An Efficient Quantum Factoring Algorithm

Oded Regev[*]

Abstract

We show that n-bit integers can be factorized by independently running a quantum circuit with $\tilde{O}(n^{3/2})$ gates for $\sqrt{n}+4$ times, and then using polynomial-time classical post-processing. The correctness of the algorithm relies on a number-theoretic heuristic assumption reminiscent of those used in subexponential classical factorization algorithms. It is currently not clear if the algorithm can lead to improved physical implementations in practice.

1 Introduction

Shor's celebrated algorithm [Sho99] allows to factorize n-bit integers using a quantum circuit of size (i.e., number of gates) $\tilde{O}(n^2)$. For factoring to be feasible in practice, however, it is desirable to reduce this number further. Indeed, all else being equal, the fewer quantum gates there are in a circuit, the likelier it is that it can be implemented without noise and decoherence destroying the quantum effects.

Here we show that quantum circuits of size $\tilde{O}(n^{3/2})$ are enough. More precisely, we present an algorithm that independently runs $\sqrt{n}+4$ times a quantum circuit with $\tilde{O}(n^{3/2})$ gates. The outputs are then classically post-processed in polynomial time (using a lattice reduction algorithm) to generate the desired factorization.

The quantum circuit size can be made even smaller if super-polynomial-time classical post-processing is allowed. Specifically, for any $0 < \varepsilon \leq 1/2$, it can be brought down to $\tilde{O}(n^{3/2-\varepsilon})$ using classical post-processing (solving a hard lattice problem) running in time $\exp(\tilde{O}(n^{2\varepsilon}))$. The number of times the quantum circuit needs to be applied is still small ($n^{1/2+\varepsilon}$). A curious corollary is that if lattice-based cryptography is broken classically (more precisely, if a polynomial-time classical algorithm exists for hard lattice problems), then quantum circuits of nearly-linear size $\tilde{O}(n)$ are sufficient for factoring integers. This is obtained by taking $\varepsilon = 1/2$ in the previous discussion.

A few remarks are in order. First, our algorithm relies on a number-theoretic heuristic assumption reminiscent of those used in subexponential classical factorization algorithms [CP05]. Second, the number of qubits in our quantum circuit is $O(n^{3/2})$, higher than the $O(n)$ in optimized implementations of Shor's algorithm. Third, the quantum circuit's depth is smaller than the one in Shor's original algorithm by $\tilde{O}(n^{1/2})$. (Note, though, that Shor's algorithm can be implemented using circuits of depth only $O(\log n)$ at the expense of a (much) larger number of gates $\Omega(n^5)$ [CW00].) Finally, we expect similar ideas to apply to the discrete logarithm problem; details will be provided in the full version.

[*] Courant Institute of Mathematical Sciences, New York University. Supported by a Simons Investigator Award from the Simons Foundation.

While our analysis is asymptotic, we expect a significant improvement in the number of gates also for moderately large n, say 2048 bits, possibly by up to two or three orders of magnitude. For such n, fast integer multiplication is typically not advantageous, and one instead uses highly optimized variants of naive integer multiplication, leading to a quantum circuit size of approximately n^3 for Shor's algorithm. Combined with the fact that polynomial-time lattice reduction algorithms provide surprisingly good approximation factors (e.g., 1.01^d for a d-dimensional lattice [GN08]), this suggests that our approach can potentially achieve a circuit size closer to n^2 without fast integer multiplication.

It is important to note, though, that an improvement in the number of gates does not necessarily translate into an improved practical implementation. Indeed, in most architectures currently being considered by industry, the space (or number of qubits) plays an important role. Shor's algorithm is amenable to extensive optimizations, allowing implementations with a very small number of qubits (see [GE21] and references therein). It is currently not clear if our algorithm can benefit from all these optimizations, the main issue being our use of repeated squaring. It therefore remains to be seen whether the algorithm can lead to improved physical implementations in practice.

Statement of the result: Fix some n-bit number $N \leq 2^n$ to be factorized. For some $d > 0$, let $b_1, \dots, b_d$ be some small $O(\log d)$-bit integers (say, b_i is the ith prime number) and let $a_i = b_i^2$. Define the lattice

$$\begin{aligned}\mathcal{L} &= \left\{(z_1, \dots, z_d) \in \mathbb{Z}^d \;\middle|\; \Big(\prod_i b_i^{z_i}\Big)^2 = 1 \bmod N\right\} \\ &= \left\{(z_1, \dots, z_d) \in \mathbb{Z}^d \;\middle|\; \prod_i a_i^{z_i} = 1 \bmod N\right\} \subset \mathbb{Z}^d \,. \end{aligned} \tag{1}$$

Also define the sublattice of $\mathcal{L}$ given by

$$\mathcal{L}_0 = \left\{(z_1, \dots, z_d) \in \mathbb{Z}^d \;\middle|\; \prod_i b_i^{z_i} \in \{-1, 1\} \bmod N\right\} \subseteq \mathcal{L} \,.$$

Assuming N is odd and not a prime power (since factoring is easy otherwise), we heuristically expect at least half the vectors in $\mathcal{L}$ to not be in $\mathcal{L}_0$. For instance, when N is a product of two distinct odd primes, there are 4 square roots of 1 modulo N, so heuristically, half of the vectors in $\mathcal{L}$ should not be in $\mathcal{L}_0$.

Given a vector $z \in \mathcal{L} \setminus \mathcal{L}_0$, we have that $b = \prod_i b_i^{z_i} \bmod N$ is a square root of unity modulo N (because $z \in \mathcal{L}$) yet it is a non-trivial square root of 1, i.e., not equal to ± 1 modulo N (because $z \notin \mathcal{L}_0$). In this case, N divides the product $(b-1)(b+1)$ but does not divide either of the terms, and we therefore must have that $\gcd(b-1, N)$ is a non-trivial factor of N, as desired. Therefore, it suffices to find a vector in $\mathcal{L} \setminus \mathcal{L}_0$.

