

Quantum Counting: Algorithm and Error Distribution

Zijian Diao · Chunfeng Huang · Ke Wang

Received: 29 April 2011 / Accepted: 3 October 2011 / Published online: 10 February 2012
© Springer Science+Business Media B.V. 2012

Abstract Counting is one of the most basic procedures in mathematics and statistics. In statistics literature it is usually done via the proportion estimation method. In this article we manifest a radically different counting procedure first proposed in the late 1990's based on the techniques of quantum computation. It combines two major tools in quantum computation, quantum Fourier transform and quantum amplitude amplification, and shares similar structure to the quantum part of the celebrated Shor's factoring algorithm. We present complete details of this quantum counting algorithm and the analysis of its error distribution. Comparing it with the conventional proportion estimation method, we demonstrate that this quantum approach achieves much faster convergence rate than the classical approach.

Keywords Quantum counting · Quantum Fourier transform · Quantum amplitude amplification · Proportion estimation

Mathematics Subject Classification (2010) 81P68 · 68Q12 · 62E20 · 62F10

To Prof. Goong Chen on the occasion of his 60th birthday.

Z. Diao (✉)

Ohio University Eastern Campus, 45425 National Road West, St Clairsville, OH 43950, USA
e-mail: diao@ohio.edu

C. Huang

Department of Statistics, Indiana University-Bloomington, Statistics House, 309 North Park Ave,
Bloomington, IN 47408, USA
e-mail: huang48@indiana.edu

K. Wang

Department of Statistics, Fudan University, 670 Guoshun Road, Yangpu District, Shanghai, 200633,
P.R. China
e-mail: kewang@fudan.edu.cn

1 Introduction

Counting is one of the most basic procedures in mathematics and statistics. It answers the following question: given a population with N items, how many are there satisfying certain criteria? The counting problem can be formulated mathematically as: given an oracle function

$$f : \{1, 2, 3, \dots, N\} \rightarrow \{0, 1\}, \quad (1)$$

where $N \in \mathbb{N}$, find t , the number of $x \in \{1, 2, 3, \dots, N\}$ such that $f(x) = 1$. N is known, thus, finding t is equivalent to finding $p = \frac{t}{N}$, the probability of getting an x with $f(x) = 1$ when x is picked randomly. The obvious classical approach is to use a sample of size M , then check them one by one by evaluations of f (oracle calls) and count how many of them satisfy $f(x) = 1$. Call this number s . The estimation \hat{p} for p is simply $\frac{s}{M}$. This approach is usually termed proportion estimation in statistics literature. The property of such estimation can be described through binomial/hypergeometric distribution [3] and approximated via the Central Limit Theorem (CLT) when M is large. In this article, we examine the error distribution of a different estimator \tilde{p} , which is obtained from a quantum counting procedure developed in [2] and [8]. We demonstrate that this quantum counting approach achieves much faster convergence rate than the classical one.

2 A Primer to Quantum Computation

To facilitate the understanding of the quantum method in this article, we first give a brief primer on the general principles of quantum computation. Please see [4] and [9] for more detailed introduction to quantum computation. There are fundamental differences between classical computation, which can be modeled by classical Turing machines, and quantum computation, which also incorporates the principles of quantum mechanics. We summarize them here at two levels: computation model and hardware.

At the computation model level, quantum computation is very close to the classical probabilistic computation at first glance. Starting from the input, i.e., a complete specification of the initial state of the computer and data, both models for quantum computation and classical probabilistic computation allow more than one possible computation paths, which might lead to more than one possible outputs. Each output happens with certain probability. The difference in these two models lies in how the probabilities of different computation paths are specified. In the classical version, we specify directly the transition probability from one state to another, while in the quantum version, we specify the *probability amplitude* instead, which is a complex number. In quantum computation, it is the square of the modulus of the probability amplitude that gives rise to the transition probability. This seemingly insignificant difference actually results in surprising outcomes, as shown in Example 1.

Example 1 Suppose that from the same input, there are two paths that lead to the same output (cf. Fig. 1) via two different intermediate states, where P_i 's and A_i 's are respectively the transition probabilities and probability amplitudes with $|A_i|^2 = P_i$, $i = 1, 2, 3, 4$. In the classical probabilistic computation model, the probability of this output is $P_1P_2 + P_3P_4$. In the quantum computation model, the probability amplitude for this output is $A_1A_2 + A_3A_4$, which gives rise to probability $P_1P_2 + P_3P_4 + 2\text{Re}(A_1A_2A_3^*A_4^*)$. The additional term $2\text{Re}(A_1A_2A_3^*A_4^*)$ results in different behavior from the classical probabilistic computation.

