Commun. Theor. Phys. **77** (2025) 065108 (7pp)
https://doi.org/10.1088/1572-9494/ad9730
CSTR: 32041.14.CTP.ad9730

Cryptanalysis and improvement to the quantum private query protocol for enhancing database privacy

Zhengda Shen[1], Wenzhu Shao[1,2,3], Zhigang Li[2,3], Xiaoyu Peng[1], Nankun Mu[1], Mahabubur Rahman Miraj[1] and Bin Liu[1,2,3]

[1] College of Computer Science, Chongqing university, Chongqing 400044, China
[2] Xinjiang Production and Construction Corps Key Laboratory of Computing Intelligence and Network Information Security, Shihezi University, Shihezi, 832003, China
[3] Center for Network and Information, Shihezi University, Shihezi, 832003, China

E-mail: shao@sszu.edu.cn

Received 13 September 2024, revised 22 November 2024
Accepted for publication 26 November 2024
Published 18 March 2025

Abstract

In order to protect the privacy of the query user and database, some QKD-based quantum private query (QPQ) protocols were proposed. One example is the protocol proposed by Zhou *et al*, in which the user makes initial quantum states and derives the key bit by comparing the initial quantum state and the outcome state returned from the database by ctrl or shift mode, instead of announcing two non-orthogonal qubits as others which may leak part secret information. To some extent, the security of the database and the privacy of the user are strengthened. Unfortunately, we find that in this protocol, the dishonest user could be obtained, utilizing unambiguous state discrimination, much more database information than that is analyzed in Zhou *et al*'s original research. To strengthen the database security, we improved the mentioned protocol by modifying the information returned by the database in various ways. The analysis indicates that the security of the improved protocols is greatly enhanced.

Keywords: quantum private query, quantum key distribution, quantum cryptography

(Some figures may appear in colour only in the online journal)

1. Introduction

The advent of quantum technology has brought about a paradigm shift in the way that we approach cryptography and secure communication. As quantum computers threaten the security of the widely used classical encryption schemes, new cryptographic protocols including ones utilizing quantum technology are emerging that promise to maintain privacy even against quantum adversaries. Among these, the quantum private query (QPQ) stands out for its potential to protect both user privacy and database security in database queries.

The need for private query protocols is well-recognized in the field of computer science, as they enable users to retrieve information from a database without revealing their queries or any sensitive information about the database itself. Traditional private query protocols, such as those based on oblivious transfer and homomorphic encryption, provide computational security against classical attackers [1–4]. However, with the advent of quantum computing [5–7], there is a need for secure private query protocols that can provide information-theoretic security against quantum adversaries.

The QPQ addresses this need by leveraging the principles of quantum mechanics, such as superposition and entanglement, to ensure that the user can retrieve information from the database while revealing virtually no information about the query or the database itself. The protocol has gained significant attention due to its potential applications in secure data retrieval, anonymous authentication, and privacy-preserving data mining [8–10].

Several studies have explored the theoretical foundations and practical implementations of QPQ. For instance, Gao *et al* [11] demonstrated the security against a malicious quantum

database; Francesco *et al* [12] proposed an efficient implementation of using linear optics. Additionally, many new QPQ protocols [13–15] were proposed. These works have laid the groundwork for understanding and developing secure private query protocols, which are expected to play a vital role in the future of privacy-preserving information retrieval.

In order to solve the problem of lacking support for a large database, the first QPQ protocol based on Quantum Key Distribution (QKD) was proposed by Jakobi *et al* [16]. From then on, more and more research focuses on QKD-based QPQ and many new protocols are proposed [17–24]. In 2018, Zhou *et al* proposed such a protocol [25], denoted as the ZBLSY protocol, where Alice prepares and sends quantum states to improve the user privacy. Additionally, it also employs the specific two-way communication approach where Bob returns the measured photons by Ctrl or Shift mode instead of declaring two non-orthogonal qubits, which may leak half of the correct qubits. Therefore, the database security is also strengthened. However, we found the advantages that the dishonest user can gain are significantly greater than which was initially analyzed in the ZBLSY protocol.

