Shor’s algorithm

Shor’s algorithm

- **Shor’s algorithm is a polynomial-time quantum computer algorithm for integer factorization, which is the order of $O((logN)^2(loglogN)(logloglogN))$, where N is an integer.**
- **Classically, it would take $O(e^{1.9(logN)^{\frac{1}{3}}(loglogN)^{\frac{2}{3}}})$, and the algorithm is general number field sieve**

Procedure

- **Here, we would restrict number to be a composite number N, to find a non-trivial divisor of N. Before finding such divisor, we could use primality-testing algorithms to verify whether N is indeed composite.**
- **We could further restrict this number to be odd (otherwise 2 is a divisor) and not to be any power of a prime (otherwise that prime is a divisor). Thus, we could use two process to accomplish this requirement.**
- **Firstly, we could check whether N is odd, i.e., $N \equiv 0\ (mod\ 2)$.**
- **Second, we check whether N is of the form q^k for any $k \geq 2$; notice that q need not be prime. Since any N which passes through the first test is odd, N is only likely to be the power of a number q with $q \geq 3$. Here, we merely need to check whether $\sqrt[k]{N}$ is a whole number for any k with $2 \leq k \leq log_3^N$.**

Explanation for the range of k

Of course, 2 is the smallest value to consider. On the other hand,

$$N = q^k \rightarrow log_3 N = log_3 q^k = klog_3 q \rightarrow k = \frac{log_3 N}{log_3 q} \leq \frac{log_3 N}{log_3 3} = log_3 N$$

Here, we have used $q \geq 3$.

- Thus, if N passes both these tests, we could assure that N isn't even and that N doesn't satisfy $N = p^k$ for any prime number p.
- Eventually, we may assume that N is the product of two coprime integers greater than 2. It follows from the Chinese remainder theorem that there are at least four distinct square roots of 1 modulo N.

Digress: Chinese remainder theorem

- **Given a set of modular equations s.t.**

$$if \begin{cases} x \equiv a_1 \ (mod\ m_1) \\ x \equiv a_2 \ (mod\ m_2) \\ \vdots \\ x \equiv a_n \ (mod\ m_n) \end{cases} \forall i.j, i \neq j, \qquad gcd(m_i, m_j) = 1,$$

$$then\ x\ has\ a\ unique\ sol\ in \{0, 1, \dots, m-1\}, where\ m = m_1 m_2 \dots m_n$$

- **Then, by the assumption stating that *N* is the product of two coprime integers greater than 2, hence, we could get the following equation:**

$$\begin{cases} b^2 \equiv 1 \ (mod\ p) \\ b^2 \equiv 1 \ (mod\ q) \end{cases}, where\ pq = N\ and\ gcd(p, q) = 1$$

Then, furthermore, we could get the following two modulo equations:

$$\begin{cases} b \equiv \pm 1 \ (mod\ p) \\ b \equiv \pm 1 \ (mod\ q) \end{cases}$$

By Chinese remainder theorem, we could at least get four different roots of 1 modulo N.

Why are we so care about such number?

If $b^2 \equiv 1 \ (mod\ N)$, $b^2 - 1 \equiv (b+1)(b-1) \ (mod\ N) \equiv 0 \ (mod\ N)$. Besides, if we require that $N \nmid b-1$ & $b+1$, then N must have a nontrivial common factor with each of $b-1$ and $b+1$. i.e., $gcd(N, b-1) \neq 1$ & $gcd(N, b+1) \neq 1$.

<proof>:

Here, for simplicity, we denote $b-1$ and $b+1$ by u and v, respectively.

Since $N \mid uv$, there exists some integer k s.t. $uv = kN$. We would prove it by contradiction. Suppose $gcd(u, N) = 1$; then by Bézout's identity, there exist some integers m and n s.t. $mu + nN = 1$. Now, multiplying both sides by v, we find that $muv + nvN = mkN + nvN = v$, so $N \mid v$, which violates our premises. By contradiction, $gcd(u, N) \neq 1$. By a similar argument, $gcd(v, N) \neq 1$.