By the pigeon-hole principle (or Minkowski's first theorem) and using the fact that there are at most $N \leq 2^n$ possible values for the product modulo N in Eq. (1) (i.e., the determinant of $\mathcal{L}$ is at most 2^n), $\mathcal{L}$ is guaranteed to have nonzero vectors of norm at most $\sqrt{d}2^{n/d}$.[1] While we

[1] To see this, consider all vectors $z \in \{-2^{n/d-1}, \dots, 2^{n/d-1}\}^d$. Since there are more than $2^n \geq N$ such vectors, there are two that lead to the same product in Eq. (1). Their difference is therefore in $\mathcal{L}$, and is of norm at most $\sqrt{d}2^{n/d}$.

expect some (in fact, at least half) of them to be in $\mathcal{L} \setminus \mathcal{L}_0$, we do not know how to prove it.[2] Instead, we will make the heuristic assumption that there exists a vector in $\mathcal{L} \setminus \mathcal{L}_0$ of norm at most $T = \exp(O(n/d))$. With this assumption, the algorithm is guaranteed to provide a factorization of N, as in our main result, stated next.

Theorem 1.1. *Let N be an n-bit number and assume that for $d = \sqrt{n}$ and $O(\log n)$-bit numbers $b_1, \dots, b_d$, there exists a vector in $\mathcal{L} \setminus \mathcal{L}_0$ of norm at most $T = \exp(O(\sqrt{n}))$. Then, there is a classical polynomial-time algorithm that outputs a non-trivial factor of N using $\sqrt{n} + 4$ calls to a quantum circuit of size $O(n^{3/2} \log n)$.*

Related work: The idea of reducing the cost of the quantum circuit at the expense of applying it several times independently and classically post-processing the outputs was already suggested by Seifert [Sei01] (see also [EH17]). However, the improvement obtained by his algorithm is only by a constant factor.

Acknowledgements: The author is grateful to Martin Ekerå, Craig Gidney, Minki Hhan, Igor Shparlinski, Noah Stephens-Davidowitz, Thomas Vidick, and Ronald de Wolf for their comments on an early draft.

2 High level overview of the algorithm

The algorithm can be thought of as a multidimensional analogue of Shor's algorithm. At the core of the algorithm is the quantum procedure presented in Section 3. It starts by creating a quantum superposition over $|z\rangle$ for $z \in \mathbb{Z}^d$. For convenience we use a discrete Gaussian state of some radius R. It then uses an elementary classical procedure (described below) in superposition to compute $\prod_i a_i^{z_i} \bmod N$ in a new register. This register will be ignored, so effectively can be thought of as being measured. This creates an $\mathcal{L}$-periodic state over the $|z\rangle$ register, or more precisely, a superposition over a random coset of $\mathcal{L}$, truncated at radius R. We then apply the quantum Fourier transform and measure. This results in vectors from the dual lattice $\mathcal{L}^*$. However, because the original $\mathcal{L}$-periodic state only extends up to radius R, what we obtain is an *approximation* of vectors in $\mathcal{L}^*$ up to error roughly $1/R$. The reader might recall a similar phenomenon occurring in Shor's algorithm, and the use of continued fractions there to recover the exact periodicity. In Section 4 we show how to use these approximations to vectors in $\mathcal{L}^*$ to recover a basis of a lattice $\mathcal{L}'$ that is essentially the same as $\mathcal{L}$ for vectors up to norm roughly $2^{-n/d}R$. Recalling that we heuristically expect to have vectors in $\mathcal{L} \setminus \mathcal{L}_0$ of norm at most $T = \exp(O(n/d))$, we then use the LLL algorithm (which achieves an approximation factor of 2^d) to obtain a vector in $\mathcal{L} \setminus \mathcal{L}_0$ of norm at most $2^d \cdot T$. For this to work, we need this norm to be below the bound of $2^{-n/d}R$, i.e., we need $R > 2^{d+n/d} \cdot T$. By taking $d = \sqrt{n}$, it suffices to take $R = \exp(C\sqrt{n})$ for some constant $C > 0$.

As in Shor's algorithm, the quantum circuit size required for this algorithm is dominated by the classical exponentiation procedure. It is here that we benefit from the fact that the a_i are small numbers. Indeed, with R as above, we can compute $\prod_i a_i^{z_i} \bmod N$ using a circuit of size only

[2] It was, however, communicated to us by I. Shparlinski that assuming the Generalized Riemann Hypothesis, when N is not a prime power and does not have any small prime divisors, $d = \tilde{O}(n^2)$, and b_i is the ith prime number, $\mathcal{L} \setminus \mathcal{L}_0$ contains a vector with all coordinates in $\{0, 1\}$, and in particular, of norm at most $\sqrt{d}$. Unfortunately, this does not seem strong enough for an improved algorithmic result.

$\tilde{O}(n^{3/2})$. The elementary but crucial idea is to perform all multiplications on the small numbers a_i directly, so that the only operations we have to perform on large n-bit numbers are $\log_2 R = O(\sqrt{n})$ squaring operations.

