Shor's Algorithm Does Not Factor Large Integers in the Presence of Noise

Jin-Yi Cai
University of Wisconsin-Madison
jyc@cs.wisc.edu

(Extended Abstract)

DRAFT

Abstract

We consider Shor's quantum factoring algorithm in the setting of noisy quantum gates. Under a generic model of random noise for (controlled) rotation gates, we prove that the algorithm does not factor integers of the form pq when the noise exceeds a vanishingly small level in terms of n — the number of bits of the integer to be factored, where p and q are from a well-defined set of primes of positive density. We further prove that with probability $1 - o(1)$ over random prime pairs (p, q), Shor's factoring algorithm does not factor numbers of the form pq, with the same level of random noise present.

1 Introduction

One of the most stunning achievements of computer science in the last several decades is Shor's quantum algorithm to factor large integers [36, 37]. The algorithm provably can factor an n-bit integer in polynomial time with high probability, assuming certain quantum operations can be performed. These are called quantum logic gates. In particular, they include the familiar Hadamard gate $H = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, the rotation gates (Phase) $S = \left[\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix}\right]$, ($\pi/8$ gate) $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{2\pi i/8} \end{smallmatrix}\right]$, and more generally $R_k = \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{smallmatrix}\right]$, and their controlled versions. Note that $S = R_2$ and $T = R_3$.

It has often been pointed out that the availability of these quantum gates at high precision (with arbitrarily small angles in R_k with $k \to \infty$) is a challenge, both intellectually and practically on engineering grounds [16, 23, 24, 1]. To a large extent, such concerns motivated another great intellectual achievement that is the development of quantum error correcting codes [38, 40, 8, 19, 32]. There is a substantial body of work on fault tolerant quantum computing, starting with Shor's work [39]. Strong threshold theorems are proved which show that in certain error models, if the error rate is below a certain threshold, quantum computation can achieve arbitrarily high accuracy [3, 17, 6, 19, 41, 18, 5]. These are beautiful mathematical theorems. But they fundamentally assume that the group SU(2) *exactly* corresponds to operations on a qubit in reality, especially in its composition—that group composition (in its infinite precision defined over $\mathbb{C}$) exactly corresponds to sequential application of realizable quantum operations. Opinions differ, as to whether such arbitrary precision is ever achievable. It is certainly a possibility. However, this author is skeptical about this, based on the belief that quantum mechanics itself (just as any other physical theory) is not, and is not meant to be, infinitely accurate when comparing reality with what the mathematical statements say in the theory (some speculations are in Section 4). Meanwhile, enormous efforts have been underway in the past few decades, and with much renewed momentum and enthusiasm more recently, to achieve ever increasingly accurate hardware implementations of quantum circuitry.

In this paper, we consider Shor's quantum factoring algorithm in the setting where each quantum controlled rotation gate is subject to a small random noise in the angle. We assume each application of the controlled-R_k gate is given an independent random error of angle $e^{2\pi i \epsilon r/2^k}$. Thus, when the control bit is 1, the operator R_k is substituted by $\widetilde{R_k} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & e^{2\pi i(1+\epsilon r)/2^k} \end{smallmatrix}\right]$, where r is an *independent* noise random variable distributed $r \sim N(0,1)$, and ϵ is a global magnitude parameter. So, the controlled-$\widetilde{R_k}$ gate is $\left[\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \rho_k \xi_k \end{smallmatrix}\right]$, where $\rho_k = \rho_{k,\epsilon} = e^{2\pi i \epsilon r/2^k}$ and $\xi_k = e^{2\pi i/2^k}$. We show that there exist positive constants $c, c' > 0$ such that if $\epsilon > cn^{-1/3}$, then Shor's algorithm does not factor n-bit integers of the form pq, where p and q are from a well-defined set of primes of density $> c'$. To the best knowledge of this author, this is the first provable statement of such failure of Shor's algorithm under any error model.

The noise model is similar to that of [30] (see also [14, 29, 31]). The specific random noise model including the independent normal distribution picked in this paper is not essential, as the proof will clearly show, but it is chosen to present the essential idea of the proof most transparently. For example, the noise r.v. r being distributed $\sim N(0,1)$ (standard normal distribution) can be replaced by any reasonable alternative distribution such as uniform $U[-1,1]$ or uniform bits from $\{-1, 0, 1\}$.

While each individual controlled-R_k gate is assumed to be accompanied by an independent r.v. r for noise, when an individual controlled-R_k gate is applied, the same randomly perturbed controlled-R_k gate is applied to each term in a sum of superpositions of quantum states. Regarding the random noise model, we do not make any claim that this model accurately reflects "reality"; our purpose is only to show that *some* vanishing amount of noise can already provably destroy the algorithm.

An important modification of Shor's algorithm by Coppersmith [10] shows that if we just ignore (not to perform) all (controlled-) R_k-gates for sufficiently large $k \geq b$, where b is some global parameter, then Shor's algorithm still retains its effectiveness (and uses a reduced number of quantum gates). The specific suggested change [10] for 500-qubits, which would require rotations of magnitude $2\pi/2^{500}$ in Shor's original algorithm, is to ignore all rotations of angle smaller than $2\pi/2^{20}$. It is estimated that this would incur an error on the order of 1% in the probability of each desirable final state. Asymptotically, Coppersmith improves the precision requirement of exponentially small angles to just slightly less than π/n. This is of enormous practical implications. This version of Shor's algorithm is called the "banded" version with parameter b, which is set to be slightly greater than $\log n$, rather than n in the original version. Nonetheless, rotation gates (as primitive steps of the algorithm) of asymptotically infinitely small angles would still be required as n, the number of bits to be factored, tends to infinity.

Our result is consistent with Coppersmith's improvement. Indeed we will present our proof in the "banded" version, with perfect controlled-R_k-gates for all $k < b$, but every controlled-R_k-gate is replaced by a controlled-$\widetilde{R_k}$-gate for all $k \geq b$, i.e., it is independently perturbed by a random noise. Our negative result will be stated in terms of $b + \log_2(1/\epsilon)$. When $b + \log_2(1/\epsilon) < \frac{1}{3}\log_2 n - c$ for some constant $c > 0$, the noise takes hold so as to destroy the desired peak in the probability of observing a useful state that leads to factorization. This condition is essentially equivalent to having both b being less than a small constant multiple of $\log n$ and ϵ greater than the reciprocal of a small positive power of n. We prove that, under this condition in this noise model, Shor's algorithm does not factor n-bit integers of the form pq, where p and q are from a well-defined set of primes of positive density $c' > 0$. [1] The proof will in fact show that, the same result holds under the same condition $b + \log_2(1/\epsilon) < \frac{1}{3}\log_2 n - c$, even if the noise gates are applied only at the single level R_b, with all other controlled-R_k-gates applied perfectly for $k \neq b$ (*or alternatively*, no controlled-R_k-gates are applied at all for $k > b$ as in the banded version by Coppersmith).

Theorem 1. *There exist constants $c, c' > 0$, such that if each controlled-R_k-gate in the quantum Fourier transform circuit is replaced by controlled-$\widetilde{R_k}$-gate for all $k \geq b$, where $b + \log_2(1/\epsilon) < \frac{1}{3}\log_2 n - c$, then with exponentially small exceptional probability, Shor's algorithm does not factor n-bit integers of the form pq, where p and q are from a well-defined set of primes of density $> c'$.*

Here "exceptional probability" is over the random choices of Shor's algorithm as well as proba-

[1] We note that the results from [14, 29, 30, 31] are generally stated in the opposite direction. Under plausible, but ultimately heuristic, assumptions for the bahavior of various sums, augmented by numerical simulations, they suggest that if b is not too large compared to n, Shor's algorithm *can* tolerate imprecisions of rotation angles. Some small concrete values of n are on the order of 10-qubits ($n = 10, 14$). These values are quite outside the range where our proof applies. Their numerical simulation does seem to suggest a logarithmic threshold of b. Thus, these positive results are not logically inconsistent with, and in fact, complement our proof. N.B. the notation b in [29] is our $b - 2$.

bilistic outcomes of quantum measurements. More precisely, the expectation over random noise r's, of the success probability (over the random choices of the algorithm and quantum measurements) of the algorithm is exponentially small in n. This will be the meaning of "does not factor" below.

Theorem 2. *If $b + \log_2(1/\epsilon) < \frac{1}{3}\log_2 n - c$, then the statement in Theorem 1 still holds, if only each controlled-R_b-gate is replaced by a controlled-$\widetilde{R_b}$-gate while all other controlled-R_k-gates remain unchanged. Alternatively, the same statement holds if each controlled-R_k-gate is: (1) applied perfectly for $k < b$, (2) replaced by a controlled-$\widetilde{R_b}$-gate for $k = b$, and (3) deleted for $k > b$.*

Our proof focuses on the essential "period-finding" part using quantum Fourier transform (QFT) in Shor's algorithm. In our proof, we use a theorem of Fouvry [13]. This theorem states that the set of all primes p such that the largest prime factor in $p - 1$ is greater than $p^{2/3}$ has positive density among all primes. We use this theorem to produce candidate inputs of the form $N = pq$ to Shor's algorithm where p and q are of this type, and argue that a random element $x \in \mathbb{Z}_N^*$ has (exponentially) large order $\omega = \omega_N(x)$ as an element of the multiplicative group $\mathbb{Z}_N^*$. This large order ω allows us to give a lower bound for a lattice counting argument, which leads to a sufficiently large number of independent perturbations in the complex arguments (in the exponent) in a crucial sum of exponentials, (which would have been a perfect geometric sum without noise) in the analysis of Shor's algorithm. This perturbation, at the appropriate setting of parameters, destroys this geometric sum, and degrades the probability of observing any useful quantum state to *negligible*, and thus fails to gain any useful information on the period ω.