The rest of this paper is organized as follows. The ZBLSY protocol is briefly reviewed in section 2. In section 3, we describe the attack against the ZBLSY protocol. Section 4 presents the improvements of the ZBLSY protocol and the security is analyzed in section 5. Finally, we make a conclusion according to our analysis in section 6.

2. Brief review of the ZBLSY protocol

In this section, we briefly review the ZBLSY protocol.

Step 1 Alice prepares a lengthy series of photons in one of the four states, $|0\rangle$, $|1\rangle$, $|+\rangle$ or $|-\rangle$, and transmits them to Bob.

Step 2 Bob generates a random string $a \in \{0, 1\}^n$ as the key K^r. In the cases where Bob has successfully received the photons, he measures the i-th photon in Z basis $|0\rangle$, $|1\rangle$ if $a_i = 0$ or X basis $|+\rangle$, $|-\rangle$ if $a_i = 1$. Then, Bob randomly selects one of the following modes to apply to his photons.

In Ctrl mode, Bob returns the measured photon directly to Alice.

In Shift mode, Bob makes the measured photon pass the X gate if the measurement result is $|0\rangle$ or $|1\rangle$, and the Z gate if the measurement result is $|+\rangle$ or $|-\rangle$. Therefore, after the operation, $|0\rangle$ is converted into $|1\rangle$, while $|1\rangle$ is converted into $|0\rangle$; $|+\rangle$ is altered to $|-\rangle$ while $|-\rangle$ is altered to $|+\rangle$. Then, Bob sends back the converted photon to Alice.

Step 3 Alice announces which instances she has successfully detected. For each photon, Bob announces the mode he has selected.

Suppose Bob selects Ctrl mode. Alice measures the detected photon in Z basis if the sent state is $|0\rangle$ or $|1\rangle$, and in X basis if the sent state is $|+\rangle$ or $|-\rangle$. If the measurement result is different from the state of the sent photon, this indicates that the measurement basis used by Alice is different from the one used by Bob. If the measurement result is the same as the state of the sent photon, Alice cannot infer the measurement basis which Bob uses. Therefore, Alice can get the value of K_i^r only when the measurement result is different from the state of the sent photon.

For example, suppose that the state of the photon which Alice sends to Bob is $|1\rangle$ in Step 1. Then, Bob uses Z basis or X basis to measures it in Step 2. If Bob uses Z basis, the measurement result must be $|1\rangle$; if Bob uses X basis, the measurement result is $|+\rangle$ or $|-\rangle$ randomly. When Alice receives the returned photon from Bob, she uses Z basis to measure it. If the measurement result is $|0\rangle$, Alice can infer that Bob uses X basis while Alice cannot obtain any information if the measurement result is $|1\rangle$. Therefore, Alice can get Bob's encoded bit only when the measurement result is $|0\rangle$ with the probability $1/4$. That is, Alice can obtain a quarter of K^r.

When Bob selects Shift mode, the approach is similar to Ctrl mode.

Step 4 Alice randomly selects half of K^r and requires Bob to announce his corresponding measurement results in Step 2. Then, Alice can detect the honesty of Bob based on the outcome states of Bob, her initial states and outcome states. If Alice finds Bob cheating her, she would abort the protocol; otherwise, they will continue to Step 5.

Step 5 At this point, Alice and Bob obtain half of K^r, denoted as K^M. Then, Bob and Alice carry out the classical post-processing step which is similar to Yang's protocol [26]. They transform K^M into

$$\boldsymbol{S} = \begin{bmatrix} q_1 & q_2 & \cdots & q_N \\ q_{N+1} & q_{N+2} & \cdots & q_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{N(k-1)+1} & q_{N(k-1)+2} & \cdots & q_{kN} \end{bmatrix}. \tag{1}$$

Here, q_i is K_i^M and k is a security parameter.