3 The quantum procedure

Fix $d > 0$, $a_1, \ldots, a_d$, and $\mathcal{L}$ as before. Let $R > \sqrt{2d}$ be another parameter to be chosen later, and take D to be a power of 2 somewhat larger than R; say, $D \in [2\sqrt{d} \cdot R, 4\sqrt{d} \cdot R)$. We will show a quantum procedure that outputs a uniform sample from $\mathcal{L}^*/\mathbb{Z}^d$ (a set of cardinality $\det \mathcal{L}$) perturbed by Gaussian noise of roughly $1/R$ and discretized to the grid $\{0, 1/D, \ldots, (D-1)/D\}^d$. In more detail, the output will be within statistical distance $1/\mathrm{poly}(d)$ of the distribution Q supported on $\{0, 1/D, \ldots, (D-1)/D\}^d$ whose mass at point w is equal to

$$Q(w) := (\det \mathcal{L})^{-1} \sum_{v \in \mathcal{L}^*/\mathbb{Z}^d} Q_v(w) \; , \tag{2}$$

where for $v \in \mathcal{L}^*/\mathbb{Z}^d$, Q_v is the probability distribution defined as

$$Q_v(w) := \frac{\rho_{1/\sqrt{2}R}(v - w + \mathbb{Z}^d)}{\rho_{1/\sqrt{2}R}(v - D^{-1}\mathbb{Z}^d)} \; . \tag{3}$$

Here, we use the Gaussian function $\rho_s : \mathbb{R}^d \to \mathbb{R}$ defined for $s > 0$ as

$$\rho_s(z) = \exp(-\pi \|z\|^2/s^2) \; .$$

In other words, a sample from Q can be described as the output of the following process: let $v \in \mathcal{L}^*/\mathbb{Z}^d$ be a uniform coset; then, output a sample from the Gaussian distribution on $\{0, 1/D, \ldots, (D-1)/D\}^d$ whose mass at point w is proportional to $\rho_{1/\sqrt{2}R}(v - w + \mathbb{Z}^d)$. Importantly, as we show in Claim A.7, with all but probability $O(2^{-d})$, $\mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(w, v) \le \sqrt{d}/(\sqrt{2}R)$, where $\mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(w, v) := \min_{z \in \mathbb{Z}^n} \mathrm{dist}(w, v + z)$ denotes distance in the torus, i.e., modulo 1.

The quantum procedure starts by approximating to within $1/\mathrm{poly}(d)$ the state proportional to

$$\sum_{z \in \{-D/2, \ldots, D/2-1\}^d} \rho_R(z) |z\rangle \; , \tag{4}$$

similarly to how it was done in, e.g., [Reg09]. Namely, to generate this "discrete Gaussian state", first note that it can be written as the tensor product of d copies of the one-dimensional state ($d = 1$). It therefore suffices to generate the one-dimensional state. To do this, we use a standard technique which basically puts each qubit from the most-significant to the least-significant into the appropriate superposition of $|0\rangle$ and $|1\rangle$ conditioned on the values of the previous qubits [GR02]. To obtain a small circuit size, notice that beyond the $O(\log d)$ most significant qubits, all remaining qubits are within distance $1/\mathrm{poly}(d)$ of the "plus state" $(|0\rangle + |1\rangle)/\sqrt{2}$, no matter what values we condition on. (This uses that $D = O(\mathrm{poly}(d) \cdot R)$ and that ρ_R changes slowly, namely, that for all $z \in \{-D/2, \ldots, D/2 - 1\}$ and $0 \le k \le D/\mathrm{poly}(d)$, $\rho_R(z)$ is within $1 \pm 1/\mathrm{poly}(d)$ of $\rho_R(z + k)$.) We can therefore simply initialize the remaining qubits to the plus state using one Hadamard gate per qubit. In summary, we can approximate the state in (4) using a quantum circuit of size only

$d(\log D + \mathrm{poly}(\log d))$, where the $\mathrm{poly}(\log d)$ term is for computing the rotation needed for each of the $O(\log d)$ most significant qubits, and the $\log D$ term is due to the Hadamard gates on the remaining qubits.

The next step is the most costly one. Here we apply a classical procedure in superposition in order to compute the value $\prod_i a_i^{z_i+D/2} \bmod N$ into a new register $|e\rangle$. (We added $D/2$ for convenience so we do not need to worry about negative exponents.) Notice that $h(z) := \prod_i a_i^{z_i} \bmod N$ is a homomorphism from $\mathbb{Z}^d$ to $\mathbb{Z}_N^*$, the multiplicative group of integers modulo N, and that its kernel is $\mathcal{L}$. Therefore, there is a bijection between $\mathbb{Z}^d/\mathcal{L}$ and the image of h. As a result, since $|e\rangle$ will be ignored, we can equivalently write the resulting state up to normalization as

$$\sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{z \in (\mathcal{L}+e) \cap [-D/2, D/2)^d} \rho_R(z) |z\rangle |e\rangle \ . \tag{5}$$

To compute $\prod_i a_i^{z_i+D/2} \bmod N$, first notice that when all the exponents are in $\{0,1\}$, we can compute the product of the d numbers in a binary tree fashion, leading to the recurrence $T(d) = 2T(d/2) + M(d \log d)$, where $M(k)$ is the number of gates needed to compute the product of two k-bit numbers, and here we are using that $a_1, \dots, a_d$ are all small $O(\log d)$ bit numbers. Using fast integer multiplication [HvdH21], $M(k) = O(k \log k)$, leading to a circuit of size $O(d \log^3 d)$. The general case of exponents in $\{0, \dots, D-1\}$ can be handled using a repeated squaring-like idea. More specifically, for $j = 0, \dots, \lfloor \log_2(D-1) \rfloor$, let z_{ij} denote the jth bit of $z_i + D/2$, with $j = 0$ being the most significant. Then, letting e be a register initialized to 1, we do the following for $j = 0, \dots, \lfloor \log_2(D-1) \rfloor$: square e, then compute the product of the subset of the a_i determined by the z_{ij}, and multiply e by the result. To summarize, the circuit size needed for this step is $O(\log D \cdot (d \log^3 d + n \log n))$, where we use that e is an n-bit number and can therefore be squared in time $O(n \log n)$.