- Our main goal would be finding out such b satisfying the above two requirements

Overview of Shor’s algorithm

- **Shor’s algorithm consists of two parts:**

1. **A reduction, which can be done on a classical computer, of the factoring problem to the problem of order-finding (which is to find out previous discussed b).**
2. **A quantum algorithm to solve the order-finding problem.**

Classical part

1. **Pick a random number $a < N$**
2. **Compute $gcd(a, N)$, the greatest common divisor of a and N, which could be done by Euclidean algorithm**

Digress: Euclidean Algorithm

Given a, b, where $a > b \geq 0$, our goal is to find out $gcd(a, b)$

Let $r_0 = a, r_1 = b$. Then we could get $gcd(a, b) = gcd(r_0, r_1)$

$$\begin{cases} r_0 = r_1 q_1 + r_2 & \rightarrow gcd(r_0, r_1) = gcd(r_1, r_2) \\ r_1 = r_2 q_2 + r_3 & \rightarrow gcd(r_1, r_2) = gcd(r_2, r_3) \\ & \vdots \\ r_{i-1} = r_i q_i & \rightarrow gcd(r_{i-2}, r_{i-1}) = gcd(r_{i-1}, r_i) \end{cases}$$

Continuing

3. If $gcd(a, N) \neq 1$, then this number is a nontrivial factor of N, so we done.
4. Otherwise, use the quantum period-finding subroutine (introduced latter) to find r, which denotes the period of the following function:

$$f(x) = a^x \ (mod\ N)$$

This is the order r of a in the group $(\mathbb{Z}_N)^\times$ (which is $\{0, 1, \ldots, N-1\}$), which is the smallest positive integer r for which $f(x+r) = f(x)$, or $f(x+r) = a^{x+r}\ mod\ N \equiv a^x\ mod\ N$. i.e., $a^r \equiv 1\ (mod\ N)$ as $gcd(a,\ N) = 1$. Besides, $r < N$, since $(\mathbb{Z}_N)^\times$ contains finite elements, thus, at most $r = N - 1$.

5. If r is odd, then go back to step 1.
6. If $a^{\frac{r}{2}} \equiv -1\ (mod\ N)$, then go back to step 1.
7. Otherwise, both $gcd(a^{\frac{r}{2}} + 1, N)$ and $gcd(a^{\frac{r}{2}} - 1, N)$ is a nontrivial factor of N, so we are done.

For example…

- Given $N = 15, a = 7$, and $f(x) = 7^x \ (mod\ 15)$, then:

x	$f(x)$
1	7
2	4
3	13
4	1
5	7

from the left hand side example, we could know that $r = 4$. Moreover, we have $gcd(7^2 \pm 1, 15) = gcd(49 \pm 1, 15)$, where $gcd(48, 15) = 3$ and $gcd(50, 15) = 5$. We factorize out $N = 15$.

Why don't we have to check whether $a^{r/2} - 1$ is divided by N?

Ans:

Since r is the smallest number for one circle, thus, if $a^{r/2} - 1$ is divided by N, then r isn't the smallest number for one circle, which contradicts to the requirement.

Quantum part: Overview of period-finding

- **It is just like Simon's algorithm. First, we would set up all zero state at first registers, then apply a unitary transform, lastly, apply a quantum Fourier transform to first register, and read out. Here $U_f = \sum_{x\in\{0,1\}^q, y\in\{0,1\}^n} |x\rangle\langle x| \otimes |y \oplus f(x)\rangle\langle y|$**
- **Given N, find $Q = 2^q$ s.t. $N^2 \leq Q < 2N^2$, which implies that $\frac{Q}{r} > N$. Then, the second register only needs n ($> log_2 N$) output qubits.**

Proceed as follows:

- **Initialize the registers, and apply the Hadamard transform to the first register, as follows:**

$$\boldsymbol{H^{\otimes q} I^{\otimes n} |0^q, 0^n\rangle = \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle \otimes |0^n\rangle}$$

- **Then apply unitary transform on the register, which is as following:**

$$\boldsymbol{U_f \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle \otimes |0^n\rangle}$$

$$\boldsymbol{= \left(\sum_{x \in \{0,1\}^q, y \in \{0,1\}^n} |x\rangle\langle x| \otimes |y \oplus f(x)\rangle\langle y|\right) \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x\rangle \otimes |0^n\rangle = \frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} |x, f(x)\rangle}$$

