Tight Bounds for Simon's Algorithm

Xavier Bonnetain

Institute for Quantum Computing, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON, Canada

Abstract. Simon's algorithm is the first example of a quantum algorithm exponentially faster than any classical algorithm, and has many applications in cryptanalysis. While these quantum attacks are often extremely efficient, they are generally missing some precise cost estimate. This article aims at resolving this issue by presenting precise query estimates for the different use cases of Simon's algorithm in cryptanalysis, and shows that Simon's algorithm requires in most cases little more than n queries to succeed.

Keywords: Simon's algorithm, quantum cryptanalysis, complexity analysis

1 Introduction

Simon's algorithm [Sim94] is the first example of a quantum algorithm exponentially faster than any classical ones, and lead the way to Shor's algorithm [Sho94]. Nevertheless, for a long time, there wasn't any concrete application for this algorithm.

This began to change in 2010, with a polynomial-time quantum distinguisher on the 3-round Feistel cipher [KM10], a construction that is classically provably secure. This first application opened the way to many other attacks in symmetric cryptography [KM12; RS15; Kap+16; SS17; Bon17; HA17; LM17; DLW19; BNS19; ND19; II19; Ito+19; Bon19; HRK20]. All these attacks have an important restriction: they only fit in the *quantum query model*, that is, they require access to a quantum circuit that can compute the corresponding construction *including its secret material.*

This last restriction was overcomed in 2019, with the offline Simon's algorithm [Bon+19] that allows to apply some of the previous attacks when we only have access to classical queries to the secret function, albeit only with a polynomial gain.

We can separate the previous attacks in two families:

- quantum distinguishers, in which we want to identify wether an oracle-accessible function has a specific structure or has been chosen uniformly at random,
- key recoveries, in which we want to recover some secret information used in an oracle-accessible function.

The main difference between the two types of attacks is that in the case of key recoveries, the knowledge of the secret generally allows to compute the oracle-accessible function, which makes it easier to check if a result is correct.

In the literature, the query cost of the attack is often left as a $\mathcal{O}(n)$. Very few concrete estimates are proposed, and they tend to be loose estimates. Hence, in the current situation, we have asymptotically efficient attacks, but their effiency in practice is less clear.

We focus here on giving precise bounds on the number of queries required by Simon's algorithm. Note that we do not give here concrete lower bounds for Simon's *problem*. To our knowledge, the sole work in this direction is [KNP07], where a lower bound of $n/8$ queries is proven.

Previous works. The first concrete estimate comes from [Kap+16, Theorem 1 and 2], where it is shown that cn queries, for $c > 3$ that depend on the function, is enough to have a success probability exponentially close to 1. In [LM17, Theorem 2], a bound of $2(n + \sqrt{n})$ queries for the Grover-meets-Simon algorithm with a perfect external test is shown, for a success probability greater than 0.4, and assuming that the periodic function has been sampled uniformly at random. In [Bon17], a heuristic cost of $1.2n + 1$ for Simon's algorithm is used. In the recent [MS19, Theorem 3.2], it is shown that for 2-to-1 functions, Simon's algorithm needs on average less than $n + 1$ queries, and that for functions with 1 bit of output, on average less than $2(n + 1)$ queries are needed, if each query uses a function sampled independently uniformly over the set of functions with the same period.

Contributions. In this article, we propose some precise concrete cost estimates for Simon's algorithm, in the ideal case of periodic permutations and in the case of random functions, and without any constraint on the subgroup size. We show that in most cases, the overhead induced by the use of random periodic functions is less than 1 query. We also give the first concrete estimates for a variant of the exact algorithm proposed in [BH97] and for the offline Simon's algorithm [Bon+19]. Finally, we show some improved estimates for the Grover-meet-Simon algorithm, and we prove that the overhead induced by a lack of external test is in fact extremely small.

These estimates allow us to propose the following heuristics on Simon's algorithm, the Grover-meet-Simon algorithm and the Offline Simon's algorithm. These heuristics are respectively supported by Theorems 8 and 9, Corollary 3 and Corollary 4, and Corollary 5.

Heuristic 1 (Simon's algorithm in cryptography) *For all cryptographic use cases, Simon's algorithm succeeds in $n + 3$ queries on average and with $n + \alpha$ queries, it succeeds with probability $1 - 2^{-\alpha}$.*

Heuristic 2 (Grover-meet-Simon in cryptography) *For key recoveries, Grover-meet-Simon with a perfect external test succeeds with probability greater than one half in $n + 2$ queries per test, plus the number of queries of the external test.*

For all cryptographic use cases, Grover-meet-Simon with a periodicity test succeeds with probability greater than one half in $n + 2 + 2\left\lceil \frac{k}{n} \right\rceil$ queries per test.

Heuristic 3 (The offline Simon's algorithm in cryptography) *For all cryptographic use cases, the offline Simon's algorithm succeeds with probability greater than one half in $n + k + 5$ queries per test.*

2 Quantum algorithms

This section presents the quantum algorithms we will study, that is, Simon's algorithm and its different uses in a quantum search.

2.1 Simon's algorithm

Simon's algorithm [Sim94] tackles the Hidden Subgroup problem when the group is $\{0,1\}^n$.

We can formulate the problem as follows:

Problem 1 (Simon's Problem). Let n be an integer, $\mathcal{H}$ a subgroup of $\{0,1\}^n$ and X a set. Let $f : \{0,1\}^n \to X$ be a function with the promise that for all $(x, y) \in (\{0,1\}^n)^2$, $[f(x) = f(y) \Leftrightarrow x \oplus y \in \mathcal{H}]$. Given oracle access to f, find a basis of $\mathcal{H}$.

The promise in the problem can also be relaxed:

Problem 2 (Relaxed Simon's Problem). Let n be an integer, $\mathcal{H}$ a subgroup of $\{0,1\}^n$ and X a set. Let $f : \{0,1\}^n \to X$ be a function with the promise that for all $(x, y) \in (\{0,1\}^n)^2$, $[x \oplus y \in \mathcal{H} \Rightarrow f(x) = f(y)]$ and that for all $h \notin \mathcal{H}$, there exists an x such that $f(x) \neq f(x \oplus h)$. Given oracle access to f, find a basis of $\mathcal{H}$.

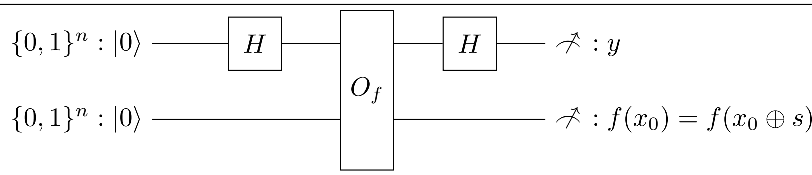
We consider two types of functions, depending on the problem we need to solve:

Definition 1 (Periodic permutations, periodic functions). *We call a* periodic permutation *a function f that fulfills the promise of Problem 1, that is, is constant over the cosets of $\mathcal{H}$ and is injective over $\{0,1\}^n/(\mathcal{H})$. If f is constant over the cosets of $\mathcal{H}$, but not necessarily injective over $\{0,1\}^n/(\mathcal{H})$, we say that f is a periodic function.*

Remark 1 (Aperiodic function). By the previous definition, an aperiodic function (resp. permutation) is periodic over the trivial subgroup.

Note that Problem 1 tackles periodic permutations while Problem 2 tackles periodic functions.

Circuit 1 Simon's circuit



Algorithm 1 Simon's routine

Input: n, $O_f : |x\rangle |0\rangle \mapsto |x\rangle |f(x)\rangle$ with $f : \{0,1\}^n \to X$ of hidden subgroup $\mathcal{H}$

Output: y with $y \cdot s = 0$

1: Initialize two n-bits registers : $|0\rangle |0\rangle$

2: Apply H gates on the first register, to compute $\sum_{x=0}^{2^n-1} |x\rangle |0\rangle$

3: Apply O_f, to compute $\sum_{x=0}^{2^n-1} |x\rangle |f(x)\rangle$

4: Reapply H gates on the register, to compute

$$\sum_x \sum_{j=0}^{2^n-1} (-1)^{x \cdot j} |j\rangle |f(x)\rangle$$

5: The register is in the state

$$\sum_{x_0 \in \{0,1\}^n/(\mathcal{H})} \sum_{x_1 \in \mathcal{H}} \sum_{j=0}^{2^n-1} (-1)^{(x_0 \oplus x_1) \cdot j} |j\rangle |f(x_0)\rangle$$

6: Measure $j, f(x_0)$, return them.

Algorithm description The quantum circuit of Simon's algorithm is presented in Circuit 1. It produces a random value orthogonal to $\mathcal{H}$, and can be described as Algorithm 1.

After measuring $f(x_0)$ the state becomes

$$\sum_{j=0}^{2^n-1} (-1)^{x_0 \cdot j} \sum_{x_1 \in \mathcal{H}} (-1)^{x_1 \cdot j} |j\rangle .$$

We now use the following lemma, proven in appendix:

Lemma 1 (Adapted from [Kap+16, Lemma 1]). *Let $\mathcal{H}$ be a subgroup of $\{0,1\}^n$. Let $y \in \{0,1\}^n$. Then*

$$\sum_{h \in \mathcal{H}} (-1)^{y \cdot h} = \begin{cases} |\mathcal{H}| & \text{if } y \in \mathcal{H}^\perp \\ 0 & \text{otherwise} \end{cases}$$

Hence, the amplitude for a j is nonzero if and only if $j \in \mathcal{H}^\perp$. Hence, this routine samples uniformly a value j orthogonal to $\mathcal{H}$.

The complete algorithm calls the routine until the values span a space of maximal rank or, if the rank is unknown, a fixed T times. In practice, we'll see in the next sections that $T = n + \mathcal{O}(1)$ is sufficient to succeed.