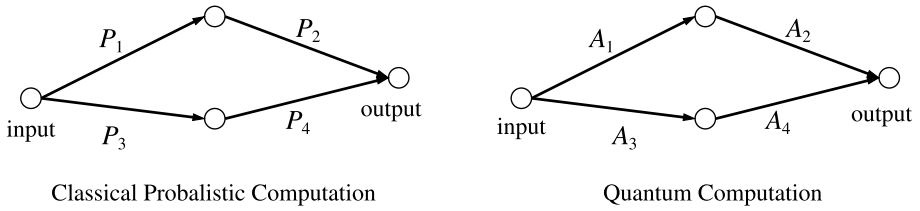
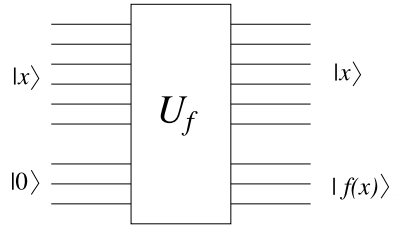


Fig. 1 Classical probabilistic computation vs. quantum computation

Fig. 2 Quantum parallelism



We can design a constructive inference such that $2Re(A_1 A_2 A_3^* A_4^*) > 0$ to enhance the correct outputs. We can also design a destructive inference such that $2Re(A_1 A_2 A_3^* A_4^*) < 0$ to suppress the wrong outputs.

At the hardware level, quantum computers also share many features of classical computers. A classical computer stores data in bits and manipulate data with logical gates. A quantum computer stores data in quantum bits (*qubits*) and manipulate data with quantum logical gates. A qubit is a microscopic system such as a spin- $\frac{1}{2}$ particle, a photon with polarization, or an atom with multiple energy levels. Similar to a classical bit, a qubit can encode 0 and 1 in two distinct states (written as $|0\rangle$ and $|1\rangle$ in the bra-ket notation). For example, an atom in the ground state and the excited state can represent 0 and 1, respectively. But the similarity ends here. A qubit can also be in a superposition of the basis states $|0\rangle$ and $|1\rangle$, i.e., $a|0\rangle + b|1\rangle$, where $a, b \in \mathbb{C}$ are probability amplitudes with $|a|^2 + |b|^2 = 1$. A measurement of the qubit yields either $|0\rangle$ or $|1\rangle$, with probability $|a|^2$ and $|b|^2$, respectively. A register with n qubits can be in the superposition of 2^n basis states ranging from $|00\dots 0\rangle$ to $|11\dots 1\rangle$. So the amount of numbers a quantum register can store *simultaneously* grows exponentially relative to its size, which is impossible for a classical register. A quantum logical gate is an elementary quantum device which preforms a unitary operation on qubits. The unitary restriction is rooted in Schrodinger’s equation, which governs the evolution of quantum systems. A simple example of one-bit quantum gate is NOT gate, also a rudimentary classical logic gate, which maps $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$. Other examples includes Walsh-Hadamard gate $H: |0\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), |1\rangle \rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, and Controlled-NOT gate: $|00\rangle \rightarrow |00\rangle, |01\rangle \rightarrow |01\rangle, |10\rangle \rightarrow |11\rangle, \text{ and } |11\rangle \rightarrow |10\rangle$. One can show that the set of one-qubit quantum gates and Controlled-NOT gate suffices to implement quantum circuits of any size.

Finally, let’s highlight one important source of the unorthodox power of quantum computation, *quantum parallelism*. Suppose we have a function evaluation operator $U_f: |x\rangle|0\rangle \rightarrow |x\rangle|f(x)\rangle$, which reads the input value x from the first register and computes the corresponding function value $f(x)$ in the second register (cf. Fig. 2). If the first register has n qubits in it and $|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{i=00\dots 0}^{11\dots 1} |i\rangle$, the uniform superposition of all 2^n basis state ranging from

$|00\dots 0\rangle$ to $|11\dots 1\rangle$, after one application of U_f , we obtain $\frac{1}{\sqrt{2^n}} \sum_{i=00\dots 0}^{11\dots 1} |i\rangle |f(i)\rangle$, where the second register holds the function values of all 2^n input values. In other words, we have completed exponentially many function evaluations in one shot via quantum parallelism. The quantum algorithm that we present later utilizes similar construction.

3 Quantum Fourier Transform

The first major ingredient we need for the quantum counting algorithm is quantum Fourier transform. It is the quantum version of the standard discrete Fourier transform. For $x \in \{0, 1, 2, \dots, M - 1\}$, define quantum Fourier transform and inverse quantum Fourier transform by

$$QFT : |x\rangle \rightarrow \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{x}{M}y} |y\rangle, \tag{2}$$

$$QFT^{-1} : |x\rangle \rightarrow \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{-2\pi i \frac{x}{M}y} |y\rangle. \tag{3}$$

Clearly,

$$QFT^{-1} \left(\frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \frac{x}{M}y} |y\rangle \right) = QFT^{-1}(QFT|x\rangle) = |x\rangle. \tag{4}$$

Suppose that instead of $\frac{x}{M}$, we have ω , a real number between 0 and 1, in (4). What is the result of applying QFT^{-1} on $|\Omega\rangle \triangleq \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \omega y} |y\rangle$? Let $|\tilde{x}\rangle = QFT^{-1}|\Omega\rangle = \sum_{x=0}^{M-1} \alpha_x |x\rangle$. There are two scenarios.