$\boldsymbol{T}$ is a row vector selected randomly by Bob,

$$\boldsymbol{T} = \begin{bmatrix} t_1 & t_2 & \ldots & t_k \end{bmatrix}. \tag{2}$$

Here, t_j is a positive integer, $j = 1, 2, \ldots, k$.

Bob and Alice multiply $\boldsymbol{T}$ by $\boldsymbol{S}$, and obtain the matrix $\boldsymbol{Q}$,

$$\begin{aligned} \boldsymbol{Q} = \boldsymbol{TS} &= \begin{bmatrix} t_1 & t_2 & \ldots & t_k \end{bmatrix} \\ &\times \begin{bmatrix} q_1 & q_2 & \cdots & q_N \\ q_{N+1} & q_{N+2} & \cdots & q_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{N(k-1)+1} & q_{N(k-1)+2} & \cdots & q_{kN} \end{bmatrix} \\ &= \begin{bmatrix} Q_1 & Q_2 & \ldots & Q_N \end{bmatrix}, \end{aligned} \tag{3}$$

where $Q_j = \sum_{i=1}^{k} t_i q_{N(i-1)+j}$.

Then, they perform a piecewise function to obtain the N-bit final key O,

$$O_j = \begin{cases} 0, & Q_j \leqslant t, \\ 1, & Q_j > t, \end{cases} \tag{4}$$

where the threshold should be set to $t = [\sum_{i=1}^{k} t_i/2]$. In the function, the value k and t should be set appropriately so that Alice can obtain about one bit of the final key O.

Step 6 Suppose Alice knows the j-th key bit O_j and tries to retrieve the i-th item X_i. She declares the number $s = j - i$. Then, Bob shifts O by s and encrypts his database with using

the new key O'. Thus, Alice can decrypt X_i correctly with the key bit O_j.

In the ZBLSY protocol, after Alice prepares the initial state, the threat of Bob's spoofing is greatly reduced because Bob could not get the information of initial state accurately.

3. An unambiguous state discrimination attack to the ZBLSY protocol

Although the ZBLSY protocol has the advantage of strong user privacy, unfortunately, we found that the advantages to the user are significantly greater than which is initially analyzed in the ZBLSY protocol. In this section, we present an attack method aimed at the ZBLSY protocol. This method enables Alice to increase the probability of obtaining every key bit beyond the stated probability in the ZBLSY protocol. The detailed analysis is described as follows.

We assume that Alice is dishonest. She prepares a fake entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$, and sends one photon of the entangled state to Bob. Then, Bob uses Z basis $|0\rangle, |1\rangle$ or X basis $|+\rangle, |-\rangle$ to measure it.

Here, the transformation formulas for Bell states are given, which is necessary preparation for the following discussion,

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}} = \frac{|++\rangle + |--\rangle}{\sqrt{2}}, \tag{5}$$

$$\frac{|00\rangle - |11\rangle}{\sqrt{2}} = \frac{|+-\rangle + |-+\rangle}{\sqrt{2}}, \tag{6}$$

$$\frac{|01\rangle + |10\rangle}{\sqrt{2}} = \frac{|++\rangle - |--\rangle}{\sqrt{2}}, \tag{7}$$

$$\frac{|01\rangle - |10\rangle}{\sqrt{2}} = \frac{|+-\rangle - |-+\rangle}{\sqrt{2}}. \tag{8}$$

Next, we will discuss the basis that Bob uses in different scenarios.

Scenario 1 Assume that Bob uses the Z basis. After the measurement, the system state retained by Alice will collapse to $|0\rangle$ if the result of the measurement is $|0\rangle$ and $|1\rangle$ if the result of the measurement is $|1\rangle$. As the probabilities that the result is $|0\rangle$ or $|1\rangle$ are both $1/2$ and Bob will send back the state corresponding to the measurement result to Alice in the subsequent steps, the system state retained by Alice will become either $|00\rangle$ or $|11\rangle$ randomly. Thus, the density operator becomes $\rho_0 = (|00\rangle\langle 00| + |11\rangle\langle 11|)/2$.