Algorithm 2 Simon's algorithm [Sim94]

Input: n, $O_f : |x\rangle |0\rangle \mapsto |x\rangle |f(x)\rangle$ with $f : \{0,1\}^n \to X$ of hidden subgroup $\mathcal{H}$, T
Output: a basis of $\mathcal{H}$

```
V = ∅
for i from 1 to T do
    Get y, f(x_0) from Algorithm 1
    Add y to V
end for
return a basis of V^⊥
```

2.2 Amplitude amplification

We will use amplitude amplification [Bra+02] with Simon's algorithm as a test function in the following sections. We first recall some standard lemmas, and then present a variant for the case where we do not have access to the test function, but only to an approximation.

Lemma 2 (Amplitude amplification[Bra+02]). *Let C be a quantum circuit such that $C|0\rangle = \sqrt{p}\,|Good\rangle + \sqrt{1-p}\,|Bad\rangle$,*

Let O_g be an operator that fulfills $O_g|Good\rangle = -|Good\rangle$ and $O_g|Bad\rangle = |Bad\rangle$, and O_0 the operator $I - 2|0\rangle\langle 0|$, let $\theta = \arcsin(\sqrt{p})$. Then $(CO_0C^\dagger O_g)^t C|0\rangle = \sin((2t+1)\theta)|Good\rangle + \cos((2t+1)\theta)|Bad\rangle$.

Lemma 3 (Exact amplitude amplification for $p = 1/2$ [BH97]). *Let C be a quantum circuit such that $C|0\rangle = \frac{1}{\sqrt{2}}|Good\rangle + \frac{1}{\sqrt{2}}|Bad\rangle$,*

Let S_g be an operator that fulfills $S_g|Good\rangle = i|Good\rangle$ and $S_g|Bad\rangle = |Bad\rangle$, and O_0 the operator $I - (1-i)|0\rangle\langle 0|$.

Then $CS_0C^\dagger S_g C|0\rangle = |Good\rangle$.

Remark 2. The operators O_g and S_g can be easily implemented given a quantum circuit T that computes 1 for the good elements and 0 for the bad ones.

In practice, we may have a bound on the success probability instead of the exact value. The following theorem tackles this case.

Theorem 1 (Amplitude amplification with unprecise success probability). *Let C be a quantum circuit such that $C|0\rangle = \sqrt{p}\,|Good\rangle + \sqrt{1-p}\,|Bad\rangle$, If $p \in [(1-\beta)p_0, p_0]$, then after $\frac{\pi}{4\arcsin\sqrt{p_0}}$ iterations, the probability of measuring an element in $|Good\rangle$ is at least $1 - (2\beta + 2p_0 + \sqrt{p_0})^2$. If $p \in [p_0, p_0(1+\beta)]$, then* *after* $\frac{\pi}{4 \arcsin \sqrt{p_0}}$ *iterations, the probability of measuring an element in* $|Good\rangle$ *is at least* $1 - (\beta + \sqrt{(1+\beta)p_0} + 2\sqrt{1+\beta}^3 p_0)^2$.

Proof. After $\frac{\pi}{4 \arcsin \sqrt{p_0}}$, the amplitude of $|\text{Bad}\rangle$ is

$$\cos\left(\frac{\pi \arcsin \sqrt{p}}{2 \arcsin \sqrt{p_0}} + \arcsin \sqrt{p}\right)$$

This is equal to

$$\sin\left(\frac{\pi}{2}\left(1 - \frac{\arcsin \sqrt{p}}{\arcsin \sqrt{p_0}}\right) - \arcsin \sqrt{p}\right)$$

Now, we need to bound the norm of this expression. As $|\sin(x)| \leq |x|$, this is bounded by

$$\left|\frac{\pi}{2}\left(1 - \frac{\arcsin \sqrt{p}}{\arcsin \sqrt{p_0}}\right)\right| + \arcsin \sqrt{p}$$

Now, if $(1-\beta)p_0 \leq p \leq p_0$, as $x \leq \arcsin x \leq x + \frac{\pi^3}{8}\frac{x^3}{6}$, this is bounded by

$$\frac{\pi}{2} \frac{\sqrt{p_0} + \frac{\pi^3}{8}\frac{\sqrt{p_0}^3}{6} - \sqrt{(1-\beta)p_0}}{\sqrt{p_0}} + \sqrt{p_0} + \frac{\pi^3}{8}\frac{\sqrt{p_0^3}}{6}$$

As $\sqrt{1-\beta} \geq 1 - \beta$, this is bounded by

$$\frac{\pi}{2}\beta + \frac{\pi^4}{16}\frac{p_0}{6} + \sqrt{p_0} + \frac{\pi^3}{8}\frac{\sqrt{p_0^3}}{6}$$

Hence as $p_0 \leq 1$, this is bounded by $2\beta + 2p_0 + \sqrt{p_0}$.

Finally, if $p_0 \leq p \leq (1+\beta)p_0$, this is bounded by

$$\frac{\pi}{2} \frac{\sqrt{(1+\beta)p_0} + \frac{\pi^3}{8}\frac{\sqrt{(1+\beta)p_0}^3}{6} - \sqrt{p_0}}{\sqrt{p_0}} + \sqrt{(1+\beta)p_0} + \frac{\pi^3}{8}\frac{\sqrt{(1+\beta)p_0}^3}{6}$$

As $\sqrt{1+\beta} \leq 1 + \frac{\beta}{2}$, this is bounded by

$$\frac{\pi}{4}\beta + \sqrt{(1+\beta)p_0} + \frac{\pi^4}{16}\frac{\sqrt{1+\beta}^3 p_0}{6} + \frac{\pi^3}{8}\frac{\sqrt{(1+\beta)p_0}^3}{6}$$

Hence, as $p_0 \leq 1$, this is bounded by $\beta + \sqrt{(1+\beta)p_0} + 2\sqrt{1+\beta}^3 p_0$. ◻

Now, for the offline Simon's algorithm, we will present a slight generalization, with an approximate test circuit.

Theorem 2 (Amplitude amplification with approximate test). *Let* S*,* f*,* C*,* O_g *and* θ *be defined as in Lemma 2, let* $\widehat{O_g}$ *be an approximation of* O_g*, such that for all* $|x\rangle$*,* $\widehat{O_g}\,|x\rangle = O_g\,|x\rangle + |\delta\rangle$*, with* $|\delta\rangle$ *an arbitrary vector such that* $|\,|\delta\rangle\,| \leq \varepsilon$*. Then after* t *iterations of amplitude amplification using* $\widehat{C}$ *instead of* C*, a measurement of the quantum state will give an element in* $|Good\rangle$ *with probability at least* $(\sin((2t+1)\theta) - t\varepsilon)^2$*.*

Proof. In the amplitude amplification procedure, a call to O_g is replaced by a call to $\widehat{O_g}$. Hence, each call adds a noise vector $|\delta\rangle$ to the state, and after t iterations, as the operators are linear, the total noise $|\psi_{err}\rangle$ is of amplitude at most $t\varepsilon$. Without any noise, after t iterations we would be in the state

$$|\psi_t\rangle = \sin((2t+1)\theta)\,|\text{Good}\rangle + \cos((2t+1)\theta)\,|\text{Bad}\rangle$$

but due to it, we are in the state $|\psi_t\rangle + |\psi_{err}\rangle$. We have

$$|\langle \text{Good}|\psi_{err}\rangle| \leq |\,|\text{Good}\rangle\,| \times |\,|\psi_{err}\rangle\,| \leq t\varepsilon.$$

Hence, the amplitude of $|\text{Good}\rangle$ is at least $\sin((2t+1)\theta) - t\varepsilon$, and the probability of obtaining a value in $|\text{Good}\rangle$ is greater than $(\sin((2t+1)\theta) - t\varepsilon)^2$. □

2.3 Exact version of Simon's algorithm

An exact version of Simon's algorithm was proposed by Brassard and Høyer in 1997 [BH97]. We propose here a more efficient variant of this algorithm, and prove that its query complexity is bounded by $3n - h + 1$ if the function operates on n bits and the hidden subgroup has dimension h. Note that this algorithm is only exact for periodic permutations.

The algorithm of Brassard and Høyer. The idea of Algorithm 3 is to ensure that any measurement we perform gives us some information about the period. This is done by doing exact amplitude amplification over a subset of the values outputted by Simon's algorithm. Moreover, in some cases, the subset we seek might be empty. In the original algorithm, the empty case meant that we should try with another subset until we find a non-empty one, or, if there is none, that the algorithm can end. As there is at most n such subsets, the algorithm is polynomial.

Our improved variant. We improve over the previous algorithm by remarking that the knowledge that a subset is empty actually gives some information on the hidden subgroup: it shows that a given vector is not in he subgroup's dual. Moreover, we show that this case is actually better, that is, we can reuse the quantum state, and save 1 query each time this occurs. This is Algorithm 4.

Theorem 3 (Complexity and correctness of Algorithm 4). *Let f be a periodic permutation from $\{0,1\}^n$ with a hidden subgroup of dimension h. Algorithm 4 returns a basis of $\mathcal{H}$ in less than* $\min(3n - h + 1, 3n)$ *queries.*

Proof. We will show that at each step of the for loop, either a linearly independent vector from $\mathcal{H}^\perp$ is added to V, or a linearly independent vector from $\{0,1\}^n/(\mathcal{H}^\perp)$ is added to W, and in the latter case, the quantum query can be reused in the next loop iteration. We also ensure that $V \subset \mathcal{H}^\perp$ and $W \subset \{0,1\}^n/(\mathcal{H})$.