Continuing…

Then, apply the quantum Fourier transform to the first register, showing that:

$$U_{QFT}\frac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x,f(x)\rangle=\frac{1}{Q}\sum_{x=0}^{Q-1}\sum_{y=0}^{Q-1}\omega^{xy}|y,f(x)\rangle$$

$$=\frac{1}{Q}\sum_{z=0}^{N-1}\sum_{y=0}^{Q-1}\sum_{x\in\{0,1,\ldots,Q-1\};f(x)=z}\omega^{xy}|y,z\rangle$$

Where $\omega=e^{\frac{2\pi i}{Q}}$ and r is the period of f

- Define x_0 to be the smallest of the $x\in\{0,1,\ldots,Q-1\}$ for which $f(x)=z$ (we actually have $x_0<r$), $m-1=\left\lfloor\frac{Q-x_0-1}{r}\right\rfloor$, and b to index these x, running from 0 to $m-1$, so that $x_0+rb<Q$
- Therefore, for the specific $|y,z\rangle$, its coefficient is

$$\frac{1}{Q}\sum_{x\in\{0,\ldots,Q-1\};f(x)=z}\omega^{xy}=\frac{1}{Q}\sum_{b=0}^{m-1}\omega^{(x_0+rb)y}=\frac{1}{Q}\omega^{x_0y}\sum_{b=0}^{m-1}\omega^{rby}$$

Continuing…

- **Perform a measurement. We obtain some outcome y in the input register and some outcome z in the output register. The probability of measuring some state $|y, z\rangle$ is given by the following:**

$$Pr(|y, z\rangle) = \left|\frac{1}{Q}\omega^{x_0 y}\sum_{b=0}^{m-1}\omega^{rby}\right|^2 = \frac{1}{Q^2}\left|\sum_{b=0}^{m-1}\omega^{rby}\right|^2 = \frac{1}{Q^2}\left(\frac{\omega^{mry}-1}{\omega^{ry}-1}\right)^2$$

$$= \frac{1}{Q^2}\left(\frac{\sin\frac{\pi mry}{Q}}{\sin\frac{\pi ry}{Q}}\right)^2$$

- **The above result shows that this probability is higher as the closer the unit vector ω^{ry} is to the positive real axis, or the closer $\frac{yr}{Q}$ is to an integer.**

Continuing…

- Since $\frac{yr}{Q}$ is close to some integer c, the known value $\frac{y}{Q}$ is close to the unknown value $\frac{c}{r}$. Performing [classical] continued fraction expansion on $\frac{y}{Q}$ allows us to find the approximations $\frac{d}{s}$ of it that satisfy two conditions:

A. $s < N$

B. $\left|\frac{y}{Q} - \frac{d}{s}\right| \leq \frac{1}{2Q}$

s is very likely to be the appropriate period r, or at least a factor of it.

Digress: continued fraction expansion

For example, $\frac{427}{512} = 0 + \frac{1}{1+\frac{1}{5+\frac{1}{42+\frac{1}{2}}}}$ $[0, 1, 5, 42, 2]$, and if x is a rational number, then this would converge to some sequence.

Realizing the criterion

- Recall that the final state is $\frac{1}{Q}\sum_{x=0}^{Q-1}\sum_{y=0}^{Q-1}\omega^{xy}|y,f(x)\rangle$. Besides, the probability of getting such $|y,z\rangle$ is $Pr(|y,z\rangle)=\left|\frac{1}{Q}\omega^{x_0 y}\sum_{b=0}^{m-1}\omega^{rby}\right|^2=\frac{1}{Q^2}\left|\sum_{b=0}^{m-1}\omega^{rby}\right|^2$. Here, we could replace ry with $\{ry\}_Q$, where $\{ry\}_Q$ is the residue which is congruent to $ry\ (mod\ Q)$ and is in the range $-\frac{Q}{2}<\{ry\}_Q\le\frac{Q}{2}$. This leaves us with the expression $\frac{1}{Q^2}\left|\sum_{b=0}^{m-1}\omega^{b\{ry\}_Q}\right|^2$, which could be turned into an integral

$$\frac{1}{Q}\int_0^{\left\lfloor\frac{Q-x_0-1}{r}\right\rfloor}\omega^{b\{ry\}_Q}\,db+O(\left\lfloor\frac{Q-x_0-1}{r}\right\rfloor/Q(\omega^{b\{ry\}_Q}-1))$$