Alice chooses Bell bases to measure ρ_0. The result of the measurement is

$$P\left(\text{result} = \frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) = P\left(\text{result} = \frac{|00\rangle - |11\rangle}{\sqrt{2}}\right) = \frac{1}{2}. \tag{9}$$

Scenario 2 Assume that Bob uses the X basis. After the measurement, the system state retained by Alice will collapse to $|+\rangle$ if the result of the measurement is $|+\rangle$ and $|-\rangle$ if the result of the measurement is $|-\rangle$. As the probabilities that the result is $|+\rangle$ or $|-\rangle$ are both $1/2$ and Bob will send back the state corresponding to the measurement result to Alice in the subsequent steps, the system state retained by Alice will become either $|++\rangle$ or $|--\rangle$ randomly. Thus, the density operator is $\rho_1 = (|++\rangle\langle ++| + |--\rangle\langle --|)/2$.

Alice chooses transformed Bell bases to measure the state ρ_1. The result of the measurement is

$$P\left(\text{result} = \frac{|++\rangle + |--\rangle}{\sqrt{2}}\right) = P\left(\text{result} = \frac{|++\rangle - |--\rangle}{\sqrt{2}}\right) = \frac{1}{2}. \tag{10}$$

According to equation (5), by comparing formula 9 and formula 10, we can draw the following conclusions. If the final result is $(|00\rangle - |11\rangle)/\sqrt{2}$, we can infer that Bob uses the Z basis; if the final is $(|++\rangle - |--\rangle)/\sqrt{2}$, we can infer that Bob uses the X basis. During our calculations, we designate the following events: event A occurs when the final outcome is $(|00\rangle - |11\rangle)/\sqrt{2}$; event B happens when the end result is $(|++\rangle - |--\rangle)/\sqrt{2}$; event C transpires when Bob employs the Z basis; and event D takes place when Bob opts for the X basis. Thus, the probability that Alice gets the key bit is

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(AB) \\ &= P(AC) + P(BD) - P(AB) \\ &= P(A|C) \times P(C) + P(B|D) \times P(D) - P(AB) \\ &= \frac{P(C) + P(D)}{2} - 0 \\ &= \frac{1}{2}. \end{aligned} \tag{11}$$

The simulation comparison of attack and non-attack scenarios can be seen in table 1.

Obviously, $1/2$ in the attacked circumstance is greater than $1/4$ in the normal circumstance. Therefore, the dishonest user can get more information by an unambiguous state discrimination attack. Additionally, the analysis results here are different from the results in 3.1.2 EPR Attacks in Zhou's paper [25] because of overlooking the application of unambiguous state discrimination. Table 2 demonstrates the expected number of keys $\overline{n}_1$ and $\overline{n}_2$ that Alice can infer with different data size N and different K in normal and attacked circumstances.

By analysis, the user can obtain 2^K times the amount of database information than which is normally allowed by launching our attack. Therefore, the attack we propose here poses a significant threat to database security.

For the ZBLSY protocol, there may be other attack methods, but the attack method we are using here is quite effective. Therefore, in this paper, we will only analyze the unambiguous state discrimination attack.

Table 1. The simulation comparison of attack and non-attack scenarios.

Attack?	The photon sent by Alice	The measure basis chosen by Bob	The measurement result of Alice	The probability
No	$\|0\rangle$	Z	$\|1\rangle$	$\frac{1}{4}$
Yes	one photon of $\frac{\|00\rangle+\|11\rangle}{\sqrt{2}}$	X	$\|0\rangle$ or $\|1\rangle$	$\frac{1}{2}$
		Z	$\frac{\|00\rangle+\|11\rangle}{\sqrt{2}}$ or $\frac{\|00\rangle-\|11\rangle}{\sqrt{2}}$	
		X	$\frac{\|++\rangle+\|--\rangle}{\sqrt{2}}$ or $\frac{\|++\rangle-\|--\rangle}{\sqrt{2}}$	

Table 2. The expected number of keys $\overline{n}_1$ and $\overline{n}_2$ with different data size N and different K in normal and attacked circumstances.