Algorithm 3 Exact Simon's algorithm, from [BH97]

Input: n, $O_f : |x\rangle |0\rangle \mapsto |x\rangle |f(x)\rangle$ with $f : \{0,1\}^n \to X$ hiding $\mathcal{H}$
Output: a basis of $\mathcal{H}$

$V = \emptyset$ ▷ Basis of $\mathcal{H}^\perp$
for i from 1 to n **do**
  Choose a set W such that V, W forms a basis of $\{0,1\}^n$.
  for j from i to n **do**
    Apply Algorithm 1, without measuring, for the state

$$\sum_{x_0 \in \{0,1\}^n/(\mathcal{H})} \sum_{y=0}^{2^n-1} (-1)^{x_0 \cdot y} \sum_{x_1 \in \mathcal{H}} (-1)^{x_1 \cdot y} |y\rangle |f(x_0)\rangle$$

    amplify the state with the following test: ▷ Exact amplification
      Decompose y as $\sum_{v_k \in V} \delta_k v_k + \sum_{w_j \in W} \gamma_\ell w_\ell$
      return $\gamma_j = 1$
    end amplify
    Measure $|\gamma_j\rangle$
    if $\gamma_j = 1$ **then** measure the first register, add the result to V
      break
    end if
  end for
  if V has not been updated **then** break
  end if
end for
return a basis of $V^\perp$

Algorithm 4 Improved variant of the exact Simon's algorithm

Input: n, $O_f : |x\rangle |0\rangle \mapsto |x\rangle |f(x)\rangle$ with $f : \{0,1\}^n \to X$ hiding $\mathcal{H}$
Output: a basis of $\mathcal{H}$

$V = \emptyset$ ▷ Basis of $\mathcal{H}^\perp$
$W = \emptyset$ ▷ Basis of $\{0,1\}^n/(\mathcal{H}^\perp)$
for i from 1 to n **do**
  if No quantum state is available **then**
    Apply Algorithm 1, without measuring, for the state

$$\sum_{x_0 \in \{0,1\}^n/(\mathcal{H})} \sum_{y=0}^{2^n-1} (-1)^{x_0 \cdot y} \sum_{x_1 \in \mathcal{H}} (-1)^{x_1 \cdot y} |y\rangle |f(x_0)\rangle$$

  end if
  amplify the state with the following test: ▷ Exact amplification
    Choose a set Z such that V, W, Z forms a basis of $\{0,1\}^n$.
    Decompose each y as $\sum_{v_k \in V} \delta_k v_k + \sum_{z_j \in Z} \gamma_j z_j$, put γ_1 in an ancilla register.
    return $\gamma_1 = 1$
  end amplify
  Measure $|\gamma_1\rangle$
  if $\gamma_1 = 1$ **then**
    Measure the first register, add the result to V, discard the quantum state
  else
    Add z_1 to W, uncompute the value of $|\gamma_1\rangle$
  end if
end for
return a basis of $V^\perp$

For step 8, note that $H^\perp \subset \langle V, Z \rangle$, hence each j can be uniquely decomposed as $\sum_{v_k \in V} \delta_k v_k + \sum_{z_\ell \in Z} \gamma_\ell z_\ell$.

The amplification is done over the elements which have the component z_1 in their decomposition over the basis (v_i, z_i).

Now, there are two cases:

- For all $j \in \mathcal{H}^\perp$, $\gamma_1 = 0$. This implies $z_1 \notin \mathcal{H}^\perp$. Hence, it represents a non-trivial coset of $\mathcal{H}^\perp$. As it is linearly independent from W by construction, we can add it to W. Note that any amplitude amplification over the quantum state will leave it invariant, as the success probability would be 0 here. Hence, amplitude amplification and the measurement of γ_1 at step 12 leaves the state invariant, which allows to uncompute and recover the original state given by Algorithm 1.
- There exists $j \in \mathcal{H}^\perp$ such that $\gamma_1 = 1$. As $\mathcal{H}^\perp$ is a vector space, we will have $\gamma_1 = 1$ for exactly half of the vectors. Hence, we can apply one step of amplitude amplification, and we will measure 1 in $|\gamma_1\rangle$. This amplitude amplification requires overall 3 calls to Algorithm 1, 1 for the initial state, and 2 for the amplitude amplification step. Once this is done, we can measure j. We will obtain a vector such that $\gamma_1 = 1$. As by construction, all $v_i \in V$ fulfill $\gamma_1 = 0$, the value we measure is linearly independent from V. Hence, we can add it.

At each iteration of the loop, we use 3 queries, except in the case where the previous iteration added a vector in W, in which case it uses 2 queries. Hence, if the last query samples a value in V, the cost will be of $3n - h$, and if the last query samples a value in W, the cost will be $3n - h + 1$. Note that the latter case can only happen if $h \geq 1$, which gives the expected bound. □

2.4 Grover-meets-Simon

The Grover-meets-Simon algorithm [LM17] aims at solving the following problem[1]:

Problem 3 (Grover-meets-Simon). Let $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^m$ be a function such that there exists a unique k_0 such that $f(k_0, \cdot)$ hides a non-trivial subgroup $\mathcal{H}$. Find k_0 and h.

The idea to solve this problem is to use Simon's algorithm as a distinguisher: for the wrong k, Simon's algorithm should return the trivial subgroup, while for k_0 it will always return a non-trivial subgroup. We can then use it as a test function to find k_0, as in Algorithm 5.

The issue with this test is that it allows many false positive: indeed, even for some incorrect k, Simon's algorithm may return a non-trivial subgroup, and these bad cases will also be amplified.

The idea to overcome this issue is to use a test outside of Simon's algorithm to check if its output is correct. This corresponds to the following problem:

[1] In [LM17], the algorithm they introduce is in fact a special case of what we call here Grover-meets-Simon with a perfect external test.

Algorithm 5 Grover-meets-Simon

Input: n, $O_f : |k\rangle\,|x\rangle\,|0\rangle \mapsto |x\rangle\,|f(k,x)\rangle$ with $f(k_0,\cdot)$ of Hidden subgroup $\mathcal{H}$
Output: k_0, a basis of $\mathcal{H}$.

```
amplify over k with the following test:
    H ← Simon's algorithm over f(k, ·)
    if H is empty then
        return 0
    else
        return 1
    end if
end amplify
```

Problem 4 (Grover-meets-Simon with external test). Let $k_0 \in \{0,1\}^n$, $\mathcal{H}$ be a subgroup of $\{0,1\}^n$ and let T be a test function that can check if $(k,H) = (k_0,\mathcal{H})$. Let $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^m$ be a function such that $f(k_0,\cdot)$ hides $\mathcal{H}$. Find k_0 and $\mathcal{H}$.

Hence, instead of only checking the dimension of the subgroup, we compute the subgroup and check if it is correct. This is something we can always do if $\dim(\mathcal{H}) \geq 1$ and $f(k,\cdot)$ is aperiodic if $k \neq k_0$, as presented in Algorithm 6. This test is only perfect with periodic permutations, otherwise it may happen that these equalities hold for an aperiodic function, but it can in general rule out false positive more efficiently than by adding more queries in Simon's algorithm.

Algorithm 6 Periodicity test

Input: k, H
Output: $(k,H) = (k_0,\mathcal{H})$

```
for some x, h ∈ {0,1}^n × H \ {0} do
    if f(k, x) ≠ f(k, x ⊕ h) then
        return False
    end if
end for
return True
```

2.5 The offline Simon's algorithm

The offline Simon's algorithm [Bon+19] is the only known application of Simon's algorithm that do not require quantum access to a secret function. It can be seen as variant of the Grover-meets-Simon algorithm that leverage a special structure of the periodic function. Concretely, the algorithm can solve the following problem:

Algorithm 7 Grover-meets-Simon with external test

Input: n, $O_f : |k\rangle |x\rangle |0\rangle \mapsto |x\rangle |f(k,x)\rangle$ with $f(k_0, \cdot)$ of Hidden subgroup $\mathcal{H}$, a test function T
Output: k_0, s.
amplify over k with the following test:
    $H \leftarrow$ Simon's algorithm over $f(k, \cdot)$
    return $T(k, H)$
end amplify

Problem 5 (Constructing and Finding a Periodic Function). Let $E_k : \{0,1\}^n \to \{0,1\}^\ell$ be a function, $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^\ell$ be a family of functions. Let P be a quantum circuit such that

$$P |i\rangle \sum_{x \in \{0,1\}^n} |x\rangle |E_k(x)\rangle = |i\rangle \sum_{x \in \{0,1\}^n} |x\rangle |f(i,x)\rangle$$

Assume that there exists a unique $i_0 \in \{0,1\}^k$ such that $f(i_0, \cdot)$ hides a non-trivial subgroup. Given oracle access to E_k and P, find i_0 and the period of $f(i_0, \cdot)$.

Here, the compared to Grover-meets-Simon, we add the assumption that the family of functions can be efficiently computed from a fixed function E_k. In most cases, we can restrict ourselves to the following simpler problem:

Problem 6 (Asymmetric Search of a Periodic Function). Let $f : \{0,1\}^m \times \{0,1\}^n \to \{0,1\}^\ell$ be a family of functions and $E_k : \{0,1\}^n \to \{0,1\}^\ell$ be a function.

Assume that there exists a unique $i_0 \in \{0,1\}^m$ such that $f(i_0, \cdot) \oplus E_k$ hides a non-trivial subgroup. Given oracle access to f and E_k, find i_0.

The idea to solve Problem 5 is to see Simon's algorithm slightly differently than usual. Instead of querying an oracle to a function, we suppose that the algorithm is given as input a *database* of E_k, a set of superpositions $\sum_x |x\rangle |E_k(x)\rangle$. It then computes the periodic function from this set, and finally extracts the period. We note $|\psi^m_{E_k}\rangle = \bigotimes_{j=1}^m \sum_x |x\rangle |E_k(x)\rangle$, a state which contains m copies of the superpositions of input/outputs of E_k.

Hence, the algorithm is very similar to Algorithm 5, but the test function fetches $|\psi^m_{E_k}\rangle$, uses it to check if the function is periodic, and finally uncomputes everything, to get back a state close to $|\psi^m_{E_k}\rangle$, which can then be reused in the next iteration.

Now, it remains to estimate the deviation due to the fact that we are not exactly testing if i is equal to i_0.

Lemma 4 (Deviation for the offline Simon's algorithm). *If Simon's algorithm fails with probability at most $2^{-\alpha}$, then the test function in Algorithm 8 tests if $i = i_0$ and adds a noise of amplitude smaller than $2^{-\alpha/2+1}$.*

Algorithm 8 The offline Simon's algorithm

Input: n, An oracle O_{E_k} and a quantum circuit P that fullfils the constraints of Problem 5

Output: i_0.