1. For some $k \in \{0, 1, 2, \dots, M - 1\}$, $\omega = \frac{k}{M}$. This case is reduced back to (4). We have $|\tilde{x}\rangle = |k\rangle$. In other words, $\alpha_x = \delta_{xk}$. If we make a measurement on $|\tilde{x}\rangle$, we will get $|k\rangle$ for sure, which tells us what ω is, namely, $\frac{k}{M}$.
2. For all $k \in \{0, 1, 2, \dots, M - 1\}$, $\omega \neq \frac{k}{M}$. We do not have a closed form expression for α_x as clean as the one in the previous case any more. Nevertheless, we still expect the α_x 's for those x 's such that $\frac{x}{M}$ is close to ω to be dominant. The following calculation verifies this claim.

$$\begin{aligned} |\tilde{x}\rangle &= QFT^{-1}|\Omega\rangle \\ &= \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \omega y} QFT^{-1}|y\rangle \\ &= \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i \omega y} \frac{1}{\sqrt{M}} \sum_{x=0}^{M-1} e^{-2\pi i \frac{y}{M}x} |x\rangle \\ &= \sum_{x=0}^{M-1} \frac{1}{M} \sum_{y=0}^{M-1} e^{2\pi i (\omega - \frac{x}{M})y} |x\rangle. \end{aligned} \tag{5}$$

Thus,

$$\alpha_x = \frac{1}{M} \sum_{y=0}^{M-1} e^{2\pi i(\omega - \frac{x}{M})y} = \frac{1 - e^{2\pi i M(\omega - \frac{x}{M})}}{M(1 - e^{2\pi i(\omega - \frac{x}{M})}}, \tag{6}$$

$$|\alpha_x| = \left| \frac{\sin M\pi(\omega - \frac{x}{M})}{M \sin \pi(\omega - \frac{x}{M})} \right|. \tag{7}$$

It is easy to see that $\lfloor M\omega \rfloor$ and $\lceil M\omega \rceil$ are the two integer values of x such that $\frac{x}{M}$ is the closest to ω . Let $\Delta = \frac{M\omega - \lfloor M\omega \rfloor}{M} = \omega - \frac{\lfloor M\omega \rfloor}{M}$ and $\frac{1}{M} - \Delta = \frac{\lceil M\omega \rceil - M\omega}{M} = \frac{\lceil M\omega \rceil}{M} - \omega$. The probability of getting an x after measurement that can best indicate ω is,

$$\begin{aligned} P\left(\left|\frac{x}{M} - \omega\right| \leq \frac{1}{M}\right) &= P(|x - M\omega| \leq 1) \\ &= P(x = \lfloor M\omega \rfloor) + P(x = \lceil M\omega \rceil) \\ &= |\alpha_{\lfloor M\omega \rfloor}|^2 + |\alpha_{\lceil M\omega \rceil}|^2 \\ &= \frac{\sin^2 M\Delta\pi}{M^2 \sin^2 \Delta\pi} + \frac{\sin^2 M(\frac{1}{M} - \Delta)\pi}{M^2 \sin^2(\frac{1}{M} - \Delta)\pi}, \end{aligned} \tag{8}$$

which attains minimum at $\Delta = \frac{1}{2M}$. Thus,

$$\begin{aligned} P\left(\left|\frac{x}{M} - \omega\right| \leq \frac{1}{M}\right) &\geq \frac{1}{M^2} \left(\frac{1}{\sin^2(\frac{\pi}{2M})} + \frac{1}{\sin^2(\frac{\pi}{2M})} \right) \\ &= \frac{2}{M^2 \sin^2(\frac{\pi}{2M})} \\ &> \frac{2}{M^2(\frac{\pi}{2M})^2} \\ &= \frac{8}{\pi^2}. \end{aligned} \tag{9}$$

If we make a measurement on $|\tilde{x}\rangle$, the probability of getting either $\lfloor M\omega \rfloor$ or $\lceil M\omega \rceil$, providing an estimation for ω within the error $\frac{1}{M}$, is at least $\frac{8}{\pi^2}$.

In both cases, a measurement after applying QFT^{-1} on $|\Omega\rangle$ yields an integer very close or equal to $M\omega$ with high probability, allowing us to estimate ω .