N	10^3	5×10^3	10^4	5×10^4	10^5	10^6
K	4	5	6	7	7	9
$\overline{n}_1$	3.91	4.88	2.44	3.05	6.10	3.81
$\overline{n}_2$	62.50	156.25	156.25	390.63	781.25	1953.12

4. Improvement to the ZBLSY protocol

As analyzed in the previous section, the security of the database in the ZBLSY protocol is severely threatened. The main reason is that when a dishonest user employs an entanglement attack, the state of the system returned by the database and the state of the system retained by the user can be distinguished with a higher probability by unambiguous state discrimination. The key to improving this is to prevent the user from using similar methods to distinguish between the states. Three improvement methods are given below. The primary enhancements focus on the second and third steps. In addition, the fourth step has been removed in all the improvements below to prevent a recently proposed attack in [27].

4.1. The first improved version of ZBLSY protocol

Step 2 Bob generates a random string $a\in\{0,1\}^n$ as the key K^r. In the cases where Bob has successfully received the photons, he measures the i-th photon in Z basis $|0\rangle, |1\rangle$ if $a_i=0$ or X basis $|+\rangle, |-\rangle$ if $a_i=1$.

Step 3 If the measurement result is $\{|0\rangle, |-\rangle\}$, Bob announces bit 0; if the measurement result is $\{|1\rangle, |+\rangle\}$, he announces bit 1.

Suppose Alice sends $|0\rangle$ to Bob in Step 1. Then, Bob will choose the X basis or Z basis to measure the photon in Step 2. If Bob uses the Z basis, the measurement result is definitely $|0\rangle$ and Bob definitely announces bit 0. If Bob uses the X basis, the measurement result is $|+\rangle$ or $|-\rangle$ randomly, and Bob will announce bit 0 or 1 with the probability $1/2$ respectively. In summary, Bob will announce bit 0 with the probability $3/4$, and bit 1 with the probability $1/4$. Alice can get Bob's encoded bit only when the announced bit is 1. That is, the probability of Alice getting the key bit is $1/4$.

In fact, the main process of this improved version is similar to the protocol in [28]. Nevertheless, the step for the user checking the honesty of the database is omitted in the improved version, which makes the proposed improvement provide better protection for the security of the database.

4.2. The second improved version of the ZBLSY protocol

Step 2 Bob generates a random string $a\in\{0,1\}^n$ as the key K^r. In the cases where Bob has successfully received the photons, he measures the i-th photon in the Z basis $|0\rangle, |1\rangle$ if $a_i=0$ or X basis $|+\rangle, |-\rangle$ if $a_i=1$.

Step 3 For each photon that Bob successfully measures, he announces a pair of two states. One is the measurement result and the other is a random state from the other basis. For example, if the measurement result is $|0\rangle$, Bob would announce $\{|0\rangle, |+\rangle\}$ or $\{|0\rangle, |-\rangle\}$. Alice could infer the basis Bob has used based on his announcement. The probability of success is $1/4$.

Suppose Alice sends $|0\rangle$ to Bob in Step 1. Then, Bob will choose the X basis or Z basis to measure the photon in Step 2. If Bob uses the Z basis, the measurement result is definitely $|0\rangle$ and Bob announces $\{|0\rangle, |+\rangle\}$ or $\{|0\rangle, |-\rangle\}$ with the probability $1/2$ respectively. If Bob uses the X basis, the measurement result is $|+\rangle$ or $|-\rangle$ randomly, and Bob will announce $\{|0\rangle, |+\rangle\}$, $\{|1\rangle, |+\rangle\}$, $\{|0\rangle, |-\rangle\}$ or $\{|1\rangle, |-\rangle\}$ with the probability $1/4$ respectively. In summary, Bob will announce $\{|0\rangle, |+\rangle\}$ with the probability $3/8$, $\{|0\rangle, |-\rangle\}$ with the probability $3/8$, $\{|1\rangle, |+\rangle\}$ with the probability $1/8$ and $\{|1\rangle, |-\rangle\}$ with the probability $1/8$. Alice can get Bob's encoded bit only when the announcement is $\{|1\rangle, |+\rangle\}$ or $\{|1\rangle, |-\rangle\}$. Therefore, the probability that Alice can get the key bit is $1/4$.