In the final step, we apply the quantum Fourier transform (over $\mathbb{Z}_D^d$) to the $|z\rangle$ register, and then output the vector in $\{0, 1/D, \dots, (D-1)/D\}^d$ obtained by measuring that register and dividing by D. As we show using a standard calculation in Appendix A (specifically, in Claim A.5 and Proposition A.6), the resulting distribution is within $O(2^{-d})$ distance of the distribution Q, as desired. The circuit size needed for this step is only $O(d \log D \cdot \log((\log D)/\varepsilon))$ by using approximate QFT with error ε [Cop02]. Taking $\varepsilon = 1/\mathrm{poly}(d)$, we get circuit size $O(d \cdot \log D \cdot (\log \log D + \log d))$.

To summarize, the quantum procedure uses a circuit of size

$$O(\log D \cdot (d \log^3 d + d \log \log D + n \log n) + d \cdot \mathrm{poly}(\log d)) \tag{6}$$

and outputs a point $w \in [0,1)^d$ that is within distance $\sqrt{d}/(\sqrt{2}R)$ of a uniformly chosen $v \in \mathcal{L}^*/\mathbb{Z}^d$ (with all but probability $1/\mathrm{poly}(d)$).

4 Recovering a lattice from noisy samples of the dual

Theorem 4.1 ([Pom01]). *Suppose G is a finite abelian group with minimal number of generators r. Then, when choosing elements from G independently and uniformly, the expected number of elements needed to generate G is less than $r + \sigma$, where $\sigma = 2.118456563\dots$.*

Corollary 4.2. *In the setting of Theorem 4.1, $r + 4$ uniformly random elements of G generate G with probability at least $1/2$.*

Proof. Otherwise, with probability at least $1/2$, $r+5$ elements are needed to generate G. Since this random variable is never smaller than r by assumption, its expectation is at least $r + 5/2 > r + \sigma$, in contradiction. □

Lemma 4.3. *Let $\mathcal{L} \subset \mathbb{Z}^d$, $m \geq d+4$, and assume $v_1, \ldots, v_m$ are uniformly chosen cosets from $\mathcal{L}^*/\mathbb{Z}^d$. With probability at least $1/4$, it holds that for all nonzero $u \in \mathbb{Z}^d/\mathcal{L}$, there exists an i such that $\langle u, v_i \rangle \notin [-\varepsilon, \varepsilon] \bmod 1$, where $\varepsilon = (4 \det \mathcal{L})^{-1/m}/3$.*

Proof. First, notice that the group $\mathcal{L}^*/\mathbb{Z}^d$ can be generated by at most d elements, e.g., by taking a basis of $\mathcal{L}^*$. Therefore, by Corollary 4.2, with probability at least $1/2$, $v_1, \ldots, v_m$ generate $\mathcal{L}^*/\mathbb{Z}^d$. Assume that this is the case. Fix some nonzero $u \in \mathbb{Z}^d/\mathcal{L}$, and consider the distribution of $\langle u, v \rangle \bmod 1$ where v is uniformly chosen from $\mathcal{L}^*/\mathbb{Z}^d$. This distribution is not identically zero (as otherwise u would be the zero coset $\mathcal{L}$). In fact, it must be equal to the uniform distribution over the set $\{0, 1/t, 2/t, \ldots, (t-1)/t\}$ for some $t \geq 2$. This follows, e.g., from the invariance of the uniform distribution over $\mathcal{L}^*/\mathbb{Z}^d$ to shifts by elements from that same group. If $t < 1/\varepsilon$ then by our assumption, there must exist an i such that $\langle u, v_i \rangle \neq 0 \bmod 1$ which in particular implies that $\langle u, v_i \rangle \notin [-\varepsilon, \varepsilon] \bmod 1$. Otherwise, assume $t \geq 1/\varepsilon$ and notice that for any fixed i, the probability of $\langle u, v_i \rangle \in [-\varepsilon, \varepsilon] \bmod 1$ is

$$(1 + 2\lfloor t\varepsilon \rfloor)/t \leq 3\varepsilon \ .$$

Therefore, the probability that $\langle u, v_i \rangle \in [-\varepsilon, \varepsilon] \bmod 1$ for all $i \in \{1, \ldots, m\}$ is at most $(3\varepsilon)^m$. We complete the proof by applying the union bound over all $\det \mathcal{L} - 1$ nonzero elements u in $\mathbb{Z}^d/\mathcal{L}$. □

Lemma 4.4. *Let $\mathcal{L} \subset \mathbb{Z}^d$ and $m \geq d+4$. Assume $v_1, \ldots, v_m$ are uniformly chosen cosets from $\mathcal{L}^*/\mathbb{Z}^d$. For some $\delta > 0$ let $w_1, \ldots, w_m \in [0,1)^d$ satisfy that $\mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(w_i, v_i) < \delta$ for all i. For some $S > 0$, define the $d+m$-dimensional lattice $\mathcal{L}'$ generated by the columns of*

$$B = \left(\begin{array}{c|c} I_{d \times d} & 0 \\ \hline S \cdot w_1 & \\ \vdots & S \cdot I_{m \times m} \\ S \cdot w_m & \end{array} \right) .$$

Then, for any $u \in \mathcal{L}$, there exists a vector $u' \in \mathcal{L}'$ whose first d coordinates are equal to u and whose norm is at most $\|u\| \cdot (1 + m \cdot S^2 \cdot \delta^2)^{1/2}$. Moreover, with probability at least $1/4$ (over the choice of the v_i), any nonzero $u' \in \mathcal{L}'$ of norm $\|u'\| < \min(S, \delta^{-1}) \cdot \varepsilon/2$ satisfies that its first d coordinates are a nonzero vector in $\mathcal{L}$, where $\varepsilon = (4 \det \mathcal{L})^{-1/m}/3$.