Query m times O_{E_k}, to construct $|\psi^m_{E_k}\rangle$
amplify over i with the following test:
    Compute m copies of $\sum_x |x\rangle |f(i, x)\rangle$ from $|\psi^m_{E_k}\rangle$ and P.
    Apply a Hadamard gate on the first register of each copy.
    Compute in superposition the rank r of the values in each first register.
    $b \leftarrow r \neq n$
    Uncompute everything but the value of b, to recover $|\psi^m_{E_k}\rangle$.
    Return b
end amplify

Proof. The ideal circuit we want takes $\sum_i |i\rangle |\psi^m_{E_k}\rangle |0\rangle$ as input, and produces the output $\sum_i |i\rangle |\psi^m_{E_k}\rangle |i = i_0\rangle$. This is however not exactly what we have. If $i = i_0$, then the answer is correct with probability 1, as the periodicity reduces the maximal rank, and it behaves like the ideal circuit.

For $i \neq i_0$, the circuit first constructs

$$|i\rangle \left(\bigotimes_{j=1}^{m} \sum_{x_j} |x_j\rangle |f(i, x_j)\rangle \right) |0\rangle \quad .$$

It then applies a Hadamard gate on $|x\rangle$, to obtain

$$|i\rangle \left(\bigotimes_{j=1}^{m} \sum_{x_j} \sum_{y_j} (-1)^{x_j \cdot y_j} |y_j\rangle |f(i, x_j)\rangle \right) |0\rangle \quad .$$

Once this is done, we can compute the rank of the m values y_j, and obtain

$$|i\rangle \left(\bigotimes_{j=1}^{m} \sum_{\substack{(y_1,\dots,y_m) \\ \text{of maximal rank}}} \sum_{x_j} (-1)^{x_j \cdot y_j} |y_j\rangle |f(i, x_j)\rangle \right) |0\rangle$$
$$+ |i\rangle \left(\bigotimes_{j=1}^{m} \sum_{\substack{(y_1,\dots,y_m) \\ \text{of lower rank}}} \sum_{x_j} (-1)^{x_j \cdot y_j} |y_j\rangle |f(i, x_j)\rangle \right) |1\rangle \quad .$$

The second term is the annoying one. The state deviates from the state we want by 2 times the amplitude of the second term, as the terms with the correct results are missing, and the terms with the wrong result have been added. Hence, the norm of the noise is 2 times the amplitude of this term. As the probability for the

m words of n bits to be of rank n is exactly Simon's algorithm success probability with the function $f(i_0, \cdot)$, the amplitude of the second term is bounded by $2^{-\alpha/2}$. With the factor 2, we obtain $2^{-\alpha/2+1}$. $\square$

Algorithm 8 allows to only make the quantum queries to E_k once, at the beginning. Now, the idea to use only classical queries is to construct manually the quantum superposition over E_k *from the classical queries of all its* 2^n *possible inputs*. This is presented in Algorithm 9. As this is can only be done in $\mathcal{O}(2^n)$ time, we really need to reuse the queries in order to still have a time-efficient algorithm.

Algorithm 9 Generation of $|\psi^m_{E_k}\rangle$ from classical queries

Input: A classical oracle to E_k, m
Output: $|\psi^m_{E_k}\rangle$

$|\phi\rangle \leftarrow \bigotimes_m \sum_x |x\rangle |0\rangle$
for $0 \leq i < 2^n$ **do**
  Query $E_k(i)$
  Apply to each register in $|\phi\rangle$ the operator

$$|x\rangle |y\rangle \mapsto \begin{cases} |x\rangle |y \oplus E_k(i)\rangle & \text{if } x = i \\ |x\rangle |y\rangle & \text{otherwise} \end{cases}$$

end for
return $|\phi\rangle$

3 Concrete cost for Simon's algorithm

3.1 Ideal case: periodic permutations

For all our analyses, we will use the following lemma to study only the aperiodic case in Simon's algorithm.

Lemma 5 (Simon reduction). *Let $n \in \mathbb{N}$, let X be a set. There exists a bijection φ_c from $\{f \in \{0,1\}^n \to X | f$ hides a subgroup of dimension $c\}$ to*

$$\{\mathcal{H} \subset \{0,1\}^n | \dim(\mathcal{H}) = c\} \times \left\{ \widehat{f} \in \{0,1\}^{n-c} \to X \middle| \widehat{f} \text{ is aperiodic} \right\}$$

such that, with $\varphi_c(f) = (\mathcal{H}, \widehat{f})$, the behaviour of Simon's algorithm with f and $\widehat{f}$ is identical, up to isomorphism. Moreover, f is a periodic permutation if and only if $\widehat{f}$ is a permutation.

Proof. See Appendix A.

Theorem 4 (Simon's algorithm success probability). *Let f be a function on n bits that fullfils the promise of Problem 1, $\mathcal{H}$ its hidden subgroup. With $T \geq n - \dim(\mathcal{H})$ queries, Simon's algorithm succeeds with probability*

$$pr_T = \prod_{i=0}^{n-\dim(\mathcal{H})-1} \left(1 - \frac{1}{2^{T-i}}\right).$$

Moreover, $\left(1 - 2^{n-T-dim(\mathcal{H})-1}\right)^2 \leq pr_T \leq 1 - 2^{n-T-dim(\mathcal{H})-1}$.

Proof. Let $\alpha \geq 0$, $y_1, \ldots, y_T$ be the $T = n + \alpha$ outputs of Simon's routine. The y_i are sampled uniformly in the set $\{y | \forall h \in \mathcal{H}, y \cdot h = 0\}$. Let

$$M = \begin{bmatrix} y_1 \\ \ldots \\ y_T \end{bmatrix}$$

be the matrix whose rows are the y_i. Then, the rank of the y_i is the rank of M.

If $\dim(\mathcal{H}) = 0$, the maximal rank is n, and M is of maximal rank if and only if its columns form a free family, as $T \geq n$. The first column is non-zero with probability $1 - \frac{1}{2^T}$. If $y_1, \ldots, y_{i-1}$ form a free family, y_i is linearly independent from them with probability $1 - \frac{1}{2^{T-i-1}}$, as $|\langle y_1, \ldots, y_{i-1}\rangle| = 2^{i-1}$. Hence, the columns form a free family with probability

$$\mathrm{pr}_T = \prod_{i=0}^{n-1} \left(1 - \frac{1}{2^{T-i}}\right).$$

Taking the log, we obtain

$$\log(\mathrm{pr}_T) = \sum_{i=0}^{n-1} \log\left(1 - \frac{1}{2^{T-i}}\right).$$

Developing in power series produces

$$\sum_{i=0}^{n-1} - \sum_{j=1}^{\infty} \frac{1}{j 2^{(T-i)j}}$$

Interchanging the sums produces

$$-\sum_{j=1}^{\infty} \frac{1}{j 2^{Tj}} \sum_{i=0}^{n-1} 2^{ji}$$

As $2^{j(n-1)} \leq \sum_{i=0}^{n-1} 2^{ji} \leq 2^{j(n-1)+1}$, we have

$$-2 \sum_{j=1}^{\infty} \frac{1}{j 2^{(T-n+1)j}} \leq \log(\mathrm{pr}_T) \leq -\sum_{j=1}^{\infty} \frac{1}{j 2^{(T-n+1)j}}$$

Factoring the power series produces

$$\log\left(\left(1-\frac{1}{2^{T-n+1}}\right)^2\right) \leq \log(\mathrm{pr}_T) \leq \log\left(1-\frac{1}{2^{T-n+1}}\right)$$

Hence, the success probability is bounded by $\left(1-\frac{1}{2^{T-n+1}}\right)^2$ and $1-\frac{1}{2^{T-n+1}}$.

If $\dim(\mathcal{H}) > 0$, we can apply Lemma 5 and the previous case to obtain the desired result. □

Theorem 5 (Simon's algorithm query complexity). *Let f be a function on n bits that fullfils the promise of Problem 1, $\mathcal{H}$ its hidden subgroup. To succeed, Simon's algorithm requires, on average:*

- *$n - dim(\mathcal{H}) + 2$ queries if $n - dim(\mathcal{H}) \geq 2$.*
- *2 queries if $n - dim(\mathcal{H}) = 1$*

Proof. If $n-\dim(\mathcal{H}) = 1$, then $\mathcal{H}^\perp = \{0, s\}$ for a given nonzero s. The algorithm succeeds if and only if s is measured, which occurs with probability one half. Hence, 2 queries on average are needed.

Now, assume $n - \dim(\mathcal{H}) \geq 2$, or, by Lemma 5, that $n \geq 2$, $\dim(\mathcal{H}) = 0$ and we want to sample a full rank set of vectors.

Now, let's consider the random variable N, which is the number of queries Simon's algorithm require to succeed. From Theorem 4, we know $\Pr[N \leq T] = \mathrm{pr}_T$. Hence, as $\dim(\mathcal{H}) = 0$, for $T > n$,

$$\begin{aligned}
\Pr[N = T] = \quad & \mathrm{pr}_T - \mathrm{pr}_{T-1} = \prod_{i=0}^{n-1}\left(1-\frac{1}{2^{T-i}}\right) - \prod_{i=0}^{n-1}\left(1-\frac{1}{2^{T-1-i}}\right) \\
= \quad & \prod_{i=0}^{n-1}\left(1-\frac{1}{2^{T-i}}\right) - \prod_{i=1}^{n}\left(1-\frac{1}{2^{T-i}}\right) = \mathrm{pr}_T\left(1-\frac{1-2^{n-T}}{1-2^{-T}}\right) \\
= \quad & \mathrm{pr}_T\left(1-\frac{2^T-2^n}{2^T-1}\right) = \mathrm{pr}_T\left(\frac{2^n-1}{2^T-1}\right).
\end{aligned}$$

Now, as $T > n$, we have the bound $\mathrm{pr}_T 2^{n-T}(1-2^{-n}) \leq \Pr[N = T] \leq \mathrm{pr}_T 2^{n-T}$.