4 Amplitude Amplification

Let \mathbb{H} be a Hilbert space with an orthonormal basis $\{|x\rangle \mid x = 1, 2, \dots, N\}$. The oracle function f defined in (1) partitions \mathbb{H} into two orthogonal subspaces by dividing $\{1, 2, \dots, N\}$ into $f^{-1}(0)$ and $f^{-1}(1)$. Define

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=1}^N |x\rangle, \tag{10}$$

$$|m\rangle = \frac{1}{\sqrt{t}} \sum_{x \in f^{-1}(1)} |x\rangle, \tag{11}$$

$$|w\rangle = \frac{1}{\sqrt{N-t}} \sum_{x \in f^{-1}(0)} |x\rangle. \tag{12}$$

In other words, $|s\rangle$ is the uniform superposition of all basis states, $|m\rangle$ the uniform superposition of all basis states satisfying the counting criteria, and $|w\rangle$ the counterpart of $|m\rangle$. Clearly $|m\rangle \perp |w\rangle$. Choose $\theta \in (0, \frac{\pi}{2})$ such that $\sin^2 \theta = p = \frac{t}{N}$. We have $\sin \theta = \sqrt{\frac{t}{N}}$, $\cos \theta = \sqrt{\frac{N-t}{N}}$, and $|s\rangle = \cos \theta |w\rangle + \sin \theta |m\rangle$.

We now define the other major ingredient of the quantum counting algorithm, the amplitude amplification operator G :

$$G = \mathcal{I}_s \mathcal{I}_f. \tag{13}$$

The operator \mathcal{I}_f is defined by

$$\mathcal{I}_f |x\rangle = \begin{cases} -|x\rangle, & \text{if } f(x) = 1 \\ |x\rangle, & \text{if } f(x) = 0 \end{cases}. \tag{14}$$

\mathcal{I}_f is sometimes referenced as the selective sign-flipping operator, because it flips the sign of the states satisfying the counting criteria. It can be implemented via a technique called *eigenvalue kickback* [1]. Each application of this operator invokes one oracle call. There is an alternative but equivalent form to define \mathcal{I}_f by

$$\mathcal{I}_f = I - 2|m\rangle\langle m|, \tag{15}$$

where I is the identity operator. The operator \mathcal{I}_s is defined by

$$\mathcal{I}_s = 2|s\rangle\langle s| - I. \tag{16}$$

\mathcal{I}_s is sometimes referenced as the inversion-around-average operator, because it inverts state vectors around the ‘‘average’’ state $|s\rangle$. The operator G is in fact exactly the iteration operator in the original Grover’s algorithm for quantum search [7]. Now let’s analyze the effect of this operator.

Lemma 1 *In the plane spanned by basis $\{|w\rangle, |m\rangle\}$, G implements a rotation of angle 2θ .*

Proof It is straightforward to derive $G|w\rangle$ and $G|m\rangle$ under basis $\{|w\rangle, |m\rangle\}$.

$$\begin{aligned} G|w\rangle &= \mathcal{I}_s \mathcal{I}_f |w\rangle = \mathcal{I}_s |w\rangle = (2|s\rangle\langle s| - I)|w\rangle = 2|s\rangle\langle s|w\rangle - |w\rangle \\ &= 2 \cos \theta |s\rangle - |w\rangle = 2 \cos \theta (\cos \theta |w\rangle + \sin \theta |m\rangle) - |w\rangle \\ &= (2 \cos^2 \theta - 1)|w\rangle + 2 \sin \theta \cos \theta |m\rangle \\ &= \cos 2\theta |w\rangle + \sin 2\theta |m\rangle. \end{aligned}$$

$$\begin{aligned} G|m\rangle &= \mathcal{I}_s \mathcal{I}_f |m\rangle = \mathcal{I}_s (-|m\rangle) = (2|s\rangle\langle s| - I)(-|m\rangle) = -2|s\rangle\langle s|m\rangle + |m\rangle \\ &= -2 \sin \theta |s\rangle + |m\rangle = -2 \sin \theta (\cos \theta |w\rangle + \sin \theta |m\rangle) + |m\rangle \end{aligned}$$

$$\begin{aligned}
 &= -2 \sin \theta \cos \theta |w\rangle + (1 - 2 \sin^2 \theta) |m\rangle \\
 &= -\sin 2\theta |w\rangle + \cos 2\theta |m\rangle.
 \end{aligned}$$

So G takes on matrix representation

$$G = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{17}$$

Obviously it is a rotation of angle 2θ oriented from $|w\rangle$ to $|m\rangle$. □

If we start from $|s\rangle$, each application of G rotates it toward $|m\rangle$ by 2θ , amplifying the amplitude of the component $|m\rangle$. Next let's define a pair of states,¹ which play a crucial role in establishing the credibility of the quantum counting algorithm.

$$|\psi_+\rangle = \frac{1}{\sqrt{2}}(|w\rangle - i|m\rangle), \tag{18}$$

$$|\psi_-\rangle = \frac{1}{\sqrt{2}}(|w\rangle + i|m\rangle). \tag{19}$$

Lemma 2 $|\psi_+\rangle$ and $|\psi_-\rangle$ are eigenvectors of G with eigenvalues $e^{2i\theta}$ and $e^{-2i\theta}$, respectively.