Proof. Take any $u \in \mathcal{L}$. Then, for any i, $\langle u, v_i \rangle = 0 \bmod 1$ (since $v_i \in \mathcal{L}^*/\mathbb{Z}^d$), and therefore, by Cauchy-Schwarz, $\langle u, w_i \rangle$ is within distance $\delta\|u\|$ of an integer. The claim now follows by taking the combination of the first d columns of B given by the coordinates of u and taking an appropriate combination of the remaining m columns of B to make the last m coordinates of the resulting vector at most $S\delta\|u\|$ in absolute value. Next, assume $v_1, \ldots, v_m$ satisfy the conclusion of Lemma 4.3, which happens with probability at least $1/4$. Take any nonzero $u' \in \mathcal{L}'$ and let $u \in \mathbb{Z}^d$ be its first d coordinates. If $u = 0$ then clearly $\|u'\| \geq S$, as desired. Assume therefore that u is not in $\mathcal{L}$, or equivalently, that the coset $u+\mathcal{L}$ is not the zero element in $\mathbb{Z}^d/\mathcal{L}$. If $\|u\| \geq \varepsilon/(2\delta)$, we are done, so assume $\|u\| < \varepsilon/(2\delta)$. By Lemma 4.3, there exists an i such that $\langle u, v_i\rangle$ is at least ε away from an integer. By Cauchy-Schwarz, this implies that $\langle u, w_i\rangle$ is at least

$$\varepsilon - \delta\|u\| > \varepsilon/2$$

away from an integer. As a result, u' has a coordinate of absolute value at least $S\varepsilon/2$, as desired. □

For convenience, we record here the special case of $m = d+4$ and $S = \delta^{-1}$.

Corollary 4.5. *Let $\mathcal{L} \subset \mathbb{Z}^d$ and $v_1, \ldots, v_{d+4}$ be uniformly chosen cosets from $\mathcal{L}^*/\mathbb{Z}^d$. For some $\delta > 0$ let $w_1, \ldots, w_{d+4} \in [0,1)^d$ satisfy that $\mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(w_i, v_i) < \delta$ for all i. Let $\mathcal{L}'$ be the $(2d+4)$-dimensional lattice defined in Lemma 4.4 (with $S = \delta^{-1}$). Then, for any $u \in \mathcal{L}$, there exists a vector $u' \in \mathcal{L}'$ whose first d coordinates are equal to u and whose norm is at most $\|u\| \cdot (d+5)^{1/2}$. Moreover, with probability at least $1/4$ (over the choice of the v_i), any nonzero $u' \in \mathcal{L}'$ of norm $\|u'\| < \delta^{-1} \cdot (4 \det \mathcal{L})^{-1/(d+4)}/6$ satisfies that its first d coordinates are a nonzero vector in $\mathcal{L}$.*

5 Proof of the main theorem

Claim 5.1. *There is an efficient classical algorithm that given a basis of a lattice $\mathcal{L} \subset \mathbb{R}^k$ and some norm bound $T > 0$, outputs a list of $\ell \leq k$ vectors $z_1, \ldots, z_\ell \in \mathcal{L}$ of norm at most $\sqrt{k}2^{k/2}T$ with the property that any vector in $\mathcal{L}$ of norm at most T must be an integer combination of them. In other words, the sublattice they generate contains all the vectors in $\mathcal{L}$ of norm at most T.*

Proof. The algorithm starts by computing an LLL reduced basis of $\mathcal{L}$, which we denote by $z_1, \ldots, z_k$. It next computes the Gram-Schmidt orthogonalization of the basis, denoted by $\tilde{z}_1, \ldots, \tilde{z}_k$. Finally, let $\ell \geq 0$ be smallest such that $\|\tilde{z}_{\ell+1}\| \geq 2^{k/2}T$ (or k if no such index exists), and output the vectors $z_1, \ldots, z_\ell$.

By properties of LLL reduced bases, we have that for all $i = 1, \ldots, k-1$, $\|\tilde{z}_{i+1}\| \geq \|\tilde{z}_i\|/\sqrt{2}$. It follows that for $i = \ell+1, \ldots, k$, $\|\tilde{z}_i\| > T$. As a result, any vector in $\mathcal{L}$ of norm at most T must be an integer combination of $z_1, \ldots, z_\ell$. We complete the proof by recalling that an LLL reduced basis is also "size reduced," implying that

$$\|z_i\|^2 \leq \sum_{j=1}^{i} \|\tilde{z}_j\|^2 < k2^kT^2 \; .$$

□

Proof of Theorem 1.1. Take $d = \sqrt{n}$ and $R = \exp(C\sqrt{n})$ for a large enough constant $C > 0$. With these parameters, and recalling that $D \in [2\sqrt{d} \cdot R, 4\sqrt{d} \cdot R)$, the quantum procedure's circuit size in (6) becomes

$$O(n^{3/2} \log n)$$

and its output is a point w within distance $\delta = \sqrt{d}/(\sqrt{2}R)$ of a uniformly chosen $v \in \mathcal{L}^*/\mathbb{Z}^d$ (with all but probability $1/\mathrm{poly}(d)$). Apply the quantum procedure $d+4$ times independently to obtain such vectors $w_1, \ldots, w_{d+4}$.