Finally, the expected number of queries is

$$\mathbb{E}[N] = \sum_{i=n}^{+\infty} i \Pr[N = i] = n + \sum_{i=1}^{+\infty} i \Pr[N = n + i]$$

Hence, we have the bound

$$n + \left(1-2^{-n}\right)\sum_{i=1}^{\infty} i 2^{-i}\mathrm{pr}_{n+i} \leq \mathbb{E}[N] \leq n + \sum_{i=1}^{\infty} i 2^{-i}\mathrm{pr}_{n+i}$$

By Theorem 4, we have

$$
\begin{aligned}
& \sum_{i=1}^{\infty} i2^{-i}(1-2^{-i}+2^{-2i-2}) \leq \sum_{i=1}^{\infty} i2^{-i}\mathrm{pr}_{n+i} \leq \sum_{i=1}^{\infty} i2^{-i}(1-2^{-i-1}) \\
\Leftrightarrow\quad & \sum_{i=1}^{\infty} i2^{-i} - \sum_{i=1}^{\infty} i4^{-i} + \frac{1}{4}\sum_{i=1}^{\infty} i8^{-i} \leq \sum_{i=1}^{\infty} i2^{-i}\mathrm{pr}_{n+i} \leq \sum_{i=1}^{\infty} i2^{-i} - \frac{1}{2}\sum_{i=1}^{\infty} i4^{-i} \\
\Leftrightarrow\quad & \frac{1}{2}\frac{1}{\left(1-\frac{1}{2}\right)^2} - \frac{1}{4}\frac{1}{\left(1-\frac{1}{4}\right)^2} + \frac{1}{32}\frac{1}{\left(1-\frac{1}{8}\right)^2} \leq \sum_{i=1}^{\infty} i2^{-i}\mathrm{pr}_{n+i} \leq \frac{1}{2}\frac{1}{\left(1-\frac{1}{2}\right)^2} - \frac{1}{8}\frac{1}{\left(1-\frac{1}{4}\right)^2} \\
\Leftrightarrow\quad & 2-\frac{4}{9}+\frac{2}{49} \leq \sum_{i=1}^{\infty} i2^{-i}\mathrm{pr}_{n+i} \leq 2-\frac{2}{9}
\end{aligned}
$$

Now, as $n \geq 2$, $1-2^{-n} \geq \frac{3}{4}$, and $(1-2^{-n})\left(2-\frac{4}{9}+\frac{2}{49}\right) > 1$.

Hence, we have that $\mathbb{E}[N]$ is between $n+1$ and $n+2$. Rounding upwards, we obtain the desired result.

3.2 A general criterion

We might want to apply Simon's algorithm on functions that are only periodic functions, and can have more preimages per image. However, we cannot expect to have a functioning algorithm in all cases. Indeed, let's consider

$$
f_s : \begin{array}{rcl} \{0,1\}^n & \to & \{0,1\} \\ x & \mapsto & \begin{cases} 1 \text{ if } x \in \{0,s\} \\ 0 \text{ otherwise} \end{cases} \end{array} .
$$

The function f_s has the hidden subgroup $\{0, s\}$. This function can be constructed from oracle access to a test function $t_s(x) = \delta_{x,s}$. Hence, as finding the hidden subgroup from f_s is equivalent to recovering s, and as quantum search is optimal to recover s given quantum oracle access to $t_s(x)$, we cannot hope for a polynomial, over even subexponential algorithm in this case.

To quantify how suitable the function is for Simon's algorithm, we define $\varepsilon(f)$ similarly to [Kap+16]

$$
\varepsilon(f) = \max_{t \in \{0,1\}^n/(\mathcal{H})} \Pr_x[f(x \oplus t) = f(x)].
$$

This value estimates the probability that any given t is present as an additional period for some of the output vectors of Algorithm 1. It allows to bound the success probability of Simon's algorithm.

Proposition 1 (Success probability with more preimages, adapted from [Kap+16, Theorem 1]). *Let f be a periodic function, $\mathcal{H}$ its hidden subgroup, and $\varepsilon(f)$ be defined as above. After $c(n-\dim(\mathcal{H}))$ steps, Simon's algorithm on f succeeds with probability greater than* $1-\left(2\left(\frac{1+\varepsilon(f)}{2}\right)^c\right)^{n-\dim(\mathcal{H})}$.

Proof (Adapted from [Kap+16]). Assume $\dim(\mathcal{H}) = 0$, thanks to Lemma 5. given a value $t \neq 0$, we will estimate the probability that we sample a vector orthogonal to t.

At the end of Simon's algorithm, the quantum state is

$$\frac{1}{2^n} \sum_{x \in \{0,1\}^n} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} |y\rangle |f(x)\rangle$$

Hence the probability of measuring a $y \in \{0,t\}^\perp$ is

$$\begin{aligned}
&\left\| \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \sum_{y \in \{0,t\}^\perp} (-1)^{x \cdot y} |y\rangle |f(x)\rangle \right\|^2 \\
=&\frac{1}{2^{2n}} \sum_{x \in \{0,1\}^n} \sum_{x' \in \{0,1\}^n} \sum_{y \in \{0,t\}^\perp} (-1)^{(x \oplus x') \cdot y} \langle f(x) | f(x') \rangle \\
=&\frac{1}{2^{2n}} \sum_{x \oplus x' \in \{0,t\}} 2^{n-1} \langle f(x) | f(x') \rangle \text{ with Lemma 1} \\
=&\frac{1}{2^{n+1}} (\sum_{x \in \{0,1\}^n} \langle f(x) | f(x) \rangle + \langle f(x) | f(x \oplus t) \rangle) \\
=&\frac{1}{2}(1 + \Pr_x[f(x \oplus t) = f(x)])
\end{aligned}$$

This is smaller than $\frac{1}{2}(1 + \varepsilon(f))$. Hence, the probability than all cn vectors are orthogonal to the same t is lower than $\left(\frac{1+\varepsilon(f)}{2}\right)^{cn}$.

By the union bound, the probability that any t is orthogonal to all the vectors is lower than $2^n \left(\frac{1+\varepsilon(f)}{2}\right)^{cn}$.

Hence, the proposition holds. □

Theorem 6 (Number of queries in general). *Let f be a periodic function, $\mathcal{H}$ its hidden subgroup, $\varepsilon(f)$ be defined as above. Simon's algorithm on f fails with probability lower than $2^{-\alpha}$ after $\frac{1}{1-\log_2(1+\varepsilon(f))} (n - \dim(\mathcal{H}) + \alpha)$ queries.*

Proof. Assume $\dim(\mathcal{H}) = 0$, thanks to Lemma 5. From Proposition 1, after cn queries, the failure probability is at most $\left(2\left(\frac{1+\varepsilon(f)}{2}\right)^c\right)^n$. Hence, we have

$$\begin{aligned}
\tfrac{1}{2^\alpha} &= \left(2\left(\tfrac{1+\varepsilon(f)}{2}\right)^c\right)^n \\
\Leftrightarrow -\alpha &= n\left(1 + c\log_2\left(\tfrac{1+\varepsilon(f)}{2}\right)\right) \\
\Leftrightarrow \quad n + \alpha &= -cn\log_2\left(\tfrac{1+\varepsilon(f)}{2}\right) \\
\Leftrightarrow \quad cn &= \tfrac{1}{\log_2\left(\frac{2}{1+\varepsilon(f)}\right)}(n + \alpha)
\end{aligned}$$

As cn is the total number of queries, the theorem holds. □

3.3 With random functions

Now, we want to estimate the value of $\varepsilon(f)$. We show below that for all but a negligible fraction of functions, as long as the domain of the function is large enough, it is too small to have any impact on the number of queries.

Lemma 6. *Let $\mathcal{F} = \{0,1\}^n \to \{0,1\}^m$. Then*

$$\left|\left\{f \in \mathcal{F} \middle| \varepsilon(f) \geq \frac{2\ell}{2^n}\right\}\right| \leq 2^n \frac{2^{(n-m)\ell}}{\ell!} |\mathcal{F}|$$

Proof. First of all, $|\mathcal{F}| = 2^{m2^n}$.

Let f be such that $\varepsilon(f) \geq \frac{2\ell}{2^n}$. There exists a t and a set L of size 2ℓ such that for all $x \in L$, $f(x) = f(x \oplus t)$.

Hence, to enumerate all the functions such that $\varepsilon(f) \geq \frac{2\ell}{2^n}$, we can choose a t and a set L' of ℓ inputs, and then choose ℓ values for $x \in L'$ and $x \oplus t$, and finally choose arbitrary values for the remaining $2^n - 2\ell$ inputs.

Hence, there is less than

$$\binom{2^n}{\ell} 2^{m\ell} 2^{m(2^n - 2\ell)}$$

such functions for a given t. As $\binom{2^n}{\ell} \leq \frac{2^{n\ell}}{\ell!}$, and there is less than 2^n possible t, the lemma holds. □

Corollary 1 (ε for aperiodic functions). *The fraction of aperiodic functions such that $\varepsilon(f) \geq \frac{2\ell}{2^n}$ is also bounded by $2^n \frac{2^{(n-m)\ell}}{\ell!}$.*

Proof. Aperiodic functions are exactly functions such that $\varepsilon(f) < 1$. Hence, by removing the functions such that $\varepsilon(f) = 1$, we can only decrease the fraction of functions such that $\varepsilon(f) \geq \frac{2\ell}{2^n}$. □

Theorem 7 (ε for large enough m). *Let $\mathcal{F} = \{0,1\}^n \to \{0,1\}^m$. The proportion of functions in $\mathcal{F}$ such that $\varepsilon \geq e2^{1-m} + 2^{1-n}(n + \alpha)$ is upper bounded by $2^{-\alpha}$.*

Proof. We bound the fraction of functions such that $\varepsilon \geq \frac{2\ell}{2^n}$, from the previous lemma. As $\ell! \geq \left(\frac{\ell}{e}\right)^\ell$, the fraction is upper bounded by

$$2^n \left(\frac{2^{n-m}e}{\ell}\right)^\ell$$

Taking the log, we obtain

$$n + \ell \log\left(\frac{2^{n-m}e}{\ell}\right)$$

Now, using the fact that $\log(x) \leq -\log(e)(1-x)$, we have the upper bound

$$n - \log(e)\ell\left(1 - \frac{2^{n-m}e}{\ell}\right)$$

This is lower than $-\alpha$ if $\ell \geq e2^{n-m} + \frac{1}{\log(e)}(n+\alpha)$. Using the bound on ε and the fact that $\frac{2}{\log(e)} < 2$, we obtain the result. $\square$

Remark 3. The previous bound is only meaningful if $m \geq 3$. Obtaining bounds for smaller m is likely to require a more precise analysis, as we have, for $n \geq m$ and all f, $\varepsilon(f) \geq 2^{-m}$.