- If $\left|\{ry\}_Q\right|\le r/2$, the error term in the above expression is bounded by $O(1/Q)$. We now show that if $\left|\{ry\}_Q\right|\le r/2$, the above integral is large, so the probability of obtaining a state $|y,z\rangle$ is large. Substituting $u=rb/q$, the above integral become

$$\frac{1}{r}\int_0^{\frac{r}{Q}\left\lfloor\frac{Q-x_0-1}{r}\right\rfloor}\omega^{\frac{\{ry\}_Q Q}{r}u}\,du$$

- If we approximate the upper limit of integration by 1, which only results in $O(1/Q)$

$$\frac{1}{r}\int_0^1 exp\left(\frac{2\pi i\{ry\}_Q}{r}u\right)du$$

Realizing criterion

- Letting $\{ry\}_Q/r$ vary between $-\frac{1}{2}$ and $\frac{1}{2}$, the absolute magnitude of the last integral is minimized when $\frac{\{ry\}_Q}{r} = \pm\frac{1}{2}$, which is $2/\pi r$. Thus, the probability is thus asymptotically bounded below by $4/\pi^2 r^2$, and so is at least $1/3r^2$ for sufficiently large N.
- The probability of getting such state $|y, z\rangle$ will thus be at least $1/3r^2$ if

$$-\frac{r}{2} \leq \{ry\}_Q \leq \frac{r}{2}$$

i.e., if there is a d s.t.

$$-\frac{r}{2} \leq ry - dQ \leq \frac{r}{2}$$

Dividing by rQ and rearranging the terms gives

$$\left|\frac{y}{q} - \frac{d}{s}\right| \leq \frac{1}{2Q}$$

Realizing the criterion

Theorem: let ξ be a real number, and let a and b be integers with $b > 0$. If

$$\left|\xi - \frac{a}{b}\right| \leq \frac{1}{2b^2}$$

then the rational number a/b is a convergent of the continued fraction expansion of ξ.

- **Therefore, by applying the theorem to our criterion,**

$$\left|\frac{y}{Q} - \frac{d}{s}\right| < \frac{1}{2Q} \leq \frac{1}{2s^2}, if\ s\ is\ the\ period\ of\ the\ function\ f$$