Consider the $(2d+4)$-dimensional lattice $\mathcal{L}'$ given in Corollary 4.5. By our assumption, there exists a vector in $u' \in \mathcal{L}'$ of norm at most $(d+5)^{1/2} \cdot T$ whose first d coordinates are a nonzero vector $u \in \mathcal{L} \setminus \mathcal{L}_0$. We next apply the classical algorithm in Claim 5.1 to $\mathcal{L}'$ with the norm bound $(d+5)^{1/2} \cdot T$. As its output, we obtain vectors $z'_1, \ldots, z'_\ell$ of norm at most

$$(2d+4)^{1/2} 2^{d+2} (d+5)^{1/2} \cdot T < \delta^{-1} (4 \det \mathcal{L})^{-1/(d+4)}/6 ,$$

where the inequality follows since $\det \mathcal{L} \leq N \leq 2^n$ and by choosing R large enough. As a result, by the second property in Corollary 4.5, except with probability 1/4, if we denote the first d coordinates of z'_i by z_i, we have that $z_i \in \mathcal{L}$ for all i. Moreover, at least one of the z_i must not be in $\mathcal{L}_0$, otherwise u, which is an integer combination of the z_i (since u' is an integer combination of the z'_i) would also be in $\mathcal{L}_0$. Finally, we apply for each of the z_i the gcd calculation outlined in Section 1. Since there exists an i such that $z_i \in \mathcal{L} \setminus \mathcal{L}_0$, one of these calculations will yield a non-trivial factor of N, as desired. □

The extension to super-polynomial classical post-processing mentioned in the introduction is similar. One needs to take $d = n^{1/2+\varepsilon}$, $R = \exp(Cn^{1/2-\varepsilon})$, and use the fact that for all $0 < \delta < 1$, one can approximate lattice problems in dimension d to within $\exp(d^{1-\delta})$ in time $\exp(\tilde{O}(d^{\delta}))$ [GN08].

A Fourier transform calculation

We will need the following useful fact by Banaszczyk. It shows that for any lattice $\mathcal{L}$, almost all the Gaussian mass in ρ_s is given by points of norm at most $\sqrt{d}s$. Here and elsewhere, for a set $A \subset \mathbb{R}^n$, we use the notation $f(A)$ to denote the sum $\sum_{x \in A} f(x)$.

Lemma A.1 ([Ban93])**.** *For any d-dimensional lattice $\mathcal{L}$, $x \in \mathbb{R}^n$, and $s > 0$,*

$$\rho_s(\{y \in \mathcal{L} + x \mid \|y\| > \sqrt{d}s\}) < 2^{-d} \cdot \rho_s(\mathcal{L}) .$$

Corollary A.2. *If $\mathcal{L}$ is a lattice containing no nonzero vectors of norm at most $\sqrt{d}s$ then $\rho_s(\mathcal{L} \setminus \{0\}) \leq 2 \cdot 2^{-d}$.*

Proof. Using Lemma A.1 with $x = 0$, write

$$\rho_s(\mathcal{L} \setminus \{0\}) = \rho_s(\{y \in \mathcal{L} \mid \|y\| > \sqrt{d}s\}) < 2^{-d} \cdot (1 + \rho_s(\mathcal{L} \setminus \{0\}))$$

and rearrange. □

We will use the following formulation of the Poisson summation formula. Here, $\hat{f}$ denotes the Fourier transform of f. For instance, $\widehat{\rho_s} = s^n \rho_{1/s}$.

Lemma A.3. *(Poisson summation formula) For any lattice $\mathcal{L}$ and any (nice enough) function $f : \mathbb{R}^n \to \mathbb{C}$,*

$$f(\mathcal{L}) = \det(\mathcal{L}^*) \hat{f}(\mathcal{L}^*) ,$$

where $\hat{f}$ denotes the Fourier transform of f.

Let $|\varphi_1\rangle$ be the state in Eq. (5) which, to recall, is given by

$$|\varphi_1\rangle = Z_1^{-1} \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{z \in (\mathcal{L}+e) \cap \{-D/2, \ldots, D/2-1\}^d} \rho_R(z) |z\rangle |e\rangle ,$$

where $Z_1 > 0$ is the normalization term.

Claim A.4. *We have that $Z_1^2 \in [1 \pm 2 \cdot 2^{-d}](R/\sqrt{2})^d$.*

Proof. Notice that

$$Z_1^2 = \sum_{z \in \{-D/2,\dots,D/2-1\}^d} \rho_R(z)^2 = \sum_{z \in \{-D/2,\dots,D/2-1\}^d} \rho_{R/\sqrt{2}}(z) \ .$$

Therefore, by Lemma A.1 (with $x = 0$) and using $D/2 \geq \sqrt{d}R/\sqrt{2}$, Z_1^2 satisfies

$$(1 - 2^{-d}) \cdot \rho_{R/\sqrt{2}}(\mathbb{Z}^d) \leq Z_1^2 \leq \rho_{R/\sqrt{2}}(\mathbb{Z}^d) \ .$$

By the Poisson summation formula, $\rho_{R/\sqrt{2}}(\mathbb{Z}^d)$ is equal to $(R/\sqrt{2})^d \rho_{\sqrt{2}/R}(\mathbb{Z}^d)$. Moreover, since $1 \geq \sqrt{d} \cdot \sqrt{2}/R$, Corollary A.2 shows that $\rho_{\sqrt{2}/R}(\mathbb{Z}^d) \in [1, 1 + 2 \cdot 2^{-d}]$. □

Also define the state

$$|\varphi_2\rangle = Z_2^{-1} \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{z \in \mathbb{Z}_D^d} \rho_R\Big((z + D\mathbb{Z}_d) \cap (\mathcal{L} + e)\Big) |z\rangle |e\rangle \ ,$$

where, again, $Z_2 > 0$ is the normalization term. In other words, whereas in $|\varphi_1\rangle$ we truncate the discrete Gaussian to the box $\{-D/2, \dots, D/2 - 1\}^d$, in $|\varphi_2\rangle$ we let it wrap around modulo D. In the next claim, we show that the two states are very close, which intuitively follows from the fact that the Gaussian mass does not extend much beyond radius $\sqrt{d}R$.

Claim A.5. *We have that*

$$\| |\varphi_1\rangle - |\varphi_2\rangle \|_2 \leq 2 \cdot 2^{-d} \ .$$

Moreover, $Z_1/Z_2 \in [1 \pm 2^{-d}]$.