Theorem 8 (Success probability for random functions). *Assume $m \geq \log_2(4e(n-h+\alpha+1))$. Then the fraction of functions in $\{0,1\}^n \to \{0,1\}^m$ with a hidden subgroup of dimension h such that, after $n-h+\alpha+1$ queries, Simon's algorithm fails with a probability greater than $2^{-\alpha}$ is bounded by $2^{n-h-\frac{2^{n-h}}{4(n-h+\alpha+1)}}$.*

Proof. Without any loss of generality, with Lemma 5, assume $h = 0$. From Theorem 6, we want to know when

$$\frac{1}{1-\log_2\left(1+\varepsilon(f)\right)}\left(n - \dim(\mathcal{H}) + \alpha\right) \geq n + \alpha + 1$$

This is equivalent to $\log_2\left(1+\varepsilon(f)\right) \geq \frac{1}{n+\alpha+1}$, which implies $\varepsilon(f) \geq \frac{1}{n+\alpha+1}$. Now, from Theorem 7, we have that with probability lower than $2^{-\beta}$,

$$\varepsilon(f) \geq e2^{1-m} + 2^{1-n}(n+\beta)$$

Hence, $\beta = 2^n(\frac{1}{2(n+\alpha+1)} + e2^{-m}) - n$ is sufficient. As $m \geq \log_2(4e(n+\alpha+1))$, we obtain the bound. $\square$

Theorem 9 (Average complexity for random functions). *Assume $m \geq \log_2(4e((n-h)+1))$ and $n-h \geq 4$. Then the fraction of functions in $\{0,1\}^n \to \{0,1\}^m$ with a hidden subgroup of dimension h such that, on average, Simon's algorithm requires more than $n-h+4$ queries is bounded by $2^{n-h-\frac{2^{n-h}}{4(2(n-h)+1)}}$.*

Proof. We use Theorem 8 with $\alpha = n$. We obtain that except for a fraction $2^{n-\frac{2^n}{4(2n+1)}}$ of the functions, Simon's algorithm will succeed with a probability greater than $1-2^{-\beta}$ with $n+1+\beta$ queries for all β in $[1;n]$.

Now, the average complexity will be smaller than in the case where the success probability after $n+1+\beta$ queries is $1-2^{-\beta}$. Furthermore, we can remark that after n queries, the success probability is greater than $1-2^{-n}$. Hence, the average complexity will be bounded by

$$\frac{1}{1-2^{-n}}\left(n+1+\sum_{i=1}^{n} i(1-2^{-\beta}-1-2^{-\beta+1})\right) \leq \frac{1}{1-2^{-n}}\left(n+1+2\right)$$

Now, as $n \geq 3$, we have that $2^n > n+4 \Leftrightarrow \frac{1}{1-2^{-n}}\left(n+3\right) < n+4$, which is the desired result. $\square$

4 Concrete cost for Grover-meets-Simon

In the Grover-meets-Simon algorithm, Simon's algorithm is used to identify one function among a family. The main issue is that there can be false positives, that is, functions that are identified as periodic while they are not.

Theorem 10 (Success probability for plain Grover-meets-Simon). *Let $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^m$ be a function such that there exists a unique i_0 such that $f(i_0, \cdot)$ hides a non-trivial subgroup $\mathcal{H}$. If for all $f(i, \cdot)$, Simon's algorithm succeeds with probability at least $1 - 2^{-\alpha}$, then Algorithm 5 succeeds in $\frac{\pi}{4 \arcsin \sqrt{2^{-k}}}$ iterations with probability at least $1 - 2^{k-\alpha} - 2^{2(k-\alpha)+1} - 2^{-k+2} - 2^{-\alpha+2} - 2^{-2k+6} + 2^{-2\alpha+7} + 2^{2k-4\alpha+6}$.*

Proof. We want to apply Theorem 1. Hence, we need to bound the success probability of the quantum circuit. As it succeeds for i_0 with probability 1, the success probability is at least 2^{-k}. Unfortunately, it also succeeds for some $i \neq i_0$, with probability at most $2^{-\alpha}$. Hence, the initial success probability will be lower than $2^{-k} + 2^{-\alpha}$.

Hence, by Theorem 1, we will measure good element with probability at least $1 - (\beta + \sqrt{(1+\beta)p_0} + 2\sqrt{1+\beta}^3 p_0)^2$ after $\frac{\pi}{4 \arcsin \sqrt{2^{-k}}}$ iterations, and we have $\beta = 2^{k-\alpha}$, $p_0 = 2^{-k}$. As $(a+b)^2 \leq 2a^2 + 2b^2$, we obtain the bound

$$1 - 2^{2(k-\alpha)+1} - 2^{-k+2} - 2^{-\alpha+2} - 2^{-2k+6} + 2^{-2\alpha+7} + 2^{2k-4\alpha+6}$$

This is however not enough, as a good element according to the test may still be incorrect. Among the amplified elements, i_0 have a probability to be measured 2^α greater than all the $2^k - 1$ wrong answers. Hence, the probability to measure the correct i_0 among the amplified values is at least $\frac{1}{1+2^{k-\alpha}} \geq 1 - 2^{k-\alpha}$.

Combining the two probabilities by the union bound allows us to obtain the desired result. □

Corollary 2 (Number of queries for plain Grover-meets-Simon). *For all but a negligible fractions of functions, the Grover-meets-Simon algorithm fails with probability lower than one half with $n + k + 2$ queries.*

Proof. We combine Theorem 10 and Theorem 8, with $h = 1$ and $\alpha = k + 2$. □

Now, we study the situation with external tests, either when we have perfect external tests, or using Algorithm 6.

Theorem 11 (Grover-meets-Simon with perfect external test). *Let $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^m$ be a function such that there exists a i_0 such that $f(i_0, \cdot)$ hides a subgroup $\mathcal{H}$, and there exists a test function T such that $T(i, H) = 1$ if and only if $(i, H) = (i_0, \mathcal{H})$. If for $f(i_0, \cdot)$, Simon's algorithm succeeds with probability at least $1 - 2^{-\alpha}$, then Algorithm 7 succeeds in $\frac{\pi}{4 \arcsin \sqrt{2^{-k}}}$ iterations with probability at least $1 - (2^{-\alpha+1} + 2^{-k/2} + 2^{-k+1})^2$.*

Proof. If we have a perfect external test, then the initial success probability will be between $2^{-k}(1-2^{-\alpha})$ and 2^{-k}, and no $i \neq i_0$ can ever be amplified. In that case, we can directly apply Theorem 1 to obtain the bound

$$1-(2^{-\alpha+1}+2^{-k/2}+2^{-k+1})^2.$$

□

Corollary 3 (Number of queries for Grover-meets-Simon with perfect external test). *For all but a negligible fractions of functions, the Grover-meets-Simon with perfect external test algorithm fails with probability lower than one half with $n+2$ queries.*

Proof. We combine Theorem 11 and Theorem 8, with $h=1$ and $\alpha=2$. □

Remark 4 (Grover-meets-Simon for periodic permutation). There is a perfect test for periodic permutations which costs only 2 queries. It amounts in testing wether or not the function fullfils $f(0)=f(h_i)$, with (h_i) a basis of the guessed hidden subgroup H. This will only be the case if the function is indeed periodic.

Remark 5. The external test allows to remove the error terms in $k-\alpha$, which means that there is no longer a dependency in k for the nimimal number of queries. We only need n plus a constant number of queries to succeed.

When there is no perfect external test (for example in the case of quantum distinguishers), we can still do better than Theorem 10 using Algorithm 6.

Theorem 12 (Grover-meets-Simon with periodicity test). *Let $k \geq 3$, $f:\{0,1\}^k \times \{0,1\}^n \rightarrow \{0,1\}^m$ be a function such that there exists a unique i_0 such that $f(i_0,\cdot)$ hides a non-trivial subgroup $\mathcal{H}$. If for all $f(i,\cdot)$, Simon's algorithm succeeds with probability at least $1-2^{-\alpha}$ and for all $i \neq i_0$, $\varepsilon(f(i,\cdot)) \leq \varepsilon$, then Algorithm 7 with γ queries in Algorithm 6 succeeds in $\frac{\pi}{4\arcsin\sqrt{2^{-k}}}$ iterations with probability at least $1-\varepsilon^\gamma 2^{k-\alpha+1}-(2^{-\alpha+1}+2^{-k/2}+2^{-k+1})^2$.*

Proof. If $i=i_0$, Simon's algorithm succeeds with a probability between $1-2^{-\alpha}$ and 1, and otherwise, it will output a success with a probability between 0 and $2^{-\alpha}$. This false positive will be catched by the periodicity test except if for all the values tested, we have $f(x)=f(x\oplus h)$, which occurs with probability at most ϵ^γ, by definition of ϵ. The success probability for the amplification is then between $2^{-k}(1-2^{-\alpha})$ and $2^{-k}+(1-2^{-k})2^{-\alpha}\varepsilon^\gamma$. Moreover, the fraction of the amplified values which correspond to $i=i_0$ is at least $\frac{2^\alpha-1}{2^\alpha+(2^k-1)\varepsilon^\gamma} \geq 1-\varepsilon^\gamma\frac{2^k-1}{2^\alpha-1}$.

