty} k(a, C)/C = 2,$$

i.e. asymptotically, Grover's algorithm provides a quadratic advantage over classical search. Here we are interested in non-asymptotic advantages.

If $a = 0$ then each Grover iteration has cost 1 and $k = 2C$. This represents the upper bound for the advantage that can be attained by circuits based on Grover's algorithm. However, as we use a time-area metric for cost, and Grover's algorithm requires at least k logical qubits, a more realistic upper bound on the advantage is provided by taking $a = 1$.

Even $a = 1$ assumes that the Grover iteration is of constant depth, however it requires at least one k-fold controlled-NOT gate. A single k-fold controlled-NOT, implemented as a circuit over Clifford+T gates, has depth $\approx \log_2 k$ [15]. If we assume that the k-fold controlled-NOT dominates the temporal overhead of the iteration this yields $a \geq 1.375$ for $k \leq 256$. This still neglects some spatial overhead required for magic state distillation, but $a = 1.375$ may be used to derive strict upper bounds, in our cost model, for the advantage of Grover search.

As an example, the AES-256 circuit from [12] has depth 130929 and requires 1336 logical qubits. This yields overhead of $a \approx 3.423$ from the reversible layer alone. In our work, we determine a via a rigorous cost analysis of a full fault-tolerant implementation of Grover's algorithm for pre-image attacks on the SHA family of cryptographic hash functions. Before summarizing our results in Table I, we state our main assumptions.

III. ASSUMPTIONS

We make the following assumptions when performing our analysis.

	SHA-256	**SHA3-256**
T-count	1.27×10^{44}	2.71×10^{44}
T-depth	3.76×10^{43}	2.31×10^{41}
Logical qubits (circuit)	2402	3200
Surface code distance	43	44
Physical qubits	1.39×10^7	1.94×10^7
Logical qubits (distillation)	3615	3615
Surface code distances	$\{33, 13, 7\}$	$\{33, 13, 7\}$
Magic state factories	1	13
Physical qubits	1.23×10^7	1.60×10^8
Surface code cycles	2^{149}	2^{146}
Total cost	2^{162}	2^{162}

TABLE I. Fault-tolerant resource counts for Grover search of SHA-256 and SHA3-256.

Assumption 1. *The resources required for any large quantum computation are well approximated by the resources required for that computation on a surface code based quantum computer.*

Assumption 2. *The classical error correction routine for the surface code on an $L \times L$ grid of logical qubits requires an $L \times L$ mesh of classical processors (i.e. $C_a = n$).*

Assumption 3. *Each classical processor performs a constant number of operations per surface code cycle.*

Assumption 4. *The temporal cost of one surface code cycle is equal to the temporal cost of one hash function invocation.*

Combining the assumptions above we arrive at the following metric for comparing the costs of classical and quantum computations.

Cost Metric 1. *The cost of a quantum computation involving ℓ logical qubits for a duration of σ surface code cycles is equal to the cost of classically evaluating a hash function $\ell \cdot \sigma$ times. Equivalently we will say that* one logical qubit cycle *is equivalent to* one hash function invocation.

IV. RESULTS AND CONCLUSIONS

We estimated the cost of a quantum pre-image attack on SHA-256 and SHA3-256 cryptographic hash functions via Grover's quantum searching algorithm. We constructed reversible implementations of both SHA-256 and SHA3-256 cryptographic hash functions, for which we then optimized their corresponding T-count and depth. We then estimated the required physical resources needed to run a brute force Grover search on a fault-tolerant surface code based architecture.

We showed that attacking SHA-256 requires approximately 2^{149} surface code cycles and that attacking SHA3-256 requires approximately 2^{146} surface code cycles. For both SHA-256 and SHA3-256 we found that the total cost when including the classical processing increases to approximately 2^{162} basic operations.

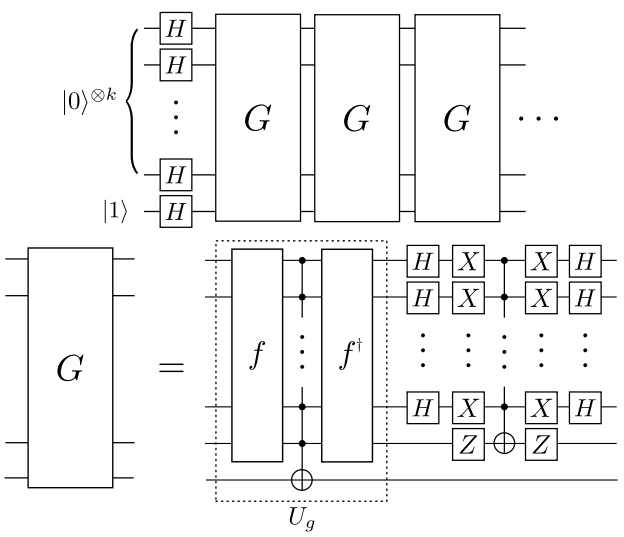

FIG. 1. Grover searching with an oracle for $f : \{0,1\}^k \to \{0,1\}^k$.