Proof. Let $\mathcal{M}$ be the $2d$-dimensional lattice given by all vectors $(z_1, z_2) \in \mathbb{Z}^{2d}$ such that both $z_1 = z_2 \bmod D$ and $z_1 = z_2 \bmod \mathcal{L}$. Then notice that

$$\begin{aligned} Z_2^2 &= \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{z \in \mathbb{Z}_D^d} \rho_R\Big((z + D\mathbb{Z}_d) \cap (\mathcal{L} + e)\Big)^2 \\ &= \sum_{\substack{z_1, z_2 \in \mathbb{Z}^d \\ z_1 = z_2 \bmod D \\ z_1 = z_2 \bmod \mathcal{L}}} \rho_R(z_1) \cdot \rho_R(z_2) \\ &= \rho_R(\mathcal{M}) \ . \end{aligned}$$

Similarly, denoting by $\mathcal{M}'$ the subset of $\mathcal{M}$ corresponding to all vectors (z_1, z_2) such that neither z_1 nor z_2 are in $\{-D/2, \dots, D/2 - 1\}^d$,

$$\begin{aligned} \|Z_1|\varphi_1\rangle - Z_2|\varphi_2\rangle\|_2^2 &= \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{z \in \mathbb{Z}_D^d} \rho_R\Big(((z + D\mathbb{Z}_d) \cap (\mathcal{L} + e)) \setminus \{-D/2, \dots, D/2 - 1\}^d\Big)^2 \\ &= \rho_R(\mathcal{M}') \\ &\leq 2^{-2d} \cdot \rho_R(\mathcal{M}) \\ &= 2^{-2d} \cdot Z_2^2 \ , \end{aligned}$$

where we used Lemma A.1 (with $x = 0$) and the fact that all vectors in $\mathcal{M}'$ are of norm at least $D/\sqrt{2} \geq \sqrt{2d}R$. Dividing both sides by Z_2^2, we get

$$\|(Z_1/Z_2)|\varphi_1\rangle - |\varphi_2\rangle\|_2 \leq 2^{-d} .$$

By the triangle inequality, this implies that $Z_1/Z_2 = \|(Z_1/Z_2) \cdot |\varphi_1\rangle\|_2 \in [1 \pm 2^{-d}]$. Using the triangle inequality again, we get that

$$\||\varphi_1\rangle - |\varphi_2\rangle\|_2 \leq \||\varphi_1\rangle - (Z_1/Z_2) \cdot |\varphi_1\rangle\|_2 + \|(Z_1/Z_2) \cdot |\varphi_1\rangle - |\varphi_2\rangle\|_2 \leq 2 \cdot 2^{-d} ,$$

as desired. □

Proposition A.6. *The distribution obtained by applying QFT to $|\varphi_2\rangle$, discarding the e register, measuring the z register, and dividing the result by D is within statistical distance $O(2^{-d})$ of the distribution Q defined in Eq. (2).*

Proof. The QFT of $|\varphi_2\rangle$ is given by

$$\begin{aligned} &D^{-d/2} \cdot Z_2^{-1} \cdot \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{w \in \{0,1/D,\ldots,(D-1)/D\}^d} \sum_{z \in \mathbb{Z}_D^d} \exp(2\pi i \langle w, z\rangle) \rho_R\Big((z + D\mathbb{Z}_d) \cap (\mathcal{L} + e)\Big)|w\rangle|e\rangle \\ &= D^{-d/2} \cdot Z_2^{-1} \cdot \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{w \in \{0,1/D,\ldots,(D-1)/D\}^d} \sum_{z \in \mathcal{L}+e} \exp(2\pi i \langle w, z\rangle) \rho_R(z)|w\rangle|e\rangle . \end{aligned}$$

Using the Poisson summation formula, this is equal to

$$R^d \cdot D^{-d/2} \cdot (\det \mathcal{L})^{-1} \cdot Z_2^{-1} \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{w \in \{0,1/D,\ldots,(D-1)/D\}^d} \sum_{v \in \mathcal{L}^*} \exp(2\pi i \langle v, e\rangle) \rho_{1/R}(v - w)|w\rangle|e\rangle .$$

Therefore, after measuring (and discarding) e, the probability of measuring w is

$$P(w) := R^{2d} \cdot D^{-d} \cdot (\det \mathcal{L})^{-2} \cdot Z_2^{-2} \cdot \sum_{e \in \mathbb{Z}^d/\mathcal{L}} \Big| \sum_{v \in \mathcal{L}^*} \exp(2\pi i \langle v, e\rangle) \rho_{1/R}(v - w)\Big|^2 .$$

The sum can be simplified as

$$\begin{aligned} &\sum_{e \in \mathbb{Z}^d/\mathcal{L}} \sum_{v_1, v_2 \in \mathcal{L}^*} \exp(2\pi i \langle v_1 - v_2, e\rangle) \rho_{1/R}(v_1 - w) \rho_{1/R}(v_2 - w) \\ &\quad = \det \mathcal{L} \cdot \sum_{\substack{v_1, v_2 \in \mathcal{L}^* \\ v_1 = v_2 \bmod \mathbb{Z}^d}} \rho_{1/R}(v_1 - w) \rho_{1/R}(v_2 - w) \\ &\quad \geq \det \mathcal{L} \cdot \sum_{v \in \mathcal{L}^*} \rho_{1/R}(v - w)^2 \\ &\quad = \det \mathcal{L} \cdot \sum_{v \in \mathcal{L}^*/\mathbb{Z}^d} \rho_{1/(\sqrt{2}R)}(v - w + \mathbb{Z}^d) , \end{aligned}$$

where in the first equality we used that $\sum_{e \in \mathbb{Z}^d/\mathcal{L}} \exp(2\pi i \langle v_1 - v_2, e\rangle)$ is equal to $\det \mathcal{L}$ if $v_1 = v_2 \bmod \mathbb{Z}^d$ and is zero otherwise, and the inequality is simply by discarding terms where $v_1 \neq v_2$.