Our proof is actually more generally applicable. In an appendix we prove the following theorem:

Theorem 3. *There exists a constant $c > 0$, such that for random primes p and q chosen uniformly from all primes of binary length m, if $b + \log_2(1/\epsilon) < \frac{1}{3}\log_2 m - c$, as $m \to \infty$ with probability $1 - o(1)$, Shor's algorithm with noisy rotation gates does not factor $N = pq$.*

A version analogous to Theorem 2 also holds for random primes.

We make a few brief remarks. Arguably, factoring integers $N = pq$ for random primes p and q is more important in cryptography than for primes that satisfy the property in Fouvry's theorem, and the statement of failure probability being $1 - o(1)$ is stronger than that of positive density guaranteed by Fouvry's theorem. We present the proof in the main text for the latter, and relegate the proof of Theorem 3 to the appendix, in order to concentrate on the main idea of how random noise degrades the performance of Shor's algorithm. The additional work needed for Theorem 3 is mainly of a number theoretic nature, and for the purpose of this paper, of secondary importance. Also, one can prove other versions of Theorem 3. E.g., we can restrict the random primes p and q to be of length m *and* both $\equiv 3 \bmod 4$, so that the numbers $N = pq$ are the so-called Blum integers, which are favored in cryptography [28]. Despite the strong failure demonstrated by the proof, our theorems do not rule out the possibility that at some future time, quantum algorithm is superior to the best "classical" factoring algorithms for factoring integers of a certain size, in practice. But our proof indicates that there is a limit to this possible superiority when n is large, if arbitrarily small random noise cannot be eliminated.

Many people have made strong arguments [32] supporting the viewpoint that Shor's algorithm presents a convincing evidence that the so-called *Strong* Church-Turing thesis needs a necessary modification. This *Strong* thesis identifies efficient computation with P or BPP. The argued-for modification states that this should be replaced by BQP. This author is personally not convinced of this. I will make some comments at the end of this paper. [2]

2 Preliminaries

Fouvry's theorem Let $N = pq$, where p and q are distinct odd primes. By the Chinese remainder theorem, the multiplicative group $\mathbb{Z}_N^* = \{m \in \mathbb{Z}_N \mid \gcd(m, N) = 1\}$ (invertible elements in $\mathbb{Z}_N$) is isomorphic to the direct product $\mathbb{Z}_p^* \times \mathbb{Z}_q^*$. Moreover, $\mathbb{Z}_p^*$ is a cyclic group of order $p-1$, and is isomorphic to a direct product of factors according to the prime factorization of $p-1$; and similarly for $\mathbb{Z}_q^*$. If $p - 1 = 2^e p_1^{e_1} \cdots p_k^{e_k}$, where $p_1 < \ldots < p_k$ are distinct odd primes, then $\mathbb{Z}_p^*$ is isomorphic to $\mathbb{Z}_{2^e} \times \mathbb{Z}_{p_1^{e_1}} \times \cdots \times \mathbb{Z}_{p_k^{e_k}}$. Let $P^+(m)$ denote the largest prime in the prime factorization of m.

Theorem 4 (Fouvry)**.** *There exist constants $c > 0$ and $n_0 > 0$, such that for all $x > n_0$,*

$$|\{p \mid p \textit{ is a prime, } p < x, \textit{ and } P^+(p-1) > p^{2/3}\}| \geq c\frac{x}{\log x}.$$

We say a prime p satisfies the Fouvry property if $P^+(p-1) > p^{2/3}$. If $N = pq$, where p and q are distinct odd primes satisfying the Fouvry property, then clearly $p' = P^+(p-1)$ appears with exponent 1 in the factorization of $p-1$, and so does $P^+(q-1)$ in the factorization of $q-1$. If $p' = P^+(p-1) > P^+(q-1)$, then $\mathbb{Z}_{p'}$ appears as an isolated factor in the direct product form of $\mathbb{Z}_N^*$. Thus, with probability $\geq 1 - 1/p' > 1 - \frac{1}{\max\{p^{2/3}, q^{2/3}\}} \geq 1 - N^{-1/3}$, a random element x in $\mathbb{Z}_N^*$ has order at least $p' > \max\{p^{2/3}, q^{2/3}\} \geq N^{1/3}$. If it so happens that $p' = P^+(p-1) = P^+(q-1)$, then $\mathbb{Z}_{p'} \times \mathbb{Z}_{p'}$ appears as a factor in the direct product form of $\mathbb{Z}_N^*$. In this case, a random element x in $\mathbb{Z}_N^*$ also has order at least $p' > N^{1/3}$ with probability $\geq 1 - 1/(p')^2 \geq 1 - N^{-2/3}$. Thus, in either case, in terms of the number of bits, such products $N = pq$ have the property that a random element x in $\mathbb{Z}_N^*$ has an exponentially large period, $\omega = \omega_N(x) \geq \max\{P^+(p-1), P^+(q-1)\} > N^{1/3}$, with exponentially small exceptional probability. Below we assume ω has this property.

Denote by $\mathrm{ord}_2(x)$ the highest power of 2 that divides x. If $e = \mathrm{ord}_2(p-1)$, and $e' = \mathrm{ord}_2(q-1)$, then we have $2^e < (P^+(p-1))^{1/2}$ and $2^{e'} < (P^+(q-1))^{1/2}$, and thus $\omega = \omega_N(x)$ satisfies $\mathrm{ord}_2(\omega) \leq \max\{e, e'\} < \frac{\log_2 \omega}{2}$, for any $x \in \mathbb{Z}_N^*$. We conclude:

Lemma 5. *Let p and q be distinct odd primes satisfying the Fouvry property, and let $N = pq$, then over a random $x \in \mathbb{Z}_N^*$,*

$$\Pr.\left(\omega_N(x) > N^{1/3} \quad \text{and} \quad \mathrm{ord}_2(\omega_N(x)) < \frac{\log_2 \omega_N(x)}{2}\right) > 1 - \frac{1}{N^{1/3}}.$$

Sum of random unit vectors Let $\xi_m = e^{2\pi i/m}$ be a primitive root of unity of order m. Let $X_i \sim N(0, 1)$, $i = 1, 2, \ldots, n$, be a finite sequence of independent and identically distributed (i.i.d.) normal random variables. Let $\{S_k \subseteq [n] \mid 1 \leq k \leq K\}$ be a finite collection of sets such that each

[2] These comments are speculative, and should not be conflated with the theorems proved in the paper.

pairwise symmetric difference $S_j \Delta S_k$ has cardinality $\geq m^2 t$, for all $j \neq k$. Let $\Sigma_k = \sum_{i \in S_k} X_i$ be the sum of X_i in S_k. We will give a simple estimate for the expectation of

$$|\xi_m^{\Sigma_1} + \xi_m^{\Sigma_2} + \ldots + \xi_m^{\Sigma_K}|^2. \tag{1}$$

Expanding the square norm expression we get

$$K + \sum_{1 \leq j < k \leq K} (\xi_m^{\Sigma_j - \Sigma_k} + \xi_m^{\Sigma_k - \Sigma_j}) = K + 2 \sum_{1 \leq j < k \leq K} \cos\left((\Sigma_j - \Sigma_k) \frac{2\pi}{m} \right).$$

Let $T_{jk} = (\Sigma_j - \Sigma_k)\frac{2\pi}{m}$. Note that $\Sigma_j - \Sigma_k = \sum_{i \in S_j \Delta S_k} (\pm X_i)$ is a sum of at least m^3 distinct (thus independent) r.v. $\pm X_i$ distributed i.i.d. $\sim N(0,1)$. Therefore, each T_{jk} is a random variable normally distributed $\sim N(0, \sigma_{jk}^2)$, with standard deviation $\sigma_{jk} = \sqrt{|S_j \Delta S_k|} \cdot \frac{2\pi}{m} \geq 2\pi\sqrt{t}$.

Moments of even orders of a normal random variable $Y \sim N(0, \sigma^2)$ are known as follows [33]

$$\mathbf{E}[Y^{2k}] = \sigma^{2k}(2k-1)!!,$$

from which we get (by the dominated convergence theorem, the exchange of orders of summation and integration is justified)

$$\begin{aligned} \mathbf{E}[\cos(T_{jk})] &= 1 - \frac{\sigma_{jk}^2}{2!}(2-1)!! + \frac{\sigma_{jk}^4}{4!}(4-1)!! - \frac{\sigma_{jk}^6}{6!}(6-1)!! + \ldots \\ &= e^{-\sigma_{jk}^2/2} \\ &\leq e^{-2\pi^2 t}. \end{aligned}$$

Hence, the expectation of (1) is at most $K + 2\binom{K}{2} e^{-2\pi^2 t}$.