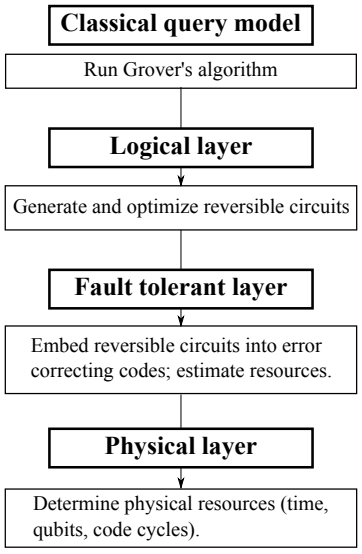

FIG. 2. Analyzing Grover's algorithm (mini-flowchart).

Our estimates are by no means a lower bound, as they are based on a series of assumptions. First, we optimized our T-count by optimizing each component of the SHA oracle individually, which of course is not optimal. Dedicated optimization schemes may achieve better results. Second, we considered a surface code fault-tolerant implementation, as such a scheme looks the most promising at present. However it may be the case that other quantum error correcting schemes perform better. Finally, we considered an optimistic per-gate error rate of about 10^{-5}, which is the limit of current quantum hardware. This number will probably be improved in the future. Improving any of the issues listed above will certainly result in a better estimate and a lower number of operations, however the decrease in the number of bits of security will likely be limited .

[1] Peter W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM Journal on Computing **26**, 1484–1509 (1997).

[2] Lov K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," Phys. Rev. Lett. **79**, 325–328 (1997).

[3] Stephen Jordan, "Quantum Algorithm Zoo," (2016).

[4] Dan Boneh and Richard J. Lipton, "Quantum Cryptanalysis of Hidden Linear Functions," in *Advances in Cryptology - CRYPTO'95*, Lecture Notes in Computer Science No. 963, edited by Don Coppersmith (Springer Berlin Heidelberg, 1995) pp. 424–437.

[5] United States National Security Agency, "NSA Suite B Cryptography - NSA/CSS," NSA website (2015).

[6] Lily Chen, Stephen Jordan, Yi-Kai Liu, Dustin Moody, Rene Peralta, Ray Perlner, and Daniel Smith-Tone, "Report on post-quantum cryptography," National Institute of Standards and Technology Internal Report 8105 (2016).

[7] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp, "Tight bounds on quantum searching," Fortschritte der Physik **46**, 493–505 (1998).

[8] Brassard Gilles, Hoyer Peter, Mosca Michele, and Tapp Alain, "Quantum amplitude amplification and estimation," Quantum Computation and Quantum Information, Samuel J. Lomonaco, Jr. (editor), AMS Contemporary Mathematics , 53–74 (2002), e-print arXiv:quant-ph/0005055.

[9] Arjen K. Lenstra, "Handbook of information security," (Wiley, 2004) Chap. Key Lengths.

[10] Arjen K. Lenstra and Eric R. Verheul, "Selecting cryptographic key sizes," Journal of Cryptology **14**, 255–293 (2001).

[11] M. Blaze, W. Diffie, R. Rivest, B. Schneier, T. Shimomura, E. Thompson, and M. Weiner, *Minimal Key Lengths for Symmetric Ciphers to Provide Adequate Commercial Security*, Tech. Rep. (An ad hoc group of cryptographers and computer scientists, 1996).

[12] M. Grassl, B. Langenberg, M. Roetteler, and R. Steinwandt, "Applying Grover's algorithm to AES: quantum resource estimates," E-print arXiv:1512.04965 [quant-ph].

[13] M. Amy, D. Maslov, and M. Mosca, "Polynomial-time t-depth optimization of Clifford+T circuits via matroid partitioning," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on **33**, 1476–1489 (2014).

[14] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland, "Surface codes: Towards practical large-scale quantum computation," Phys. Rev. A **86**, 032324 (2012).

[15] Peter Selinger, "Quantum circuits of T-depth one," Phys. Rev. A **87**, 042302 (2013), arXiv:1210.0974.

ACKNOWLEDGMENTS

We acknowledge support from NSERC and CIFAR. IQC and PI are supported in part by the Government of Canada and the Province of Ontario.