Proof Via (17),

$$\begin{aligned}
 G|\psi_+\rangle &= \frac{1}{\sqrt{2}}(G|w\rangle - iG|m\rangle) = \frac{1}{\sqrt{2}}(\cos 2\theta|w\rangle + \sin 2\theta|m\rangle + i \sin 2\theta|w\rangle - i \cos 2\theta|m\rangle) \\
 &= \frac{e^{2i\theta}}{\sqrt{2}}(|w\rangle - i|m\rangle) = e^{2i\theta}|\psi_+\rangle,
 \end{aligned}$$

$$\begin{aligned}
 G|\psi_-\rangle &= \frac{1}{\sqrt{2}}(G|w\rangle + iG|m\rangle) = \frac{1}{\sqrt{2}}(\cos 2\theta|w\rangle + \sin 2\theta|m\rangle - i \sin 2\theta|w\rangle + i \cos 2\theta|m\rangle) \\
 &= \frac{e^{-2i\theta}}{\sqrt{2}}(|w\rangle + i|m\rangle) = e^{-2i\theta}|\psi_-\rangle.
 \end{aligned}$$
□

Let $\theta = \pi\omega$, $|s\rangle = \cos \pi\omega|w\rangle + \sin \pi\omega|m\rangle = \frac{e^{i\pi\omega}}{\sqrt{2}}|\psi_+\rangle + \frac{e^{-i\pi\omega}}{\sqrt{2}}|\psi_-\rangle$. If we apply G on $|s\rangle$ for y times,

$$G^y |s\rangle = \frac{e^{i\pi(2y+1)\omega}}{\sqrt{2}}|\psi_+\rangle + \frac{e^{-i\pi(2y+1)\omega}}{\sqrt{2}}|\psi_-\rangle, \tag{20}$$

which is reminiscent of the state $|\Omega\rangle$ in Sect. 3.

5 Quantum Counting Algorithm

Now we have both major ingredients ready for the construction of a quantum counting algorithm which estimates $p = \frac{1}{N}$. Since $p = \sin^2 \theta = \sin^2 \pi\omega$, we can estimate p via ω .

¹In [8], they are misdefined to be the other way around.

The key is to create a state in the form of $|\Omega\rangle$ using G , which supplies an estimation of ω through QFT^{-1} . The algorithm below follows the path given in [8].

1. Prepare two registers in the initial state $|\psi_0\rangle = \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} |y\rangle \otimes |s\rangle$.
2. Apply C_F on $|\psi_0\rangle$, which implements $|y\rangle \otimes |s\rangle \rightarrow |y\rangle \otimes G^y|s\rangle$. Call the resultant state $|\psi_1\rangle$.
3. Apply QFT^{-1} on the first register of $|\psi_1\rangle$. Call the resultant state $|\psi_2\rangle$.
4. Measure the first register of $|\psi_2\rangle$ to obtain $|x\rangle$ and output $\tilde{p} = \sin^2(\frac{x}{M}\pi)$, the quantum estimator of p .

We can show the following result, which is slightly tighter than [2, Theorem 6].

Theorem 1 $\forall M \in \mathbb{N}$, the above algorithm outputs \tilde{p} such that

$$|p - \tilde{p}| \leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1 - 2p| \tag{21}$$

with probability at least $\frac{8}{\pi^2}$.

Proof After Step 2,

$$\begin{aligned} |\psi_1\rangle &= \frac{1}{\sqrt{2M}} \sum_{y=0}^{M-1} |y\rangle (e^{\pi i(2y+1)\omega} |\psi_+\rangle + e^{-\pi i(2y+1)\omega} |\psi_-\rangle) \tag{cf. (20)} \\ &= \frac{e^{\pi i\omega}}{\sqrt{2M}} \sum_{y=0}^{M-1} e^{2\pi i\omega y} |y\rangle |\psi_+\rangle + \frac{e^{-\pi i\omega}}{\sqrt{2M}} \sum_{y=0}^{M-1} e^{-2\pi i\omega y} |y\rangle |\psi_-\rangle \\ &= \frac{e^{\pi i\omega}}{\sqrt{2M}} \sum_{y=0}^{M-1} e^{2\pi i\omega y} |y\rangle |\psi_+\rangle + \frac{e^{-\pi i\omega}}{\sqrt{2M}} \sum_{y=0}^{M-1} e^{2\pi i(1-\omega)y} |y\rangle |\psi_-\rangle \tag{22} \end{aligned}$$

After Step 3, we have

$$|\psi_2\rangle = \frac{e^{\pi i\omega}}{\sqrt{2}} |\tilde{x}_+\rangle |\psi_+\rangle + \frac{e^{-\pi i\omega}}{\sqrt{2}} |\tilde{x}_-\rangle |\psi_-\rangle, \tag{23}$$

where

$$|\tilde{x}_+\rangle = QFT^{-1} \left(\frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i\omega y} |y\rangle \right), \tag{24}$$

$$|\tilde{x}_-\rangle = QFT^{-1} \left(\frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{2\pi i(1-\omega)y} |y\rangle \right). \tag{25}$$