We will need a slight generalization of this. Let $\sigma > 0$, and let $\varphi_k \in [0, 2\pi)$ be any angle, $1 \leq k \leq K$. We replace each Σ_k by $\varphi_k + \sigma \sum_{i \in S_k} X_i$. Then,

Lemma 6. *Let $\sigma > 0$ and $\xi_m = e^{2\pi i/m}$. Let $X_i \sim N(0,1)$, i.i.d. for $i = 1, 2, \ldots, n$, and let $\{S_k \subseteq [n] \mid 1 \leq k \leq K\}$ be a finite collection of sets. Assume, all except at most δ fraction of pairwise symmetric differences $S_j \Delta S_k$ have cardinality $\geq (m/\sigma)^2 t$ for $j \neq k$. Let $\Sigma_k = \varphi_k + \sigma \sum_{i \in S_k} X_i$, where $\varphi_k \in [0, 2\pi)$. Then,*

$$\mathbf{E}[|\xi_m^{\Sigma_1} + \xi_m^{\Sigma_2} + \ldots + \xi_m^{\Sigma_K}|^2] \leq K + 2\delta \binom{K}{2} + 2(1-\delta)\binom{K}{2} e^{-2\pi^2 t}.$$

Proof. Let $T_{jk} = \frac{2\pi\sigma}{m}(\sum_{i \in S_j} X_i - \sum_{i \in S_k} X_i)$. We only need to note in addition to the above that

$$\cos(\varphi + T_{jk}) = \cos\varphi \cos T_{jk} - \sin\varphi \sin T_{jk},$$

and we have $\cos\varphi \leq 1$ for any φ, and $\mathbf{E}[\sin T_{jk}] = 0$ since sin is an odd function and T_{jk} is symmetrically distributed. The lemma follows. □

3 Corrupted geometric sums

Suppose N is an integer we wish to factor, and $2^n \approx N^2$ as in [37] [3]. For definiteness assume $2^{n-1} < N^2 \le 2^n$. Assume ω is the period of the function $f(k) = x^k \bmod N$ for a randomly chosen $x \in \mathbb{Z}_N^*$, and by Lemma 5 we assume $\omega > N^{1/3}$ and $\mathrm{ord}_2(\omega) < \frac{\log_2 \omega}{2}$. Also $\omega < N$ clearly.

Let us write out a few terms as the controlled-R_k gates are applied successively in the QFT circuit (e.g., see [32] p.219), but now with random noise added whenever the controlled rotation gate is R_k-gates with $k \ge b$, i.e., we apply controlled-R_k-gates when $k < b$ but controlled-$\widetilde{R_k}$-gates for all $k \ge b$. (As the first controlled-R_k-gate has $k = 2$, we have $b > 1$.) Suppose we start with the state $|u\rangle = |u_{n-1} \dots u_1 u_0\rangle$. After the first gate H on the qubit $|u_{n-1}\rangle$, we have the state

$$\frac{1}{2^{1/2}} \left(|0\rangle + e^{2\pi i \ 0.u_{n-1}}|1\rangle\right) |u_{n-2} \dots u_0\rangle.$$

The next is the controlled-R_2-gate on target qubit $|u_{n-2}\rangle$ controlled by the left most qubit (which was initially $|u_{n-1}\rangle$), after which we have (assuming $b > 2$)

$$\frac{1}{2^{1/2}} \left(|0\rangle + e^{2\pi i \ 0.u_{n-1}u_{n-2}}|1\rangle\right) |u_{n-2} \dots u_0\rangle.$$

The random noise starts at the controlled-R_b-gate, after which we get

$$\frac{1}{2^{1/2}} \left(|0\rangle + e^{2\pi i \ \left[0.u_{n-1}\dots u_{n-b} + \frac{\epsilon}{2^b} u_{n-b} r_0^{(0)}\right]}|1\rangle\right) |u_{n-2} \dots u_0\rangle,$$

where $r_0^{(0)} \sim N(0,1)$.

After all the rotation gates controlled by the left most qubit (initially $|u_{n-1}\rangle$) we have

$$\frac{1}{2^{1/2}} \left(|0\rangle + e^{2\pi i \ \left[0.u_{n-1}\dots u_0 + \frac{\epsilon}{2^b} \left(u_{n-b} r_0^{(0)} + \frac{u_{n-b-1} r_1^{(0)}}{2} + \dots + \frac{u_0 r_{n-b}^{(0)}}{2^{n-b}}\right)\right]}|1\rangle\right) |u_{n-2} \dots u_0\rangle, \tag{2}$$

where $r_0^{(0)}, \dots, r_{n-b}^{(0)}$ are i.i.d. $\sim N(0,1)$.

Then, similarly, after all the rotation gates controlled by the two left most qubits (initially $|u_{n-1}u_{n-2}\rangle$) we have

$$\frac{1}{2^{2/2}} \left(|0\rangle + e^{2\pi i \ \left[0.u_{n-1}\dots u_0 + \frac{\epsilon}{2^b} \left(u_{n-b} r_0^{(0)} + \dots + \frac{u_0 r_{n-b}^{(0)}}{2^{n-b}}\right)\right]}|1\rangle\right) \otimes \left(|0\rangle + e^{2\pi i \ \left[0.u_{n-2}\dots u_0 + \frac{\epsilon}{2^b} \left(u_{n-b-1} r_0^{(1)} + \dots + \frac{u_0 r_{n-b-1}^{(1)}}{2^{n-b-1}}\right)\right]}|1\rangle\right) |u_{n-3} \dots u_0\rangle, \tag{3}$$

where $r_0^{(0)}, \dots, r_{n-b}^{(0)}, r_0^{(1)}, \dots, r_{n-b-1}^{(1)}$ are i.i.d. $\sim N(0,1)$.

The circuit continues to apply controlled rotation gates with random noise starting at the controlled-R_b-gate, producing a final expression with n tensor factors. When written out the tensor product, this is a sum indexed by $|v_{n-1} \dots v_0\rangle$, such that $v_0 = 0$ or 1 corresponds to selecting respectively the term $|0\rangle$ or $e^{2\pi i \ [\cdots]}|1\rangle$ in (2) (or equivalently, to selecting one of the two terms in

[3] Thus N has $\approx n/2$ bits, a slight change in notation from Section 1.

the first tensor factor in (3)), and $v_1 = 0$ or 1 corresponds to selecting respectively the term $|0\rangle$ or $e^{2\pi i\ [\cdots]}|1\rangle$ in the second tensor factor in (3), and similarly for $v_s = 0$ or 1, for all $0 \le s \le n-1$.

The crucial step in Shor's algorithm, after the quantum Fourier transform, is to take a quantum measurement, with the property that the probability of observing a state that is close to an integral multiple of $\frac{2^n}{\omega}$ is high. Such a state has an n-bit integer expression $v \in \{0,1\}^n$ that has value close to the rational number $\frac{2^n}{\omega}j$, for some $0 \le j \le \omega$. States $|v\rangle$ such that the number v is not close to an integral multiple of $\frac{2^n}{\omega}$ have negligible probability of being observed, while states in a small vicinity of each of the integral multiple of $\frac{2^n}{\omega}$ get observed with probability on the order of $1/\omega$ (per each multiple), and these add up to give a good probability that some such state is observed, whereby the period is deduced with good probability. (I am omitting steps of the continued fraction algorithm in the post quantum processing steps.)

For each v, the probability of $|v\rangle$ being observed has an expression as a square norm of a sum over a set of the form $u \in \{u^* + k\omega : k \ge 0, \text{ and } u^* + k\omega < 2^n\}$ (for some initial $0 \le u^* < \omega$), with cardinality K, which is approximately $2^n/\omega$. For $u^{(k)} = u^* + k\omega$, we write the n-bit integers $u^{(k)} = \sum_{s=0}^{n-1} u_s^{(k)} 2^s$ and $v = \sum_{s=0}^{n-1} v_s 2^s$. When there is no noise in the controlled-R_k-gates used in the QFT this probability expression for observing $|v\rangle = |v_{n-1} \dots v_1 v_0\rangle$ can be written as

$$\frac{1}{2^n K} \left| \sum_{k=0}^{K-1} \exp\left\{ 2\pi i \sum_{t=1}^{n} \frac{\sum_{s=0}^{n-t} u_{n-t-s}^{(k)} v_s}{2^t} \right\} \right|^2 .$$

With independent random noise present starting with controlled-R_b-gates, this becomes

$$\frac{1}{2^n K} \left| \sum_{k=0}^{K-1} \exp\left\{ 2\pi i \left[\sum_{t=1}^{n} \frac{\sum_{s=0}^{n-t} u_{n-t-s}^{(k)} v_s}{2^t} + \frac{\epsilon}{2^b} \left\{ \left(u_{n-b}^{(k)} r_0^{(0)} + \cdots + \frac{u_0^{(k)} r_{n-b}^{(0)}}{2^{n-b}} \right) v_0 + \left(u_{n-b-1}^{(k)} r_0^{(1)} + \cdots + \frac{u_0^{(k)} r_{n-b-1}^{(1)}}{2^{n-b-1}} \right) v_1 + \dots + u_0^{(k)} r_0^{(n-b)} v_{n-b} \right\} \right] \right\} \right|^2 , \tag{4}$$

where

$$r_0^{(0)}, \dots, r_{n-b}^{(0)}, r_0^{(1)}, \dots, r_{n-b-1}^{(1)}, \dots, r_0^{(n-b-1)}, r_1^{(n-b-1)}, r_0^{(n-b)}$$

are random variables i.i.d. $\sim N(0,1)$.