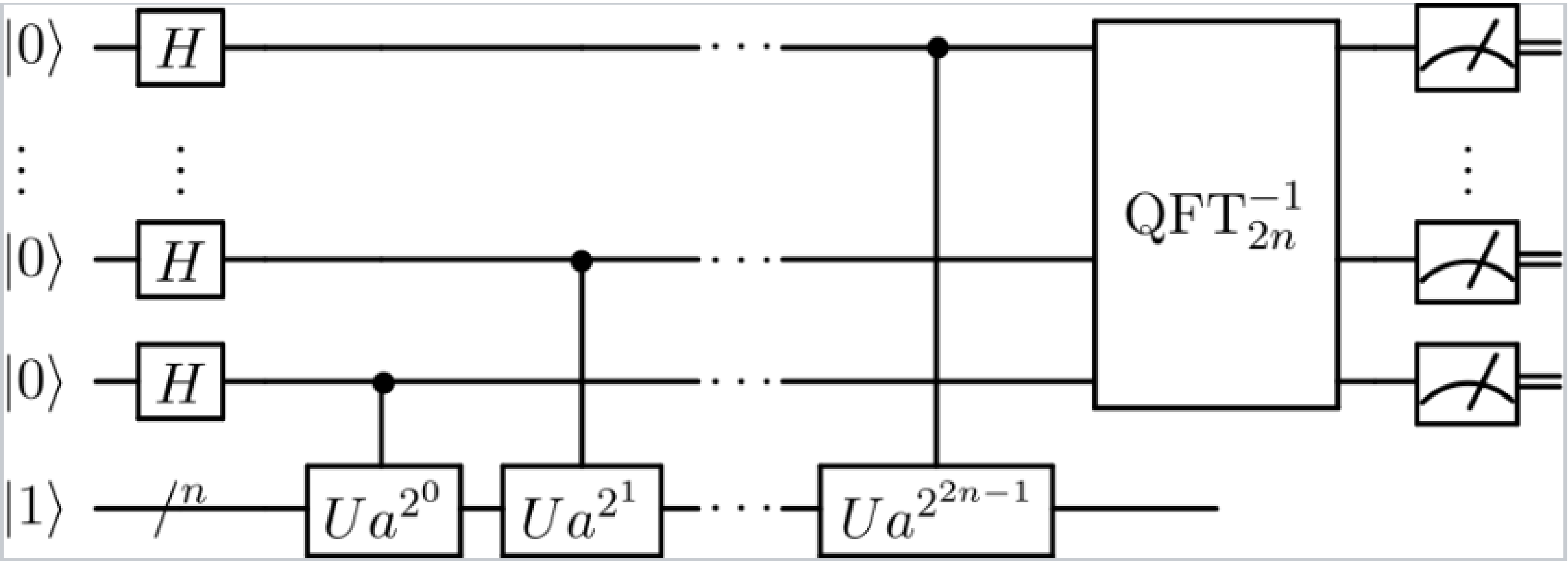
Finally,

- Check (classically) if $f(x) = f(x+s) \Leftrightarrow a^s \equiv 1\ (mod\ N)$. If so, then we are done.
- Otherwise, (classically) obtain more candidates for r by using multiples of s or by using other s with $\frac{d}{s}$ near $\frac{y}{Q}$. If any candidate works, then we are done.
- Otherwise, try again starting from step 1 of this subroutine.

For example

- If $N = 15$ (implies $Q = 2^8 = 256$) and $a = 4$, here, we know that $4^2 \equiv 1 \ (mod\ 15)$. Therefore, by the last discussion, we could know that $\frac{yr}{Q} = \frac{128\times 2}{256} = 1$. The most likely readout would be $\mathrm{y} = 128$ ($\rightarrow \frac{\mathrm{y}}{\mathrm{Q}} = \frac{128}{256} = \frac{1}{2}$). Applying continued fraction expansion to $\frac{y}{Q}$, we could know the only possible value for $\frac{d}{s}$ is $\frac{1}{2}$, which exactly is the period of function f.

Phase estimation



Unitary operator

U will be defined as:

$$U|y\rangle = |xy \mod N\rangle$$

Its eigenvalue and the eigenvector of *U* are (with proof here):

$$U|u_s\rangle = \underbrace{\exp\left[\frac{2\pi i s}{r}\right]}_{\text{eigenvalue}} |u_s\rangle \qquad \text{with} \qquad \underbrace{|u_s\rangle}_{\text{eigenvector}} = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \exp\left[\frac{-2\pi i s k}{r}\right] |x^k \mod N\rangle.$$

In the proof, we introduce the period *r* to simplify the expression. That establishes a relationship between the eigenvalue and the period of *f*. i.e.

$$\text{period of } |x^a \mod N\rangle = r$$

and

$$U|u_s\rangle = \exp\left[\frac{2\pi i s}{r}\right] |u_s\rangle$$

With U defined as:

$$U|y\rangle = |xy \mod N\rangle$$

The corresponding eigenvalue and the eigenvector of U are:

$$U|u_s\rangle = \exp\left[\frac{2\pi i s}{r}\right]|u_s\rangle \qquad \text{with} \qquad |u_s\rangle = \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\exp\left[\frac{-2\pi i s k}{r}\right]|x^k \mod N\rangle.$$

Prove:

$$\underline{U|u_s\rangle} = U\underbrace{\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\exp\left[\frac{-2\pi i k t}{r}\right]|x^k \bmod N\rangle}_{|u_s\rangle} = \underbrace{\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\exp\left[\frac{-2\pi i k t}{r}\right]|x^{k+1} \bmod N\rangle}_{\text{apply } U|y\rangle = |xy \mod N\rangle}$$

$$= \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\exp\left[\frac{-2\pi i (k-1) t}{r}\right]|x^k \bmod N\rangle = \underline{\exp\left[\frac{-2\pi i t}{r}\right]|u_s\rangle}$$

if r is the period, $x^0 = x^r$. We can shift k → k - 1

$|u_s\rangle$ is eigenvector of U

Create a superposition with all eigenvectors.

$$\frac{1}{\sqrt{r}}\sum_{t=0}^{r-1}|v_t\rangle = \frac{1}{\sqrt{r}}\sum_{t=0}^{r-1}\frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}\exp\left[\frac{-2\pi ikt}{r}\right]|x^k \bmod N\rangle$$