Now, by Theorem 1, if the success probability for the amplification is lower than 2^{-k}, the amplification will succeed with probability at least

$$1-(2^{-\alpha+1}+2^{-k/2}+2^{-k+1})^2 \quad .$$

Otherwise, the probability will be at least

$$1-(2^{-\alpha}+2^{-k/2}\sqrt{1+2^{-\alpha}\varepsilon^\gamma}+2^{-k+1}\sqrt{1+2^{-\alpha}\varepsilon^\gamma}^3)^2 \quad .$$

We will prove that the latter is higher than the former. This is the case if

$$2^{-\alpha} \geq 2^{-k/2}(\sqrt{1+2^{-\alpha}\varepsilon^{\gamma}}-1) + 2^{-k+1}(\sqrt{1+2^{-\alpha}\varepsilon^{\gamma}}^3-1) \quad .$$

As $\sqrt{1+x} \leq 1+x/2$ and, by convexity, if $x \leq 1$, $(1+x)^{3/2}-1 \leq (2^{3/2}-1)x \leq 2x$, this is true if

$$2^{-\alpha} \geq 2^{-k/2-1}2^{-\alpha}\varepsilon^{\gamma} + 2^{-k+2}2^{-\alpha}\varepsilon^{\gamma} \quad .$$

This is equivalent to

$$\varepsilon^{\gamma}(2^{-k/2-1} + 2^{-k+2}) \leq 1 \quad .$$

This is always true if $k \geq 3$ or $\varepsilon \leq 1/2$. As we assumed $k \geq 3$, this holds, and the amplitude amplification succeeds with probability at least

$$1 - (2^{-\alpha+1} + 2^{-k/2} + 2^{-k+1})^2 \quad .$$

Now, the correct result will be measured with probability greater than $1 - \varepsilon^{\gamma}\frac{2^k-1}{2^{\alpha}-1} \geq 1 - \varepsilon^{\gamma}2^{k-\alpha+1}$. By the union bound, we obtain the desired result. □

Corollary 4 (Number of queries for Grover-meets-Simon with periodicity test). *For all but a negligible fractions of functions, Grover-meets-Simon with periodicity test fails with probability lower than one half with* $n+2+2\left\lceil\frac{k+2}{\min(n,m)-\log(4n+4k+6)}\right\rceil$ *queries.*

Proof. We use Theorem 12 and Theorem 8, with $h = 1$ and $\alpha = 2$. We also need to bound the term $\varepsilon^{\gamma}2^{k-\alpha+1}$, with $\alpha = 2$. This is smaller than 2^{-3} if

$$\gamma \geq \frac{k+2}{-\log_2(\varepsilon)}$$

From Theorem 7, we have that the fraction of functions such that $\varepsilon \geq e2^{1-m} + 2^{1-n}(n+n+2k)$ is bounded by 2^{-n-2k}. Now, we use that

$$e2^{1-m} + 2^{1-n}(n+n+2k) \leq 2^{-\min(m,n)}(4n+4k+6)$$

and take the log. We obtain the bound by remarking that a call to the periodicity test costs 2 queries. This bound will not hold for all 2^k functions with probability at most 2^{-n-k}, which is negligible. □

5 Concrete cost for the offline Simon's algorithm

The offline Simon's algorithm reuses the quantum queries between the tests. This allows to make the algorithm work with classical queries, at the expense of having more noise than with the Grover-meets-Simon algorithm.

Theorem 13 (Success probability for the Offline Simon's algorithm). *Let* $f : \{0,1\}^k \times \{0,1\}^n \to \{0,1\}^m$ *be a function such that there exists a unique* i_0 *such that* $f(i_0,\cdot)$ *hides a non-trivial subgroup* $\mathcal{H}$*. If for all* $f(i,\cdot)$*, Simon's algorithm succeeds with probability at least* $1-2^{-\alpha}$*, then Algorithm 8 succeeds in* $\frac{\pi}{4\arcsin\sqrt{2^{-k}}}$ *iterations with probability at least* $1 - 2^{k-\alpha+3} - 2^{2(k-\alpha+2)} - 2^{-k+3} - 2^{-\alpha+3} - 2^{-2k+5} - 2^{-k-\alpha+8}$

Proof. We use Theorem 2 to bound the divergence between the Offline Simon's algorithm and the corresponding Grover-meets-Simon algorithm, whose success probability comes from Theorem 10.

After $\frac{\pi}{4\arcsin\sqrt{2^{-k}}}$ iterations, by Theorem 10, the amplitude of the bad elements in Grover-meets-Simon is at most $(2^{k-\alpha}+\sqrt{2^{-k}+2^{-\alpha}}+2^{-k+1}\sqrt{1+2^{k-\alpha}}^3)$. Now, we add an additional noise of amplitude $2^{-\alpha/2+1}$ for each iteration. The number of iterations is bounded by $2^{k/2}$, hence the total noise has an amplitude of at most $2^{k/2-\alpha/2+1}$.

Combining the two, the probability that the algorithm succeeds is at least $1-(2^{k-\alpha}+\sqrt{2^{-k}+2^{-\alpha}}+2^{-k+1}\sqrt{1+2^{k-\alpha}}^3+2^{k/2-\alpha/2+1})^2 \geq 1-2^{k-\alpha+2}-2^{2(k-\alpha+2)}-2^{-k+3}-2^{-\alpha+3}-2^{-2k+5}(1+2^{k-\alpha})^3$. Now, we use the fact that in order to have a probability greater than 0, we need $2^{k-\alpha}\leq 1$, which means, as x^3 is convex, that $(1+2^{k-\alpha})^3 \leq 1+7\times 2^{k-\alpha}$. Hence, the probability is bounded by

$$1-2^{k-\alpha+2}-2^{2(k-\alpha+2)}-2^{-k+3}-2^{-\alpha+3}-2^{-2k+5}-2^{-k-\alpha+8} \quad .$$

Now, we add the addionnal failure probability of $2^{k-\alpha}$ due to the fact that even if we measure one of the amplified values, the result may still be incorrect.

Overall, we obtain a success probability bounded by

$$1-2^{k-\alpha+3}-2^{2(k-\alpha+2)}-2^{-k+3}-2^{-\alpha+3}-2^{-2k+5}-2^{-k-\alpha+8} \quad .$$

□

Remark 6. Contrary to the Grover-meets-Simon case, we cannot remove the error terms in $2^{k-\alpha}$ with an external test, which means that we cannot remove the direct dependency in k in the number of queries.

Corollary 5. *For all but a negligible fractions of functions, the offline Simon's algorithm fails with probability lower than one half with $n+k+5$ queries.*

Proof. We use Theorem 13 and Theorem 8, with $h=1$ and $\alpha=k+5$. □

6 Applications in cryptanalysis

In cryptanalysis, we're interested in specific sets of parameters.

The hidden subgroup is generally of dimension 1, the only known exception being the quantum cryptanalysis of AEZ [Bon17], which uses hidden subgroups up to dimension 3. We seldom have periodic permutations, but the functions we're considering are expected to have a low $\varepsilon(f)$, as otherwise it would mean that the function exhibit a poor differential property, which may lead to some dedicated attacks that leverage this fact. Finally, in these attacks, the domain of the function is generally smaller than the codomain. One of the only counter-examples is the attack on the MAC Chaskey [Mou+14], which has 128 bits of input and 64 bits of output. The situation is even better for the offline Simon's algorithm, as we often consider functions with a partially fixed input. Hence, Theorem 8 and Theorem 9 will apply with overwhelming probability.

External tests in Grover-meets-Simon. In general, a perfect external test amounts to checking if the output of a secret function (say, a block cipher) matches the output of the guessed function. Hence, this test will be wrongfully passed with probability 2^{-n}, which means that we need $\lceil \frac{k}{n} \rceil$ tests to filter out all the false positives. Hence, we can estimate that at best, it should cost $\lceil \frac{k}{n} \rceil$ queries. However, this test may use a different function, and will at worse cost as much as a query to the periodic function. Hence, we have a gain, even if the dependency in $\frac{k}{n}$ is still there.

The situation is however different with periodic permutations, as 2 queries to the periodic permutation are always enough.

Open problems. When our approach is applicable, a gain of 1 query for periodic functions might be obtained by a more precise analysis, at least for a large enough n and a small enough error rate.

Moreover, there is still one case out of reach with our approach: periodic functions with 1 bit of output, which were proposed in [MS19] as a potential way to reduce the memory cost of Simon's algorithm.

Acknowledgements. The author would like to thank André Schrottenloher for interesting discussions on Simon's algorithm.

References

[BH97] G. Brassard and P. Hoyer. "An exact quantum polynomial-time algorithm for Simon's problem". In: *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems*. 1997, pp. 12–23.

[BNS19] Xavier Bonnetain, María Naya-Plasencia, and André Schrottenloher. "On Quantum Slide Attacks". In: *SAC 2019*. Ed. by Kenneth G. Paterson and Douglas Stebila. Vol. 11959. LNCS. Springer, Heidelberg, Aug. 2019, pp. 492–519.

[Bon+19] Xavier Bonnetain et al. "Quantum Attacks Without Superposition Queries: The Offline Simon's Algorithm". In: *ASIACRYPT 2019, Part I*. Ed. by Steven D. Galbraith and Shiho Moriai. Vol. 11921. LNCS. Springer, Heidelberg, Dec. 2019, pp. 552–583.

[Bon17] Xavier Bonnetain. "Quantum Key-Recovery on Full AEZ". In: *SAC 2017*. Ed. by Carlisle Adams and Jan Camenisch. Vol. 10719. LNCS. Springer, Heidelberg, Aug. 2017, pp. 394–406.

[Bon19] Xavier Bonnetain. *Collisions on Feistel-MiMC and univariate GMiMC*. Cryptology ePrint Archive, Report 2019/951. `https://eprint.iacr.org/2019/951`. 2019.

[Bra+02] Gilles Brassard et al. "Quantum Amplitude Amplification and Estimation". In: *Quantum Computation and Information, AMS Contemporary Mathematics 305*. Ed. by Samuel J. Lomo-naco and Howard E. Brandt. 2002.

[DLW19] Xiaoyang Dong, Zheng Li, and Xiaoyun Wang. "Quantum cryptanalysis on some generalized Feistel schemes". In: *SCIENCE CHINA Information Sciences* 62.2 (2019), 22501:1–22501:12.

[HA17] Akinori Hosoyamada and Kazumaro Aoki. "On Quantum Related-Key Attacks on Iterated Even-Mansour Ciphers". In: *IWSEC 17*. Ed. by Satoshi Obana and Koji Chida. Vol. 10418. LNCS. Springer, Heidelberg, Aug. 2017, pp. 3–18.