Our first goal is to show that among states $|v\rangle$ such that the binary number v is close to an integral multiple $\frac{2^n}{\omega}j$ (for some $0 \le j \le \omega$), it is the case that for most j, a linear number of bits in the binary expansion of v are one: $v_s = 1$. This will leave us with a linear number of terms of the form in the exponent

$$\frac{2\pi i \epsilon}{2^b} \left(u_{n-b-s}^{(k)} r_0^{(s)} + \frac{u_{n-b-s-1}^{(k)} r_1^{(s)}}{2} + \cdots + \frac{u_0^{(k)} r_{n-b-s}^{(s)}}{2^{n-b-s}} \right) v_s .$$

Eventually we will show that, fixing any such v, among those s where $v_s = 1$, for most k, there are a linear number of terms with $u_{n-b-s}^{(k)} = 1$, which will give us the perturbation as a sum of $\frac{2\pi i \epsilon}{2^b} \cdot r_0^{(s)}$.

Let us consider integers $v = \lfloor \frac{2^n}{\omega} j \rfloor$, for $0 \le j < \omega$; it will be clear from the proof below that what is proved is also true for any v in the vicinity of a polynomial range of such a number.

For $0 \le j < \omega$, the integer $v = \lfloor \frac{2^n}{\omega} j \rfloor$ has the i-th leading bit $v_{n-i} = 1$ iff the i-th most significant bit, among the first n bits, in the binary expansion of $\frac{j}{\omega}$ is 1. This is true iff for some $1 \le k \le 2^{i-1}$,

$$\frac{2k-1}{2^i} \le \frac{j}{\omega} < \frac{2k}{2^i},$$

which is equivalent to

$$(2k-1)\frac{\omega}{2^i} \le j < 2k\frac{\omega}{2^i}. \tag{5}$$

So, j needs to be placed in the alternate ("odd" indexed) segments of length $\frac{\omega}{2^i}$. This is a lattice counting problem.

Recall that $\omega > N^{1/3} \approx 2^{n/6}$. We take $i_0 = \lfloor \frac{3}{4}\log_2 \omega \rfloor \ge \lfloor \frac{1}{4}\log_2 N \rfloor = \Omega(n)$. Then $\frac{\omega}{2^{i_0}} \ge \omega^{1/4} > N^{1/12} = 2^{\Omega(n)}$. We will only count those i-th (significant) bits v_{n-i} that are one, within $1 \le i \le i_0$, and first show that for most j, even just among the first i_0 bits $v_{n-1}, \dots, v_{n-i_0}$, there are a linear number of ones. (Any additional bits that are 1 can only add more noise to the perturbation.)

Now we divide the range $[0, \omega)$ of real numbers into 2^{i_0} segments of equal length $\frac{\omega}{2^{i_0}}$

$$I_\alpha = \{x \in \mathbb{R} \mid \frac{\omega}{2^{i_0}}(\alpha)_2 \le x < \frac{\omega}{2^{i_0}}((\alpha)_2 + 1)\},$$

where $\alpha \in \{0,1\}^{i_0}$ is a binary string, and $(\alpha)_2$ is the binary number it represents [4].

Note that any real interval of the form $[A, A+B)$ has either $\lfloor B \rfloor$ or $\lfloor B \rfloor + 1$ many integers. Thus, each I_α contains either $\lfloor \frac{\omega}{2^{i_0}} \rfloor$ or $\lfloor \frac{\omega}{2^{i_0}} \rfloor + 1$ many integers, which is $\frac{\omega}{2^{i_0}} + \eta$ for some $-1 \le \eta \le 1$. We consider two distributions on the integers $0 \le j < \omega$. Let Pr. denote the uniform distribution and let $\Pr_\alpha$ denote the distribution induced by first picking $\alpha \in \{0,1\}^{i_0}$ uniformly, and then picking $j \in I_\alpha$ uniformly. They are exponentially close: For any $0 \le j < \omega$, $\Pr.(j) = 1/\omega$, and

$$\Pr_\alpha(j) = \frac{1}{2^{i_0}} \frac{1}{\frac{\omega}{2^{i_0}} + \eta} = \frac{1}{\omega + \eta 2^{i_0}} = \frac{1}{\omega} \cdot \frac{1}{1 + \eta \frac{2^{i_0}}{\omega}} = \left(1 \pm 2^{-\Omega(n)}\right) \cdot \Pr.(j). \tag{6}$$

Let $\alpha = \alpha_1 \alpha_2 \dots \alpha_{i_0}$. Consider any $j \in I_\alpha$. If $\alpha_{i_0} = 1$, then j satisfies the equation (5) for $i = i_0$. Now suppose $\alpha_{i_0 - 1} = 1$, then

$$\frac{\omega}{2^{i_0}}(\alpha_1 \dots \alpha_{i_0-1}\alpha_{i_0})_2 \ge \frac{\omega}{2^{i_0}}(\alpha_1 \dots \alpha_{i_0-1}0)_2 = \frac{\omega}{2^{i_0-1}}(\alpha_1 \dots \alpha_{i_0-1})_2$$

and

$$\frac{\omega}{2^{i_0}}((\alpha_1 \dots \alpha_{i_0-1}\alpha_{i_0})_2 + 1) \le \frac{\omega}{2^{i_0-1}} \frac{(\alpha_1 \dots \alpha_{i_0-1}0)_2 + 2}{2} = \frac{\omega}{2^{i_0-1}}((\alpha_1 \dots \alpha_{i_0-1})_2 + 1).$$

And so clearly j satisfies the equation (5) for $i = i_0 - 1$.

Similarly, we can see that every $j \in I_\alpha$ satisfies the equation (5) for every $i \in \{1, \dots, i_0\}$ such that the corresponding bit in α is 1. For any constant $0 < \delta < 1/2$, the proportion of 0-1 sequences of length i_0 that have δi_0 ones is asymptotically $2^{-(1-H(\delta))i_0}$, where $H(\cdot)$ is the entropy function. For any fixed constant $c > 0$, consider any $J = \{i : i_0' \le i \le i_0\}$ with length $i_0 - i_0' + 1 \ge cn$ indexing bit positions $\alpha_{i_0'}, \dots, \alpha_{i_0}$. Then, for a random $\alpha \in \{0,1\}^{i_0}$, with exponentially small exceptional probability $2^{-\Omega(n)}$, there are $\Omega(n)$ bits $\alpha_i = 1$ in those bit positions $i \in J$. Then any $j \in I_\alpha$ gives the corresponding bit $v_{n-i} = 1$. By (6) this is true under the uniform distribution Pr. for j as well. It follows that with exponentially small exceptional probability $2^{-\Omega(n)}$, a uniformly chosen j defines a number $v = \lfloor \frac{2^n}{\omega} j \rfloor$ with a linear number of bit $v_{n-i} = 1$, for $i \in J$.

[4] The reason we cut off at i_0 is to avoid having to deal with intervals that are too small and such odd indexed segments may just miss most integers. We can afford to cut off at i_0, and still get a linear number $\Omega(n)$ of 1's in the first i_0 bits of the n bit binary expansion. This is where we use the fact that ω is large.

Lemma 7. *For any fixed constant $c > 0$ and any $J = \{i : i'_0 \leq i \leq i_0\}$ with length $i_0 - i'_0 + 1 \geq cn$, picking a random $0 \leq j < \omega$ uniformly which defines $v = \lfloor \frac{2^n}{\omega} j \rfloor$,*

$$\text{Pr.}\left(|\{i \in J : v_{n-i} = 1\}| \geq \Omega(n)\right) = 1 - 2^{-\Omega(n)}.$$

Now back to the expression (4) for the probability of observing $|v\rangle$ when noise is present. Regardless what values

$$\sum_{t=1}^{n} \frac{\sum_{s=0}^{n-t} u^{(k)}_{n-t-s} v_s}{2^t}, \quad \text{and} \quad r^{(0)}_1, \ldots, r^{(0)}_{n-b}, r^{(1)}_1, \ldots, r^{(1)}_{n-b-1}, \ldots, r^{(n-b-1)}_1$$

are, let us consider only those terms

$$\frac{2\pi\epsilon}{2^b}\left(u^{(k)}_{n-b} v_0 r^{(0)}_0 + u^{(k)}_{n-b-1} v_1 r^{(1)}_0 + \ldots + u^{(k)}_0 v_{n-b} r^{(n-b)}_0\right) = \frac{2\pi\epsilon}{2^b} \sum_{i=b}^{n} u^{(k)}_{i-b} v_{n-i} r^{(n-i)}_0. \tag{7}$$

We will further throw away some noise terms in (7). Let $d = \text{ord}_2(\omega)$. Recall that $d < \frac{\log_2 \omega}{2}$ and $i_0 = \lfloor \frac{3}{4} \log_2 \omega \rfloor$. Thus, assuming b is $O(\log n)$, $i_0 - b - d = \Omega(n)$, and we will only consider the subsum in (7) for $i \in \{d + b, \ldots, i_0\}$, which has $\Omega(n)$ terms.

By Lemma 7, except for an exponentially small fraction $2^{-\Omega(n)}$ of j indexing $v = \lfloor \frac{2^n}{\omega} j \rfloor$ $(1 \leq j < \omega)$, each j defines a linear sized $T_j = \{d + b \leq i \leq i_0 : v_{n-i} = 1\}$ (of cardinality $> \Omega(n)$) such that $v_{n-i} = 1$ and so $u_{i-b} v_{n-i} r^{(n-i)}_0 = u_{i-b} r^{(n-i)}_0$, for $i \in T_j$. Thus we will further ignore a large portion of the above sum (7), and consider only

$$\frac{2\pi\epsilon}{2^b} \sum_{i \in T_j} u^{(k)}_{i-b} r^{(n-i)}_0. \tag{8}$$

Intuitively, any term that in fact survives (i.e., with $u^{(k)}_{i-b} v_{n-i} = 1$) that we omitted can only increase the noise. (Formally, when we eventually apply Lemma 6, these will all be part of the term φ_k.)