In Step 4, the measurement on the first register gives us an estimation of either ω or $1 - \omega$, with probability $\frac{1}{2}$ each. In the first case, we obtain $|x\rangle$ such that $\Delta = |\frac{x}{M} - \omega| \leq \frac{1}{M}$ with probability at least $\frac{8}{\pi^2}$, while

$$\begin{aligned}
 |\tilde{p} - p| &= \left| \sin^2\left(\pi \frac{x}{M}\right) - \sin^2(\pi\omega) \right| \\
 &= \left| \sin^2(\pi\omega \pm \pi\Delta) - \sin^2(\pi\omega) \right| \\
 &= \left| (\sin(\pi\omega) \cos(\pi\Delta) \pm \sin(\pi\Delta) \cos(\pi\omega))^2 - \sin^2(\pi\omega) \right| \\
 &= \left| \sin^2(\pi\Delta) \cos(2\pi\omega) \pm \sin(\pi\omega) \cos(\pi\omega) \sin(2\pi\Delta) \right| \\
 &\leq \left| \sin(2\pi\Delta) \sin(\pi\omega) \cos(\pi\omega) \right| + \sin^2(\pi\Delta) |1 - 2\sin^2(\pi\omega)| \\
 &\leq 2\pi\Delta \sqrt{p(1-p)} + (\pi\Delta)^2 |1 - 2p| \\
 &\leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1 - 2p|.
 \end{aligned} \tag{26}$$

In the second case, we obtain $|x\rangle$ such that $\Delta = \left| \frac{x}{M} - (1 - \omega) \right| \leq \frac{1}{M}$ with probability at least $\frac{8}{\pi^2}$, while

$$\begin{aligned}
 |\tilde{p} - p| &= \left| \sin^2\left(\pi \frac{x}{M}\right) - \sin^2(\pi\omega) \right| \\
 &= \left| \sin^2(\pi(1 - \omega) \pm \pi\Delta) - \sin^2(\pi\omega) \right| \\
 &= \left| \sin^2(\pi\omega \pm \pi\Delta) - \sin^2(\pi\omega) \right| \\
 &\leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1 - 2p|.
 \end{aligned} \tag{27}$$

Either way, we obtain an estimation of p with the same error behavior. □

With \tilde{p} , the estimation of p at hand, we can estimate t , the number of counting targets via $\tilde{t} = N\tilde{p}$. It is straightforward to derive the following result.

Corollary 1 $\forall M \in \mathbb{N}$,

$$|t - \tilde{t}| \leq \frac{2\pi}{M} \sqrt{t(N-t)} + \frac{\pi^2}{M^2} |N - 2t| \tag{28}$$

with probability at least $\frac{8}{\pi^2}$.

6 Comparison with the Classical Sampling Method

In this section, we compare the classical estimator \hat{p} (see Sect. 1) and the quantum estimator \tilde{p} .

In proportion estimation, $\hat{p} = \frac{s}{M}$, where s is the number of x 's satisfying $f(x) = 1$ in the sample. Sampling can be done either with or without replacement. When sampling with replacement, s follows binomial distribution, with

$$P(s = x) = \binom{M}{x} p^x (1-p)^{M-x}, \quad x = 0, 1, 2, \dots, M. \tag{29}$$

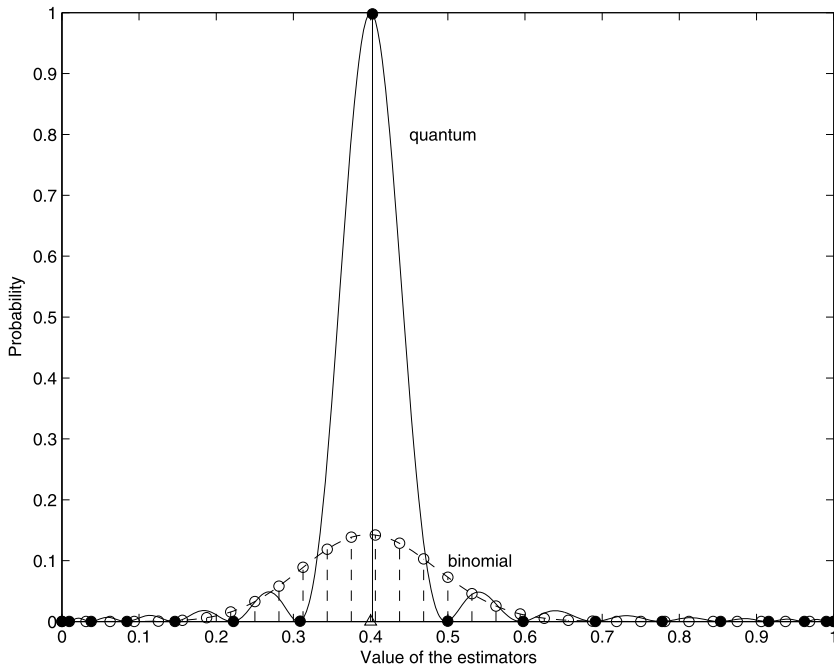


Fig. 3 The distributions of the classical estimator \hat{p} (binomial) and the quantum estimator \tilde{p} , where $N = 100$, $t = 40$, and $M = 32$. The true value of p is marked with Δ