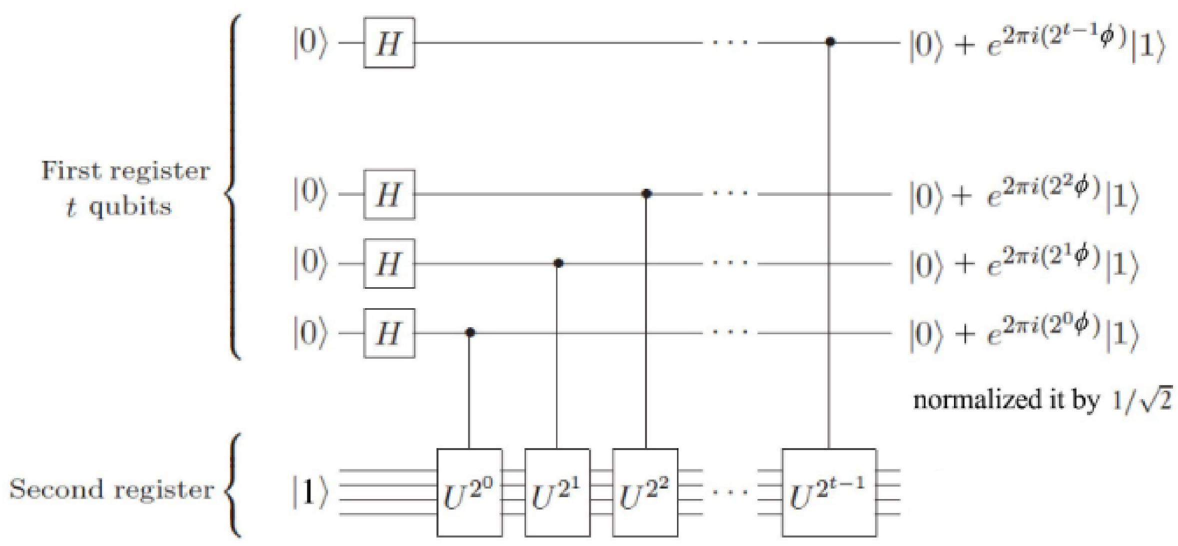
which, using $\sum_{t=0}^{r-1}\exp\left[\frac{-2\pi ikt}{r}\right] = r\delta_{k,0}$ becomes,

$$\frac{1}{\sqrt{r}}\sum_{t=0}^{r-1}|v_t\rangle = |1\rangle$$

We use t Hadamard gates to prepare the first register into a uniform superposition. And prepare the second register to be $|1\rangle$.

$$\frac{1}{\sqrt{2^t}} \sum_{j=0}^{2^t-1} |j\rangle |1\rangle$$

Then we apply a series of Controlled-U gates.



Modified from source

This will bring the system to

$$\frac{1}{\sqrt{r2^t}} \sum_{s=0}^{r-1} \sum_{j=0}^{2^t-1} e^{2\pi isj/r} |j\rangle |u_s\rangle$$

The circuit above can be viewed as a nice approximation to:

$$|\psi\rangle = \sum_{j=0}^{2^t-1} |j\rangle U^j |1\rangle = \sum_{j=0}^{2^t-1} |j\rangle |x^j \text{ mod } N\rangle$$

We apply the inverse Quantum Fourier transform to the first register, the superposition becomes

$$\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|s/r\rangle|u_s\rangle$$

Here is the summary of the whole flow:

1.	$\|0\rangle\|1\rangle$	initial state
2.	$\rightarrow \frac{1}{\sqrt{2^t}} \sum_{j=0}^{2^t-1} \|j\rangle\|1\rangle$	create superposition
3.	$\rightarrow \frac{1}{\sqrt{2^t}} \sum_{j=0}^{2^t-1} \|j\rangle\|x^j \bmod N\rangle$	apply $U_{x,N}$
	$\approx \frac{1}{\sqrt{r2^t}} \sum_{s=0}^{r-1} \sum_{j=0}^{2^t-1} e^{2\pi isj/r}\|j\rangle\|u_s\rangle$	
4.	$\rightarrow \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} \|s/r\rangle\|u_s\rangle$	apply inverse Fourier transform to first register
5.	$\rightarrow s/r$	measure first register
6.	$\rightarrow r$	apply continued fractions algorithm