Our next goal is to show that, among $i \in T_j$, most pairs of $u^{(k)} = u^* + k\omega$ and $u^{(k')} = u^* + k'\omega$, for $k \neq k'$, have a linear number of different bit values $u^{(k)}_{i-b} \neq u^{(k')}_{i-b}$, for $i \in T_j$.

To investigate the (least $i_0 - b + 1$ significant) bits $u^{(k)}_0, u^{(k)}_1, \ldots, u^{(k)}_{i_0-b}$ of $u^{(k)} = u^* + k\omega$, we consider $u^{(k)} \bmod 2^{i_0-b+1}$. If ω is odd, then $(k\omega \bmod 2^{i_0-b+1})$ will enumerate all values in $\{0, 1, \ldots, 2^{i_0-b+1} - 1\}$ exactly once, when $k = 0, 1, \ldots, 2^{i_0-b+1} - 1$. Our range of k is actually from 0 to just below $\frac{2^n - u^*}{\omega} \approx 2^n/\omega \gg 2^{i_0}$. Thus, for any u^*, $(u^* + k\omega \bmod 2^{i_0-b+1})$ enumerates every value in $\{0, 1, \ldots, 2^{i_0-b+1} - 1\}$ almost uniformly.

In general, $0 \leq d = \text{ord}_2(\omega) < \frac{\log_2 \omega}{2}$, and thus $i_0 - b - d = \Omega(n)$ for $b = O(\log n)$, and then $\omega' = \omega/2^d$ is invertible in $\mathbb{Z}_{2^{i_0-b+1-d}}$, and for any u^*, the most significant $i_0 - b + 1 - d = \Omega(n)$ bits in $(u^{(k)} \bmod 2^{i_0-b+1})$ are almost uniform. These are the most significant $i_0 - b + 1 - d$ of the $i_0 - b + 1$ least significant bits of $u^{(k)} = u^* + k\omega$.

Consider $u^{(k)} = u^* + k\omega$ and $u^{(k')} = u^* + k'\omega = u^{(k)} + (k' - k)\omega$. For any k, let k' run through $\{0, \ldots, \lfloor \frac{2^n - u^*}{\omega} \rfloor\}$, then $k' - k$ runs through $\{-k, \ldots, \lfloor \frac{2^n - u^*}{\omega} \rfloor - k\}$, a set of consecutive integers of size $\geq \lfloor \frac{2^n}{\omega} \rfloor \geq 2^{i_0-b+1-d} \left(\frac{2^n}{2^{i_0}\omega}\right)$. As $2^{i_0}\omega \leq 2^{7n/8}$ we have $\frac{2^n}{2^{i_0}\omega} \geq 2^{n/8}$. Hence, $(k' - k)\omega' \bmod 2^{i_0-b+1-d}$

picks every value in $\{0, \ldots, 2^{i_0-b+1-d}-1\}$ with probability $\frac{1}{2^{i_0-b+1-d}} \cdot (1 \pm 2^{-\Omega(n)})$. Thus, the most significant $i_0 - b + 1 - d = \Omega(n)$ of the $i_0 - b + 1$ least significant bits of $u^{(k')} - u^{(k)} = (k' - k)\omega$ are almost uniform, as k' runs through $\{0, \ldots, \lfloor \frac{2^n - u^*}{\omega} \rfloor\}$. Then using the same argument with the entropy function $H(\cdot)$, for all except a fraction of $2^{-\Omega(n)}$ of the pairs (k, k'), we have $u^{(k)}_{i-b} \neq u^{(k')}_{i-b}$, for a subset of $i \in T_j$ of cardinality $\geq c_0 n$, where the constant $c_0 > 0$ is uniform for (k, k').

Lemma 8. *Assume $|T_j| = \Omega(n)$. There exists $c_0 > 0$, such that for random pairs (k, k'),*

$$\Pr.\left(\left|\{i \in T_j : u^{(k)}_{i-b} \neq u^{(k')}_{i-b}\}\right| \geq c_0 n\right) = 1 - 2^{-\Omega(n)}.$$

It follows that, except for a $2^{-\Omega(n)}$ fraction of pairs (k, k'), the sum

$$\sum_{i=b}^{n} (u^{(k)}_{i-b} - u^{(k')}_{i-b}) v_{n-i} r_0^{(n-i)}$$

contains a linear number $n' \geq c_0 n$ of uncancelled terms $r_0^{(n-i)}$ where $v_{n-i} = 1$ and $u^{(k)}_{i-b} \neq u^{(k')}_{i-b}$. To apply Lemma 6, we require $(\frac{\epsilon}{2^b})^{-1} < (n')^{1/3}$, or equivalently $b + \log 1/\epsilon < \frac{1}{3} \log n'$.

To summarize the error estimates: (I) except with probability $2^{-\Omega(n)}$, we have $\omega > N^{1/3}$ and $\mathrm{ord}_2(\omega) < \frac{\log_2 \omega}{2}$ by Lemma 5; (II) except for a fraction of $2^{-\Omega(n)}$ of j's, all $v = \lfloor \frac{2^n}{\omega} j \rfloor$ have $|T_j| = \Omega(n)$ by Lemma 7; (III) except for a fraction of $2^{-\Omega(n)}$ of all pairs (k, k')'s, the sums (8) defined by k and k' all have a symmetric difference with cardinality $\geq (2^b/\epsilon)^3$, by Lemma 8.

Finally we estimate the sum of the expectations of the square norm sum (4) indexed by all $v = \lfloor \frac{2^n}{\omega} j \rfloor$. Note that the sum $\sum_{k=0}^{K-1}$ is over K complex numbers of unit norm, and thus has norm at most K. With probability $\leq 2^{-\Omega(n)}$, (I) may be violated and the sum over all $v = \lfloor \frac{2^n}{\omega} j \rfloor$ of (4) can be at most $\frac{\omega}{2^n K} K^2 = O(1)$. Assume (I) holds, then the sum of the terms (4) indexed by the $\leq 2^{-\Omega(n)}$ fraction of exceptional v's regarding (II) has value at most $(2^{-\Omega(n)}\omega) \frac{1}{2^n K} K^2 = 2^{-\Omega(n)}$. Assume (I) and (II) are both not violated, we apply Lemma 6. By (III), we get the estimate over all v, the expectation $\frac{\omega}{2^n K} \left(K + 2^{-\Omega(n)} K^2 + K^2 2^{-\Omega(n^{1/3})} \right) = 2^{-\Omega(n^{1/3})}$.

We conclude that the expectation (over the random noise bits r's) of the probability of observing a member in $\{|v\rangle : v = \lfloor \frac{2^n}{\omega} j \rfloor, 0 \leq j < \omega\}$ is exponentially small.

The proof carries over easily to those $|v\rangle$ that are in the vicinity of a polynomial range of $\lfloor \frac{2^n}{\omega} j \rfloor$. And since the estimate is exponentially small, the proof shows that the probability of observing any member of the set of those $|v\rangle$ that are polynomially close to an integral multiple $\lfloor \frac{2^n}{\omega} j \rfloor$ is still exponentially small in expectation.

4 Some comments

This section contains some comments and personal opinions. They are speculative, and are not to be conflated with the provable part.

Quantum mechanics is unquestionably an accurate model of microscopic physical reality. However, I believe every physical theory is an approximate description of the real world, and quantum mechanics is no exception. In particular, I believe the SU(2) description of possible operations of a qubit to be only approximately true. Specifically, I don't believe arbitrarily small angles have physical meaning.

The real numbers $\mathbb{R}$, the continuum, is a human logical construct in terms of Dedekind cut or Cauchy sequence in the language of ϵ-δ definition. SU(2) (or equivalently SO(3)) as a group, is built on top of the continuum. That these mathematical objects provide remarkable fit in some mathematical theory for physical reality, is an extraordinary fact. But this extraordinary fit is always within a certain range; its unlimited extrapolation is mathematical idealization. The Schrödinger equation $i\hbar\frac{d}{dt}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle$ suggests that small angles are related to small time periods. But physicists have suggested that time ultimately also comes in discrete and indivisible "units". The concept "chronon" has been proposed as a quantum of time [25]. It has even been proposed that one chronon corresponds to about 6.27×10^{-24} seconds for an electron, much longer than the Planck time, which is only about 5.39×10^{-44} [9] (see also [12, 4]). (Of course the literal form of the mathematical meaning of Schrödinger equation, as a differential equation, suggests time is infinitely divisible. But my personal view is that this is just mathematical abstraction.)

Thus, I view arbitrarily small angles permitted under SU(2) as mere mathematical abstraction. It is true that using a fixed finite set of rotations of reasonable angles such as $\pi/8$ along various axes *can* compose to rotations of arbitrarily small angles. But my view is that these compositional rules as specified by the group SU(2) must not be exact for physical reality. And thus, it seems to me permitting some noise in the model is reasonable. The random noise model in this paper is just a model, is not meant as reality.