When sampling without replacement, s follows hypergeometric distribution, with

$$P(s = x) = \frac{\binom{t}{x} \binom{N-t}{M-x}}{\binom{N}{M}}, \quad \max(0, M + t - N) \leq x \leq \min(M, t). \tag{30}$$

In the typical setup for the proportion estimation problem, where the population size is far greater than the sample size, i.e., $N \gg M$, these two distributions behave very similarly. However, the quantum estimator \tilde{p} takes on a drastically different distribution, where $\tilde{p} = \sin^2(\frac{x}{M}\pi)$, and x follows distribution

$$P(x) = \frac{\sin^2 M\pi(\omega - \frac{x}{M})}{M^2 \sin^2 \pi(\omega - \frac{x}{M})}, \quad x = 0, 1, 2, \dots, M - 1. \tag{31}$$

In the case of estimating $1 - \omega$ (cf. Step 4 in the proof of Theorem 1), the distribution of x is mirror symmetric to (31). But the distribution of \tilde{p} is identical, because $\sin^2(\frac{x}{M}\pi) = \sin^2(\pi - \frac{x}{M}\pi)$. Figure 3 shows the contrast between the distributions of \hat{p} and \tilde{p} (binomial) where $N = 100$, $t = 40$, and $M = 32$.

We can observe that, when M is large, the distribution of \tilde{p} concentrates more heavily around the true value of p than that of \hat{p} . This feature can be justified theoretically via the Central Limit Theorem (CLT) [6, p. 112]. In proportion estimation, by CLT, as $M \rightarrow \infty$,

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/M}} \rightarrow_d N(0, 1), \tag{32}$$

where \rightarrow_d denotes the convergence in distribution, and $N(0, 1)$ is the standard normal distribution. To compare \hat{p} with \tilde{p} , subject \hat{p} to the same bound in Theorem 1. For any fixed p bounded away from 0 and 1, as $M \rightarrow \infty$,

$$\begin{aligned} &P\left(|p - \hat{p}| \leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1-2p|\right) \\ &= P\left(\left|\frac{\hat{p} - p}{\sqrt{p(1-p)/M}}\right| \leq \frac{2\pi}{\sqrt{M}} + \frac{\pi^2}{M^{3/2}} \frac{|1-2p|}{\sqrt{p(1-p)}}\right) \\ &\rightarrow 2\Phi\left(\frac{2\pi}{\sqrt{M}} + \frac{\pi^2}{M^{3/2}} \frac{|1-2p|}{\sqrt{p(1-p)}}\right) - 1, \end{aligned} \tag{33}$$

where Φ is the standard normal cumulative distribution function. It is clear that for any fixed $p \in (0, 1)$, when the sample size M increases, this probability shrinks to 0. In contrast, the quantum estimator \tilde{p} subjected to the same bound happens with probability at least $\frac{8}{\pi^2}$ no matter what M is by Theorem 1. That is, when M is large,

$$\begin{aligned} &P\left(|p - \tilde{p}| \leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1-2p|\right) \\ &\gg P\left(|p - \hat{p}| \leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1-2p|\right). \end{aligned}$$

In addition, there is another angle which shows better error behavior of \tilde{p} than that of \hat{p} when M is large. Suppose that we enforce the probability of getting a close estimate of p within error ϵ to be at least $\frac{8}{\pi^2}$. In the case of \tilde{p} , ϵ goes down in the order of $O(\frac{1}{M})$. In the case of \hat{p} , because of (32) and $P(|X| < 1.32) \approx \frac{8}{\pi^2}$ when X is standard normal,

$$P\left(|p - \hat{p}| \leq \frac{1.32}{\sqrt{M}} \sqrt{p(1-p)}\right) \approx \frac{8}{\pi^2}, \tag{34}$$

which indicates that ϵ goes down at most in the order of $O(\frac{1}{\sqrt{M}})$.

Finally, for any fixed $p \in (0, 1)$, we may numerically compare the probability of \tilde{p} and \hat{p} (following either binomial or hypergeometric distribution) under the same bound in Theorem 1. The former is at least $\frac{8}{\pi^2}$, while the latter is

$$\begin{aligned} &P\left(|p - \hat{p}| \leq \frac{2\pi}{M} \sqrt{p(1-p)} + \frac{\pi^2}{M^2} |1-2p|\right) \\ &= P\left(|Mp - M\hat{p}| \leq 2\pi \sqrt{p(1-p)} + \frac{\pi^2}{M} |1-2p|\right) \\ &= \sum_{x=\lceil Mp-2\pi\sqrt{p(1-p)}-\frac{\pi^2}{M}|1-2p|\rceil}^{\lfloor Mp+2\pi\sqrt{p(1-p)}+\frac{\pi^2}{M}|1-2p|\rfloor} P(s = x), \end{aligned} \tag{35}$$

where $P(s = x)$ is given by either (29) or (30). These probabilities are shown in Fig. 4 for various values of p (with binomial distribution). We again observe that the probabilities related to \hat{p} shrink to 0 as M increases. Table 1 summarizes the values of M where \tilde{p} starts to outperforms \hat{p} .