[HRK20] Samir Hodzic, Lars Knudsen Ramkilde, and Andreas Brasen Kidmose. "On Quantum Distinguishers for Type-3 Generalized Feistel Network Based on Separability". In: *Post-Quantum Cryptography - 11th International Conference, PQCrypto 2020*. Ed. by Jintai Ding and Jean-Pierre Tillich. Springer, Heidelberg, 2020, pp. 461–480.

[II19] Gembu Ito and Tetsu Iwata. "Quantum Distinguishing Attacks against Type-1 Generalized Feistel Ciphers". In: *IACR Cryptology ePrint Archive* 2019 (2019), p. 327.

[Ito+19] Gembu Ito et al. "Quantum Chosen-Ciphertext Attacks Against Feistel Ciphers". In: *CT-RSA 2019*. Ed. by Mitsuru Matsui. Vol. 11405. LNCS. Springer, Heidelberg, Mar. 2019, pp. 391–411.

[Kap+16] Marc Kaplan et al. "Breaking Symmetric Cryptosystems Using Quantum Period Finding". In: *CRYPTO 2016, Part II*. Ed. by Matthew Robshaw and Jonathan Katz. Vol. 9815. LNCS. Springer, Heidelberg, Aug. 2016, pp. 207–237.

[KM10] Hidenori Kuwakado and Masakatu Morii. "Quantum distinguisher between the 3-round Feistel cipher and the random permutation". In: *IEEE International Symposium on Information Theory, ISIT 2010, June 13-18, 2010, Austin, Texas, USA, Proceedings*. 2010, pp. 2682–2685.

[KM12] Hidenori Kuwakado and Masakatu Morii. "Security on the quantum-type Even-Mansour cipher". In: *Proceedings of the International Symposium on Information Theory and its Applications, ISITA 2012, Honolulu, HI, USA, October 28-31, 2012*. 2012, pp. 312–316.

[KNP07] Pascal Koiran, Vincent Nesme, and Natacha Portier. "The quantum query complexity of the abelian hidden subgroup problem". In: *Theor. Comput. Sci.* 380.1-2 (2007), pp. 115–126.

[LM17] Gregor Leander and Alexander May. "Grover Meets Simon - Quantumly Attacking the FX-construction". In: *ASIACRYPT 2017, Part II*. Ed. by Tsuyoshi Takagi and Thomas Peyrin. Vol. 10625. LNCS. Springer, Heidelberg, Dec. 2017, pp. 161–178.

[Mou+14] Nicky Mouha et al. "Chaskey: An Efficient MAC Algorithm for 32-bit Microcontrollers". In: *SAC 2014*. Ed. by Antoine Joux and Amr M. Youssef. Vol. 8781. LNCS. Springer, Heidelberg, Aug. 2014, pp. 306–323.

[MS19] Alexander May and Lars Schlieper. "Quantum Period Finding with a Single Output Qubit - Factoring n-bit RSA with n/2 Qubits". In: *CoRR* abs/1905.10074 (2019). arXiv: `1905.10074`.

[ND19] Boyu Ni and Xiaoyang Dong. “Improved quantum attack on Type-1 Generalized Feistel Schemes and Its application to CAST-256”. In: *IACR Cryptology ePrint Archive* 2019 (2019), p. 318.

[RS15] Martin Roetteler and Rainer Steinwandt. “A note on quantum related-key attacks”. In: *Inf. Process. Lett.* 115.1 (2015), pp. 40–44.

[Sho94] Peter W. Shor. “Algorithms for Quantum Computation: Discrete Logarithms and Factoring”. In: *35th FOCS*. IEEE Computer Society Press, Nov. 1994, pp. 124–134.

[Sim94] Daniel R. Simon. “On the Power of Quantum Computation”. In: *35th FOCS*. IEEE Computer Society Press, Nov. 1994, pp. 116–123.

[SS17] Thomas Santoli and Christian Schaffner. “Using Simon’s algorithm to attack symmetric-key cryptographic primitives”. In: *Quantum Information & Computation* 17.1&2 (2017), pp. 65–78.

A Proof of Lemmas 1 and 2

Lemma 1 (Adapted from [Kap+16, Lemma 1]). *Let $\mathcal{H}$ be a subgroup of $\{0,1\}^n$. Let $y \in \{0,1\}^n$. Then*

$$\sum_{h\in\mathcal{H}} (-1)^{y\cdot h} = \begin{cases} |\mathcal{H}| & \text{if } y \in \mathcal{H}^\perp \\ 0 & \text{otherwise} \end{cases}$$

Proof. Let $h_1, \ldots, h_c$ be a basis of $\mathcal{H}$. Then each element of $\mathcal{H}$ can be uniquely written as $\sum_{i=1}^c \varepsilon_i h_i$. Hence,

$$\sum_{h\in\mathcal{H}} (-1)^{y\cdot h} = \sum_{\varepsilon_1=0}^{1} (-1)^{y\cdot\varepsilon_1 h_1} \cdots \sum_{\varepsilon_c=0}^{1} (-1)^{y\cdot\varepsilon_c h_c} = \prod_{i=1}^{c} \left(1 + (-1)^{y\cdot h_i}\right)$$

This product is nonzero if and only if each $y \cdot h_i$ is equal to 0, that is, $y \in \mathcal{H}^\perp$, in which case it is equal to $2^c = |\mathcal{H}|$. Hence, the lemma holds. □

Lemma 5 (Simon reduction). *Let $n \in \mathbb{N}$, X be a set. There exists a bijection φ_c from $\{f \in \{0,1\}^n \to X | f$ hides a subgroup of dimension $c\}$ to*

$$\{\mathcal{H} \subset \{0,1\}^n | \dim(\mathcal{H}) = c\} \times \left\{ \widehat{f} \in \{0,1\}^{n-c} \to X \middle| \widehat{f} \text{ is aperiodic} \right\}$$

such that, with $\varphi_c(f) = (\mathcal{H}, \widehat{f})$, the behaviour of Simon's algorithm with f and $\widehat{f}$ is identical, up to isomorphism. Moreover, f is a periodic permutation if and only if $\widehat{f}$ is a permutation.

Proof. First of all, a periodic function is uniquely defined by its hidden subgroup $\mathcal{H}$ and its restriction over $\{0,1\}^n/(\mathcal{H})$, which is aperiodic. As $\{0,1\}^n/(\mathcal{H}) \simeq \{0,1\}^{n-c}$, we can define an isomorphism $\psi_\mathcal{H} : \{0,1\}^{n-c} \to \{0,1\}^n/(\mathcal{H})$, and ϕ_c as $\phi_c(f) = (\mathcal{H}, f \circ \psi_\mathcal{H})$. Moreover, by construction, f is a periodic permutation if and only if $\widehat{f}$ is a permutation. Now, it remains to show that this does not impact Simon's algorithm.

After the second set of Hadamard gates, the quantum state in Simon's algorithm when we use f is

$$\frac{1}{2^n} \sum_{x\in\{0,1\}^n} \sum_{y\in\{0,1\}^n} (-1)^{x\cdot y} |y\rangle |f(x)\rangle$$

As for all $h \in \mathcal{H}$, $f(x) = f(x \oplus h)$, this is equal to

$$\frac{1}{2^n} \sum_{x\in\{0,1\}^n/(\mathcal{H})} \sum_{y} (-1)^{x\cdot y} \sum_{h\in\mathcal{H}} (-1)^{y\cdot h} |y\rangle |f(x)\rangle$$

Now, $\sum_{h\in\mathcal{H}} (-1)^{y\cdot h} = |\mathcal{H}| = 2^{\dim(\mathcal{H})}$ if $y \in \mathcal{H}^\perp$, and 0 otherwise. Hence, the state can be rewritten

$$\frac{1}{2^{n-\dim(\mathcal{H})}} \sum_{x\in\{0,1\}^n/(\mathcal{H})} \sum_{y\in\mathcal{H}^\perp} (-1)^{x\cdot y} |y\rangle |f(x)\rangle$$

Now, with $\widehat{f}$, the state can be written as

$$\frac{1}{2^{n-\dim(\mathcal{H})}} \sum_{x \in \{0,1\}^{n-\dim(\mathcal{H})}} \sum_{y \in \{0,1\}^{n-\dim(\mathcal{H})}} (-1)^{x \cdot y} |y\rangle |\widehat{f}(x)\rangle$$

We can use the isomorphism $\psi_{\mathcal{H}}$, and rewrite the state as

$$\frac{1}{2^{n-\dim(\mathcal{H})}} \sum_{x \in \{0,1\}^{n}/(\mathcal{H})} \sum_{y \in \{0,1\}^{n-\dim(\mathcal{H})}} (-1)^{\psi_{\mathcal{H}}(x) \cdot y} |y\rangle |f(x)\rangle$$

Finally, we need to exhibit a dual isomorphism ρ such that $x \cdot y = \psi_{\mathcal{H}}(x) \cdot \rho(y)$. We note $e_1, \ldots, e_\ell$ the canonical basis of $\{0,1\}^{n-\dim(\mathcal{H})}$. Define $f_i = \psi_{\mathcal{H}}(e_i)$, let $h_1, \ldots, h_d$ be a basis of $\mathcal{H}$.

Let g_i be the solution of the system of equations $\forall j, g_i \cdot f_j = \delta_{i,j}$ and $\forall k, g_i \cdot h_k = 0$. As the f_j are a basis of $\{0,1\}^n/(\mathcal{H})$ and the h_k are a basis of $\mathcal{H}$, they are linearly independent. Hence each g_i is uniquely defined. Moreover, the g_i lie in $\mathcal{H}^\perp$, as they are orthogonal to a basis of $\mathcal{H}$. Finally, they are linearly independent. Hence, they form a basis of $\mathcal{H}^\perp$.

Now, define ρ by $\rho(e_i) = g_i$.

By construction, we have $f_i \cdot g_j = \delta_{i,j}$. Hence,

$$\psi_{\mathcal{H}}(x) \cdot \rho(y) = \sum_i x_i f_i \cdot \sum_j y_j g_j = \sum_{i,j} x_i y_j f_i \cdot g_j = \sum_i x_i y_i = x \cdot y$$

□