Of course, in addition to its intrinsic interest, factoring integers of the form pq is at the heart of the RSA public-key cryptosystem [34]. But several results and conjectures in number theory suggest that the failure reported in this paper of Shor's factoring algorithm in the presence of noise can be more severe in the asymptotic sense. We used a theorem of Fouvry [13] to produce a set of primes of positive density that have the desired properties of the period of a random element. The most important property is that this period is sufficiently large. In Theorem 3 we prove a version of the theorem for primes of density one. There are deep results and many conjectures about the distribution of prime factorizations of $p - 1$. In the extreme there are the so-called Sophie Germain primes p' such that $p = 2p' + 1$ is also a prime. It is conjectured that there are $2C\frac{x}{(\log_e x)^2}$ many Sophie Germain primes up to x, where $C = \prod_{p>2} \frac{p(p-2)}{(p-1)^2} \approx 0.660161$ is the Hardy-Littlewood twin prime constant. This is just slightly less than positive density. (However, it has not been proved that there are infinitely many Sophie Germain primes.) Sophie Germain primes were studied in (the first case of) Fermat's Last Theorem. Indeed, Adleman and Heath-Brown [2], and Fouvry [13] proved that the first case of Fermat's Last Theorem holds for infinitely many primes p. (See also [22].) In [20], Håstad, Schrift and Shamir (acknowledging Noga Alon) proved that, for any fixed constant k, for randomly chosen primes p and q of equal size, and $N = pq$ of size n, a random element in $\mathbb{Z}_N^*$ has order $\geq \phi(pq)/n^k$ except with probability $O(n^{-(k-5)/5})$. We improve this slightly in Theorem 13 in the process of proving Theorem 3.

Another property we use of primes of the property of Theorem 4 is that the period of a random element in $\mathbb{Z}_N^*$ does not have high ord_2. Erdös and Odlyzko [11] proved that the set of odd divisors of $p - 1$ has a positive density.

Finally, a few comments on the Strong Church-Turing thesis. It is conceivable that some other quantum algorithm in the BQP model can factor integers (or some other seemingly difficult problem) in polynomial time, *and* withstand the random noise discussed in this paper. Separately, it is definitely conceivable that at some future time, a quantum algorithm is superior to the best "classical" factoring algorithms for integers of a certain range. But I am not convinced that BQP requires that we revise the Strong Church-Turing thesis, even if factoring is eventually known to be outside P or BPP. In Turing's careful definition of computability, he made a deliberate choice that the "primitive" steps of such a computing device must be discrete. Thus, the set of states of a TM is finite; the symbols are placed in discrete cells; the alphabet set is finite. At its most fundamental level, it is not permitted to ask the computing machine to scan a continuously deformed symbol from ξ to ζ, while a mathematical homotopy can easily be envisioned. I believe the model BQP, in its use of the full SU(2) as primitive steps (or what amounts to equivalently, the assumption that the exact rule of composition of SU(2) corresponds exactly to realizable computational steps), is a departure from the Turing model.

Acknowledgement

The author thanks Al Aho, Dan Boneh, Péter Gács, Zvi Galil, Fred Green, Steve Homer, Leonid Levin, Dick Lipton, Ashwin Maran, Albert Meyer, Ken Regan, Ron Rivest, Peter Shor, Mike Sipser, Les Valiant, and Ben Young for insightful comments. He also thanks Eric Bach for inspiring discussions on some of the number theoretic estimates, and we hope to report some further improvements soon [7]. A similar result can be proved for Shor's algorithm computing Discrete Logarithm, and will be reported later.

References

[1] Scott Aaronson. Quantum Computing Since Democritus. Cambridge University Press. 2013

[2] L. M. Adleman and D. R. Heath-Brown The first case of Fermat's last theorem Inventiones mathematicae (1985). Volume: 79, pp. 409-416.

[3] D. Aharonov and M. Ben-Or. Fault-tolerant quantum computation with constant error. In Proceedings of STOC'97, pages 176–188, 1997.

[4] Claudio Albanese and Stephan Lawi. Time Quantization and q-deformations. Journal of Physics A. 37 (8): 2983–2987.

[5] P. Aliferis, D. Gottesman and J. Preskill. Quantum accuracy threshold for concatenated distance-3 codes. Quantum Information and Computation, 6(2):97–165, 2006.

[6] P. Aliferis. Level reduction and the quantum threshold theorem. Ph.D. thesis, Caltech, 2007.

[7] Eric Bach. Private communications.

[8] Calderbank, A. R.; Rains, E. M.; Shor, P. W.; Sloane, N. J. A. (1998). "Quantum Error Correction via Codes over GF(4)". IEEE Transactions on Information Theory. 44 (4): 1369–1387.

[9] Piero Caldirola. The introduction of the chronon in the electron theory and a charged lepton mass formula. Lettere al Nuovo Cimento. 27 (8): 225–228.

[10] D. Coppersmith. (1994), An Approximate Fourier Transform Useful in Quantum Factoring, IBM Research Report RC 19642. https://arxiv.org/abs/quant-ph/0201067

[11] P. Erdös and A. Odlyzko. On the density of odd integers of the form $(p-1)/2^{-n}$ and related questions. Journal of number theory, vol. 11 (1979) pp 257-263.

[12] Ruy A. H. Farias and Erasmo Recami. Introduction of a Quantum of Time ("chronon"), and its Consequences for Quantum Mechanics. arXiv:quant-ph/9706059

[13] Étienne Fouvry. Théorème de Brun-Titchmarsh; application au théorème der Fermat. Inventiones mathematicae (1985). Volume: 79, pp. 383-408

[14] A. G. Fowler and L. C. L. Hollenberg. Scalability of Shor's algorithm with a limited set of rotation gates. Phys. Rev. A 70, 032329 (2004).

[15] A. G. Fowler and L. C. L. Hollenberg. Erratum: Scalability of Shor's algorithm with a limited set of rotation gates. [Phys. Rev. A 70, 032329 (2004)] Phys. Rev. A 75, 029905 (2007)

[16] Gil Kalai. https://gilkalai.wordpress.com/2022/05/26/waging-war-on-quantum/

[17] A. Kitaev. Quantum computations: Algorithms and error correction. Russian Mathematical Surveys, 52(6):1191–1249, 1997.

[18] E. Knill. Quantum computing with realistically noisy devices. Nature, 434:39–44, 2005.

[19] Gottesman, Daniel (2009). "An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation". arXiv:0904.2557

[20] Johan Håstad, A. W. Schrift and Adi Shamir. The Discrete Logarithm Modulo a Composite Hides $O(n)$ Bits. J. Comput. Syst. Sci. 47(3): 376-404 (1993)

[21] Christopher Hooley. Applications of sieve methods to the theory of numbers, Cambridge University Press. 1976.

[22] H. W. Lenstra and P. Stevenhagen Jr. (1997). Class Field Theory and the First Case of Fermat's Last Theorem. In: Cornell, G., Silverman, J.H., Stevens, G. (eds) Modular Forms and Fermat's Last Theorem. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-1974-3_18

[23] Leonid Levin. https://www.cs.bu.edu/fac/lnd/expo/qc.htm

[24] R. Landauer, in Quantum Computing and Communications, edited by M. Brooks (Springer, London, 1999), p. 61

[25] Henry Margenau. The Nature of Physical Reality. McGraw-Hill. (1950).

[26] H. L. Montgomery and R. C. Vaughan. The large sieve. Mathematika 20: 119-134. 1973.

[27] H. L. Montgomery and R. C. Vaughan. Multiplicative Number Theory I: Classical Theory. Cambridge Studies in Advanced Mathematics, Series Number 97. Cambridge University Press 2006.

[28] A. J. Menezes, P. C. van Oorschot and S. A. Vanstone. Handbook of applied cryptography. Boca Raton: CRC Press. 1997.

[29] Y. S. Nam and R. Blümel. Scaling laws for Shor's algorithm with a banded quantum Fourier transform. PHYSICAL REVIEW A 87, 032333 (2013)

[30] Y. S. Nam and R. Blümel. Robustness of the quantum Fourier transform with respect to static gate defects. April 2014. Physical review A, Atomic, molecular, and optical physics 89(4).

[31] Y. S. Nam and R. Blümel. Performance scaling of the quantum Fourier transform with defective rotation gates. July 2015. Quantum Information & Computation 15(9-10):721-736

[32] Michael Nielsen and Isaac Chuang. Quantum Computation and Quantum Information. Cambridge University Press. 2000 (first ed.) 2010 (second ed.)

[33] Athanasios Papoulis. Probability, Random Variables and Stochastic Processes (4th ed.). p. 148.

[34] R. Rivest, A. Shamir, and L. Adleman, A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. Communications of the ACM. 21 (2): 120–126. (February 1978).

[35] J. B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, Illinois J. Math. 6 (1962), 64–94.

[36] Peter Shor. Algorithms for quantum computation: discrete logarithms and factoring. Proceedings 35th Annual Symposium on Foundations of Computer Science. IEEE Comput. Soc. Press: 124–134. (1994).

[37] Peter Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM Journal on Computing. Vol. 26, Iss. 5 (1997)10.1137/S0097539795293172

[38] Peter Shor. Scheme for reducing decoherence in quantum computer memory. Physical Review A. 52 (4): R2493–R2496. (1995)

[39] Peter Shor. Fault-tolerant quantum computation. In Proceedings of FOCS '96, pages 56–65, 1996.

[40] A. Steane. Multiple particle interference and quantum error correction. Proceedings of the Royal Society of London Ser. A, 452: 2551–2573, 1996.

[41] A. Steane and D. Lucas. Quantum computing with trapped ions, atoms and light. Fortschritte der Physik, 48:839–858, 2000.

Appendix: Pairs of random primes

The proof in the paper exhibits a particular set of primes of positive density, and shows that if the input N to Shor's algorithm is of the form $N = pq$ for any primes p and q from that set then the algorithm does not factor with exponentially small exceptional probability, *if* the rotational gates are accompanied by a suitable level of noise.