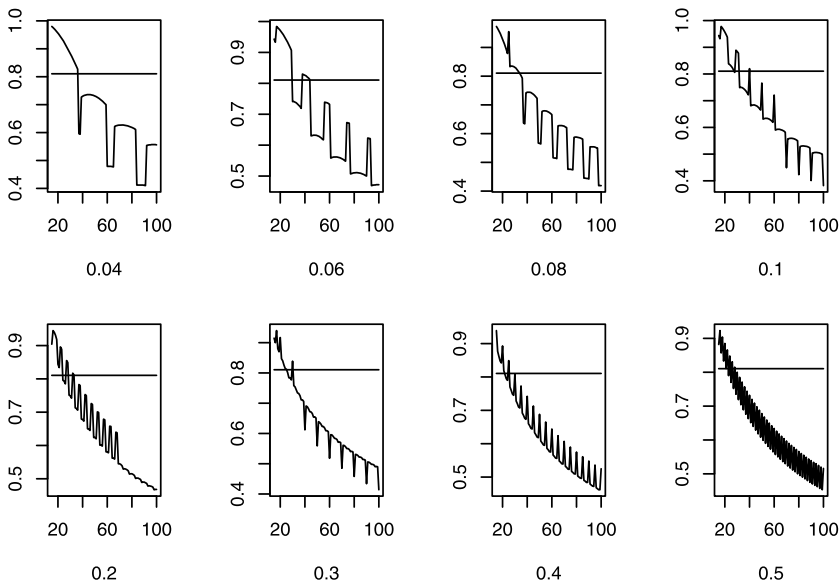


Fig. 4 The probabilities for the classical estimator \hat{p} under the bound in Theorem 1 as the sample size M increases, where $p = 0.04, 0.06, 0.08, 0.1, 0.2, 0.3, 0.4,$ and 0.5 . The horizontal line indicates the position of $\frac{8}{\pi}$

Table 1 The values of M where the quantum estimator \tilde{p} starts to outperform the classical estimator \hat{p} under the bound in Theorem 1

p	0.04	0.06	0.08	0.1	0.2	0.3	0.4	0.5
M	37	45	34	41	29	31	26	29

7 Conclusion

In this article, we have presented the details of the quantum counting algorithm and its advantage over the classical proportional estimation method in terms of their error distributions. However, we need to point out one peculiar fact unique to the quantum counting algorithm. With classical proportional estimation and sampling without replacement, when the sample size M equals the population size N , the estimator $\hat{p} = p$, the true proportion. In contrast, we can not guarantee that the quantum estimator $\tilde{p} = p$ when $M = N$. This is caused by the mismatch between $\omega = \frac{\theta}{\pi} = \frac{1}{\pi} \sin^{-1} \sqrt{\frac{L}{N}}$, which is rarely rational [5], and $\frac{x}{M}$, which is always rational. Further research is needed to come up with an exact quantum counting algorithm.

Acknowledgements The first author was supported by an Ohio University Faculty Research Activity Fund. The second author was supported by National Science Foundation Award DMS 0808993. The third author was partially supported by a grant from the Scientific Research Foundation for the Returned Overseas Chinese Scholars, Ministry of Education of China.

References

1. Boyer, M., Brassard, G., Hoyer, P., Tapp, A.: Tight bounds on quantum searching. *Fortschr. Phys.* **46**, 493–506 (1998)
2. Brassard, G., Hoyer, P., Tapp, A.: Quantum counting. In: *Proc. of the 25th Int. Colloquium on Automata, Languages and Programming*. Lecture Notes in Comp. Sci., vol. 1443, pp. 820–831. Springer, New York (1998)
3. Casella, G., Berger, R.L.: *Statistical Inference*. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove (1990)
4. Diao, Z.: Quantum Computation. *Handbook of Linear Algebra. Discrete Mathematics and Its Applications*, pp. 62–71. Chapman & Hall/CRC Press, Boca Raton (2006)
5. Diao, Z.: Exactness of the original Grover search algorithm. *Phys. Rev. A* **82**, 044301 (2010)
6. Durrett, R.: *Probability: Theory and Examples*. Duxbury Press, Belmont (1996)
7. Grover, L.K.: Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **78**, 325–328 (1997)
8. Mosca, M.: *Quantum computer algorithms*. Ph.D. Thesis, Oxford University Press (1999)
9. Nielsen, M.A., Chuang, I.L.: *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge (2000)