Therefore, $P(w)$ is bounded from below by

$$\begin{aligned} P'(w) &:= R^{2d} \cdot D^{-d} \cdot (\det \mathcal{L})^{-1} \cdot Z_2^{-2} \cdot \sum_{v \in \mathcal{L}^*/\mathbb{Z}^d} \rho_{1/(\sqrt{2}R)}(v - w + \mathbb{Z}^d) \\ &= \sum_{v \in \mathcal{L}^*/\mathbb{Z}^d} \alpha_v Q_v(w) \ , \end{aligned}$$

where Q_v is the probability distribution defined in Eq. (3), and α_v is given by

$$\begin{aligned} \alpha_v &= R^{2d} \cdot D^{-d} \cdot (\det \mathcal{L})^{-1} \cdot Z_2^{-2} \cdot \rho_{1/(\sqrt{2}R)}(v + D^{-1}\mathbb{Z}^d) \\ &= (R/\sqrt{2})^d \cdot (\det \mathcal{L})^{-1} \cdot Z_2^{-2} \cdot \sum_{x \in D\mathbb{Z}^d} \exp(2\pi i \langle v, x \rangle) \rho_{\sqrt{2}R}(x) \\ &\in (R/\sqrt{2})^d \cdot (\det \mathcal{L})^{-1} \cdot Z_2^{-2} \cdot [1 \pm 2 \cdot 2^{-d}] \ , \end{aligned} \tag{7}$$

where we used the Poisson summation formula, and in the last step we applied Corollary A.2 (using $D \geq \sqrt{2d}R$) and the triangle inequality. From Claim A.4 and the second part of Claim A.5, we get that Z_2^2 is within $1 \pm O(2^{-d})$ of $(R/\sqrt{2})^d$, and so $\alpha_v \in (\det \mathcal{L})^{-1} \cdot [1 \pm O(2^{-d})]$. This establishes that P' is within ℓ_1 distance $O(2^{-d})$ of Q. In particular, its ℓ_1 norm (i.e., total mass) is at least $1 - O(2^{-d})$, which, combined with it being a lower bound on P, implies that it is within ℓ_1 distance $O(2^{-d})$ of P. Using the triangle inequality,

$$\|P - Q\|_1 \leq \|P - P'\|_1 + \|P' - Q\|_1 = O(2^{-d}) \ ,$$

completing the proof. □

Claim A.7. *For any $v \in \mathbb{R}^d/\mathbb{Z}^d$, if w is chosen from the distribution Q_v defined in Eq. (3), then the probability that $\mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(w, v) > \sqrt{d}/(\sqrt{2}R)$ is at most $O(2^{-d})$.*

Proof. Using Lemma A.1,

$$\begin{aligned} &\rho_{1/\sqrt{2}R}(\{x \in v - D^{-1}\mathbb{Z}^d \mid \mathrm{dist}_{\mathbb{R}^d/\mathbb{Z}^d}(x, 0) > \sqrt{d}/(\sqrt{2}R)\}) \\ &\quad \leq \rho_{1/\sqrt{2}R}(\{x \in v - D^{-1}\mathbb{Z}^d \mid \|x\| > \sqrt{d}/(\sqrt{2}R)\}) \\ &\quad < 2^{-d} \rho_{1/\sqrt{2}R}(D^{-1}\mathbb{Z}^d) \\ &\quad \leq (1 + O(2^{-d})) 2^{-d} \rho_{1/\sqrt{2}R}(v - D^{-1}\mathbb{Z}^d) \ , \end{aligned}$$

where in the last inequality we used Eq. (7) (specifically, that $\alpha_0 \leq (1 + O(2^{-d}))\alpha_v$). □

References

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993. 8

[Cop02] Don Coppersmith. An approximate Fourier transform useful in quantum factoring, 2002. quant-ph/0201067. 5

[CP05] Richard Crandall and Carl Pomerance. *Prime numbers*. Springer, New York, second edition, 2005. A computational perspective. 1

[CW00] Richard Cleve and John Watrous. Fast parallel circuits for the quantum Fourier transform. In *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)*, pages 526–536. IEEE Comput. Soc. Press, Los Alamitos, CA, 2000. 1

[EH17] Martin Ekerå and Johan Håstad. Quantum algorithms for computing short discrete logarithms and factoring RSA integers. In *Post-quantum cryptography*, volume 10346 of *Lecture Notes in Comput. Sci.*, pages 347–363. Springer, Cham, 2017. 3

[GE21] Craig Gidney and Martin Ekerå. How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, April 2021. 2

[GN08] Nicolas Gama and Phong Q. Nguyen. Finding short lattice vectors within Mordell's inequality. In *STOC'08*, pages 207–216. ACM, New York, 2008. 2, 8

[GR02] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions, 2002. arXiv:quant-ph/0208112. 4

[HvdH21] David Harvey and Joris van der Hoeven. Integer multiplication in time $O(n \log n)$. *Ann. of Math. (2)*, 193(2):563–617, 2021. 5

[Pom01] Carl Pomerance. The expected number of random elements to generate a finite abelian group. *Period. Math. Hungar.*, 43(1-2):191–198, 2001. 5

[Reg09] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):Art. 34, 40, 2009. 4

[Sei01] Jean-Pierre Seifert. Using fewer qubits in Shor's factorization algorithm via simultaneous Diophantine approximation. In *Topics in cryptology—CT-RSA 2001 (San Francisco, CA)*, volume 2020 of *Lecture Notes in Comput. Sci.*, pages 319–327. Springer, Berlin, 2001. 3

[Sho99] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.*, 41(2):303–332, 1999. 1