In cryptography, an interesting question concerns the performance on $N = pq$ for random primes p and q of length m.

In this appendix, we prove Theorem 3, dealing with random pairs of primes p and q chosen uniformly from all primes of the same length.

To prove Theorem 3, we will appeal to some number theoretic estimates for

- The period $\omega_N(x)$ of a random element $x \in \mathbb{Z}_N^*$, where $N = pq$, and p and q are primes uniformly randomly chosen from all primes of length m. (The period $\omega_N(x)$ is the order of x as a group element in $\mathbb{Z}_N^*$.)
- The exact order of the prime 2 of the integer $\omega_N(x)$, i.e., $\mathrm{ord}_2(\omega_N(x))$, for a random element $x \in \mathbb{Z}_N^*$, where $N = pq$, and p and q are primes uniformly randomly chosen from all primes of length m.

For primes p and q of binary length m, $N = pq$ has binary length $\approx 2m$, and the QFT circuit uses about $4m$ qubits with $2^{4m} \approx N^2$. The statement $b + \log 1/\epsilon < \frac{1}{3}\log m - c$ for some $c > 0$ is equivalent to $b + \log 1/\epsilon < \frac{1}{3}\log(4m) - c'$ for some $c' > 0$. We note that to carry through the same proof of the Main theorem in the paper, we only need to have the property that

1. $\omega_N(x)$ is large, say $\omega_N(x) = 2^{\Omega(m)}$, and
2. $\mathrm{ord}_2(\omega_N(x))$ is not too large, say $\mathrm{ord}_2(\omega_N(x)) = o(m)$.

Håstad, Schrift and Shamir proved a version of the following theorem (Theorem 9) (acknowledging Noga Alon) [20] [Proposition 1, p. 378]. Their theorem is sufficient to address item (1) for our purpose. But we will give a minor improvement using the Brun-Titchmarsh theorem, which will be used to derive a bound for item (2) as well. The proof will be essentially the same as in [20]; the minor improvement comes from using the Brun-Titchmarsh theorem and an estimate due to Rosser and Schoenfeld [35] [Theorem 15]:

$$\frac{d}{\phi(d)} \le e^{\gamma} \cdot \log\log d \cdot \left(1 + \frac{2.5}{e^{\gamma}(\log\log d)^2}\right),$$

where $\phi(\cdot)$ is the Euler totient function, $\gamma = 0.577\ldots$ is Euler's constant, and log denotes natural logarithm (as it will be for the rest of this section). The estimate is valid for every $d \ge 3$, except one case $d = 2 \cdot 3 \cdot \ldots \cdot 23$ when the constant 2.5 should be replaced by 2.50637. We will just use $\frac{d}{\phi(d)} \leq C \log\log d$ for some universal constant C, and all $d \geq 3$.

Denote by $X = 2^m - 1$, $Y = \lceil \frac{X}{2} \rceil = 2^{m-1}$.

Theorem 9. *There exists a constant C, such that for any m and any randomly chosen distinct primes p and q of binary length m, $N = p \cdot q$, and let g be a randomly chosen element in $\mathbb{Z}_N^*$, then for all $m^2 \leq A < X$,*

$$\Pr.\left(\omega_N(g) < \frac{1}{A}\phi(N)\right) \leq C\frac{m^{2/5}}{A^{1/5}},$$

where the probability is over all random $Y \leq p \neq q \leq X$ and $g \in \mathbb{Z}_N^*$.

Note that $\phi(N) = (p-1)(q-1) \approx 2^{2m}$. If we take $A = 2^{2\epsilon m}$ then a random $\omega_N(g) \geq 2^{2(1-\epsilon)m}$ with probability $1 - O(m2^{-\epsilon m/5})$. This is more than sufficient for our required item (1) above.

The Brun-Titchmarsh theorem is a reasonably sharp estimate for the number of primes up to any upper bound x, in an arithmetic progression. The bound is applicable even when the modulus of the arithmetic progression is large. The following version is an improvement of the original Brun-Titchmarsh theorem proved by Montgomery and Vaughan [26, 21]. Suppose a and d are relatively prime. Let $\pi(x; d, a)$ denote the number of primes $p \equiv a \bmod d$, with $p \leq x$.

Theorem 10 (Montgomery-Vaughan)**.**

$$\pi(x; d, a) \leq \frac{2x}{\phi(d)\log(x/d)},$$

for all $d < x$, where $\phi(\cdot)$ is the Euler totient function, and log *denotes natural logarithm.*

Following [20] the proof of Theorem 9 is based on two lemmas. Let $O_N = \max\{\omega_N(x) : x \in \mathbb{Z}_N^*\}$ be the exponent of the finite Abelian group $\mathbb{Z}_N^* \cong \mathbb{Z}_p^* \times \mathbb{Z}_q^*$, then $O_N = \operatorname{lcm}(p-1, q-1)$, and $\omega_N(x) | O_N$ for all $x \in \mathbb{Z}_N^*$.

Lemma 11. *There exists a constant $C_1 > 0$, such that for randomly chosen distinct primes p and q of binary length m, $N = p \cdot q$, and for any $1 \leq A_1 \leq X^{1/4} < 2^{m/4}$,*

$$\Pr.\left(O_N < \frac{1}{A_1}\phi(N)\right) \leq C_1\frac{1}{A_1}.$$

Proof. It is trivial if $m \leq 2$. We will assume $m > 2$. Clearly $O_N = \phi(N)/\gcd(p-1, q-1)$. So,

$$O_N < \frac{1}{A_1}\phi(N) \Longleftrightarrow \gcd(p-1, q-1) > A_1.$$

By the Prime Number Theorem, the number of primes of length m is $\pi(X) - \pi(Y) \approx \frac{x}{2\log x}$. And so the number of ordered pairs of distinct primes of length m is approximately $(\frac{x}{2\log x})^2$. Now we bound the cardinality of

$$S = \{(p, q) : Y \leq p \neq q \leq X, p, q \text{ are primes, and } \gcd(p-1, q-1) > A_1\}.$$

For $p \neq q$ in that range, we claim that $p-1 \nmid q-1$. For otherwise $(q-1)/2 \geq p-1$, which implies that $q \geq 1+2(Y-1) = X$ and hence $q = X$. Then $p-1 \leq (q-1)/2 = Y-1 \leq p-1$, and so equality holds, and $p = Y = 2^{m-1}$, a contradiction. It follows that $\gcd(p-1, q-1) \leq (p-1)/2 < X/2$. So, $\gcd(p-1, q-1) \leq 2^{m-1} - 1$.

We have

$$\begin{aligned} |S| &= \sum_{d=\lfloor A_1 \rfloor+1}^{2^{m-1}-1} \sum_{(p,q)} \mathbf{1}_{[\gcd(p-1,q-1)=d]} \\ &\leq \sum_{d=\lfloor A_1 \rfloor+1}^{2^{m-1}-1} (\pi(X; d, 1) - \pi(X/2; d, 1))^2, \end{aligned}$$

where $\sum_{(p,q)}$ denotes the sum over primes (p,q) in the range $Y \leq p \neq q \leq X$.

Now we separate the sum into two parts, depending on whether $d > \lfloor A_1^2 X^{1/3} \rfloor$. One part is

$$H = \sum_{d=\lfloor A_1^2 X^{1/3} \rfloor+1}^{2^{m-1}-1} (\pi(X; d, 1) - \pi(X/2; d, 1))^2,$$

where we use the trivial bound $\pi(X; d, 1) - \pi(X/2; d, 1) \leq \frac{X}{2d} + 1$. In the range $d < 2^{m-1}$, it is $\leq \frac{X}{d}$. It follows that

$$H < X^2 \sum_{d=\lfloor A_1^2 X^{1/3} \rfloor+1}^{\infty} \frac{1}{d^2} < \frac{X^2}{A_1^2 X^{1/3}} = \frac{X^{5/3}}{A_1^2},$$

by a comparison to the integral $\int_K^\infty \frac{1}{x^2} dx = \frac{1}{K}$.

The other part is

$$L = \sum_{d=\lfloor A_1 \rfloor+1}^{\lfloor A_1^2 X^{1/3} \rfloor} (\pi(X; d, 1) - \pi(X/2; d, 1))^2,$$

where we use Theorem 10, to get

$$L \leq \sum_{d=\lfloor A_1 \rfloor+1}^{\lfloor A_1^2 X^{1/3} \rfloor} \left(\frac{2X}{\phi(d) \log \frac{X}{d}} \right)^2.$$

As $d \leq A_1^2 X^{1/3} \leq X^{5/6}$, we have $\frac{X}{d} \geq X^{1/6}$, and $\log \frac{X}{d} \geq (\log X)/6$. So

$$L \leq 144 \left(\frac{X}{\log X} \right)^2 \sum_{d=\lfloor A_1 \rfloor+1}^{\lfloor A_1^2 X^{1/3} \rfloor} \frac{1}{\phi(d)^2}.$$

Next we claim that

Claim: $\sum_{d>D} \frac{1}{\phi(d)^2} = O(\frac{1}{D})$, for any $D \geq 1$.

To prove this Claim we need a result from [27] [p. 61, equation (2.32)]

$$\sum_{n \leq x} \left(\frac{n}{\phi(n)} \right)^2 = O(x),$$

for all $x > 0$. Let $a_n = \frac{1}{n^2}$, $b_n = \left(\frac{n}{\phi(n)}\right)^2$, and $B_n = \sum_{k=D+1}^{n} b_k$, with $n \geq D$. Then $B_D = 0$ and $b_n = B_n - B_{n-1}$, for all $n > D$. We have for all $Z > D$,

$$\sum_{n=D+1}^{Z} \frac{1}{\phi(n)^2} = \sum_{n=D+1}^{Z} a_n b_n = a_Z B_Z + \sum_{n=D+1}^{Z-1} (a_n - a_{n+1}) B_n.$$

Now $a_Z B_Z = O(1/Z)$, $a_n - a_{n+1} < 2/n^3$ and thus $(a_n - a_{n+1})B_n = O(1/n^2)$. It follows that

$$\sum_{n=D+1}^{Z} \frac{1}{\phi(n)^2} = O(1/Z) + O(1/D).$$

Letting $Z \to \infty$ proves the Claim.

It follows that

$$L = O\left(\left(\frac{X}{\log X}\right)^2 \cdot \frac{1}{A_1}\right).$$

And

$$|S| \leq L + H = O\left(\left(\frac{X}{\log X}\right)^2 \cdot \frac{1}{A_1}\right) + \frac{X^{5/3}}{A_1^2},$$

Hence,

$$\text{Pr.}\left(O_N < \frac{1}{A_1}\phi(N)\right) = O\left(\frac{1}{A_1}\right).$$

The lemma is proved. □

Lemma 12. *There exists a constant $C_2 > 0$, such that for any $B > 1$,*

$$\text{Pr.}\left(\omega_p(g) < \frac{1}{B}\phi(p)\right) \leq C_2 \left(\frac{m}{B \log B}\right)^{1/2},$$

where the probability is over a random prime $Y \leq p \leq X$ and a random $g \in \mathbb{Z}_p^*$, and $\omega_p(g)$ is the order of g as a group element in $\mathbb{Z}_p^*$.

Proof. For any prime p, the order of any $g \in \mathbb{Z}_p^*$ divides the order of the group $\phi(p) = p - 1$,

$$|\{g \in \mathbb{Z}_p^* : \omega_p(g) < \frac{1}{B}\phi(p)\}| = \sum_{d|p-1,\ d<\phi(p)/B} \phi(d).$$

Define $F(p) = \sum_{d|p-1,\ d<\phi(p)/B} \phi(d)$ for any prime p, we have

$$\sum_{Y \leq p \leq X} F(p) = \sum_{d<X/B} \phi(d) \sum_{Y \leq p \leq X} \mathbf{1}_{[d|p-1]} \leq \sum_{d<X/B} \phi(d)\pi(X; d, 1).$$

Now we apply Theorem 10 and obtain

$$\sum_{Y \leq p \leq X} F(p) \leq \sum_{d<X/B} \frac{2X}{\log(X/d)} \leq \frac{2X^2}{B \log B}.$$

It follows that for any $B' > 0$,

$$|\{p : Y \leq p \leq X,\ p \text{ is a prime, and } F(p) \geq X/B'\}| \leq \frac{2X^2}{B \log B} \cdot \frac{B'}{X} = \frac{2XB'}{B \log B}.$$

Then, by the Prime Number Theorem,

$$\text{Pr.} \left(F(p) \geq X/B' \right) \leq O\left(\frac{B' \log X}{B \log B} \right).$$

Conditional on any p such that $Y \leq p \leq X$ and $F(p) < X/B'$, the probability over $g \in \mathbb{Z}_p^*$ of the event $\omega_p(g) < \frac{1}{B}\phi(p)$, is $\frac{F(p)}{p-1} < \frac{3}{B'}$. Thus, the conditional probability over both p and $g \in \mathbb{Z}_p^*$ given $F(p) < X/B'$ is

$$\text{Pr.} \left[\omega_p(g) < \frac{1}{B}\phi(p) \middle| F(p) < \frac{X}{B'} \right] = O\left(\frac{1}{B'} \right).$$

It follows easily that

$$\text{Pr.} \left(\omega_p(g) < \frac{1}{B}\phi(p) \right) = O\left(\frac{B' \log X}{B \log B} \right) + O\left(\frac{1}{B'} \right).$$

Setting $B' = (B \log B / \log X)^{1/2}$, gives the bound of the lemma. □

Now the proof of Theorem 9 can be completed.

Proof. [of Theorem 9] We will pick A_1 and A_2 such that $A = A_1 A_2$, then

$$\text{Pr.} \left(\omega_N(g) < \frac{1}{A}\phi(N) \right) = \text{Pr.} \left(O_N < \frac{1}{A_1}\phi(N) \right) + \text{Pr.} \left(\omega_N(g) < \frac{1}{A_2} O_N \right),$$

where the first expression is over primes $Y \leq p \neq q \leq X$ and the second expression is over p, q and $g \in \mathbb{Z}_N^*$. This is seen by the contrapositive: if $\phi(N) \leq A_1 O_N$ and $O_N \leq A_2 \omega_N(g)$ then $\phi(N) \leq A \omega_N(g)$.

By Lemma 11, the first term is $O\left(\frac{1}{A_1}\right)$.

For the second term, we know that $\omega_N(g) = \text{lcm}(\omega_p(g), \omega_q(g))$, as $\mathbb{Z}_N^* \cong \mathbb{Z}_p^* \times \mathbb{Z}_q^*$. $\omega_p(g)$ is a divisor of $p-1$, and similarly for $\omega_q(g)$. We write $\omega_p(g) = (p-1)/a$, and $\omega_q(g) = (q-1)/b$, then

$$\omega_N(g) \geq \frac{\text{lcm}(p-1, q-1)}{ab}.$$

To see this, we take any prime $r \mid \text{lcm}(p-1, q-1)$,

$$\begin{aligned} \text{ord}_r(\omega_N(g)) &= \max\{\text{ord}_r(p-1) - \text{ord}_r(a), \text{ord}_r(q-1) - \text{ord}_r(b)\} \\ &\geq \max\{\text{ord}_r(p-1), \text{ord}_r(q-1)\} - \max\{\text{ord}_r(a), \text{ord}_r(b)\} \\ &\geq \max\{\text{ord}_r(p-1), \text{ord}_r(q-1)\} - \text{ord}_r(ab) \\ &= \text{ord}_r(\text{lcm}(p-1, q-1)) - \text{ord}_r(ab). \end{aligned}$$

It follows that, after taking $B = \sqrt{A_2}$,

$$\text{Pr.} \left(\omega_N(g) < \frac{1}{A_2} O_N \right) \leq \text{Pr.} \left(\omega_p(g) < \frac{p-1}{B} \right) + \text{Pr.} \left(\omega_q(g) < \frac{q-1}{B} \right) = O\left(\frac{m}{B \log B} \right)^{1/2},$$

by Lemma 12. Equalize the two error bounds we set

$$\frac{1}{A_1} \approx \left(\frac{m}{B\log B}\right)^{1/2},$$

subject to $1 \leq A_1 \leq X^{1/4}$, $A_1B^2 = A$, $B > 1$, where A is given as $m^2 \leq A < X$.

We can set $B = \frac{(A^2m)^{1/5}}{\log A}$ to achieve the bound in Theorem 9.

□

We remark that, for polynomial bounded $A = m^k$, we can choose B slightly better, $B = \left(\frac{m^{2k+1}}{\log m}\right)^{1/5}$, and achieve

Theorem 13. *With the same setting as in Theorem 9, for any $k \geq 2$*

$$\Pr.\left(\omega_N(g) < \frac{1}{m^k}\phi(N)\right) \leq O\left(\frac{1}{m^{(k-2)/5}(\log m)^{2/5}}\right),$$

where the probability is over all random $Y \leq p \neq q \leq X$ and $g \in \mathbb{Z}_N^*$. (The constant in O depends on k.)

Finally, to finish the proof of Theorem 3, we address the required item (2), again using the Brun-Titchmarsh theorem.

For any prime p, we have the prime factorization $p - 1 = 2^{e_0}p_1^{e_1}\cdots p_k^{e_k}$. We have

$$\Pr.\left(\exists g \in \mathbb{Z}_p^* : \mathrm{ord}_2(\omega_p(g)) \geq e\right) \leq \frac{1}{\pi(X)-\pi(Y)}\frac{2X}{\phi(2^e)\log(X/2^e)},$$

where the probability is over a random $Y \leq p \leq X$.

We have $\phi(2^e) = 2^{e-1}$ for $e \geq 1$, and $\pi(X) - \pi(Y) = \Theta(X/\log X)$. Using the Rosser-Schoenfeld estimate again, we have

$$\Pr.\left(\exists g \in \mathbb{Z}_p^* : \mathrm{ord}_2(\omega_p(g)) \geq e\right) \leq O\left(\frac{\log X}{\log(X/2^e)}\frac{\log\log 2^e}{2^e}\right).$$

If we set $m^c = 2^e$, then we get an upper bound of $O\left(\frac{\log\log m}{m^c}\right)$, where the constant in O depends on c. Thus, for any $c > 0$,

$$\Pr.\left(\exists g \in \mathbb{Z}_p^* : \mathrm{ord}_2(\omega_p(g)) \geq c\log_2 m\right) \leq O\left(\frac{\log\log m}{m^c}\right).$$

As $\omega_N(g) = \mathrm{lcm}(\omega_p(g), \omega_q(g))$, it follows that,

$$\Pr.\left(\exists g \in \mathbb{Z}_N^* : \mathrm{ord}_2(\omega_N(g)) \geq c\log_2 m\right) \leq O\left(\frac{\log\log m}{m^c}\right).$$

Since both required items (1) and (2) are separately true with probability approaching 1, they are jointly true with probability approaching 1.