The improved version is roughly the inverse process of the J protocol [16]. In this improved version, Alice infers which basis Bob used by sending photons and receiving Bob's announcements, thereby obtaining the key bits. While in the J protocol [16], Alice measures the photons sent by Bob and receives Bob's announcements to infer which photon Bob sent, thus obtaining the key bits.

4.3. The third improved version of the ZBLSY protocol

Step 2 Bob generates a random string $a\in\{0,1\}^n$ as the key K^r. For the instances that Bob has successfully received, he measures the i-th photon in Z basis $|0\rangle, |1\rangle$ if $a_i=0$ or X basis $|+\rangle, |-\rangle$ if $a_i=1$. If the measurement result is $\{|0\rangle, |-\rangle\}(\{|1\rangle, |+\rangle\})$, he sends back $|0\rangle(|+\rangle)$.

Step 3 Alice measures the photon Bob returns in the Z basis if she sends $|0\rangle$ or $|-\rangle$ and in the X basis if she sends $|1\rangle$ or $|+\rangle$. Alice could infer which basis Bob uses in Step 2 with success probability $1/8$.

Suppose Alice sends $|1\rangle$ to Bob in Step 1. Then Bob will choose the X basis or Z basis to measure the photon in Step 2. If Bob uses the Z basis, the measurement result is definitely $|1\rangle$ and Bob will send back $|+\rangle$. If Bob uses the X basis, the measurement result is $|+\rangle$ or $|-\rangle$ randomly, and Bob will send back $|+\rangle$ or $|0\rangle$ with the probability $1/2$ respectively. Thus, Bob will send back $|+\rangle$ with the probability $3/4$, and $|0\rangle$ with the probability $1/4$. After Alice receives the photon, she will measure it in the X basis. Alice can get Bob's encoded bit only when the measurement result is $|-\rangle$ with the probability $1/8$.

5. Security analysis to the improved versions

After analysis, we found that the third improved version can enhance the security of the database compared to the original protocol, but it is not as effective as the other improved versions. Therefore, next we will mainly analyze and demonstrate the security of the first and second improved versions. Due to the vulnerability in the original protocol being related to the database security, we focused on analyzing the performance of the improved protocols in terms of the database security.

5.1. Security analysis to the first improved version

As mentioned before, the main process of the first improved version is similar to the protocol in [28]. In fact, the security analysis in [28] only focuses on some specific attacks. And here we will make generalized assumptions about the attack patterns, and prove the security of the improved protocol based on these assumptions. Since Bob performs a single-qubit measurement, the state Alice initially prepared can be represented as

$$|\psi\rangle = |\phi_0\rangle_a|0\rangle_b + |\phi_1\rangle_a|1\rangle_b, \tag{12}$$

where the system a is reserved by Alice and the system b is sent to Bob. Here $|\phi_0\rangle$ and $|\phi_1\rangle$ are non-normalized and satisfy

$$\langle\phi_0|\phi_0\rangle + \langle\phi_1|\phi_1\rangle = 1. \tag{13}$$

We assume that $\langle\phi_0|\phi_0\rangle = k_0$, $\langle\phi_1|\phi_1\rangle = k_1$, $\langle\phi_0|\phi_1\rangle = \lambda$ and $\langle\phi_1|\phi_0\rangle = \lambda^*$.

When Bob receives the system b, he measures it with the Z basis or X basis. When Bob chooses the Z basis, the state of system a will become $|\phi_0\rangle/\sqrt{k_0}$ if the result of measurement is $|0\rangle$ and the state of system a will become $|\phi_1\rangle/\sqrt{k_1}$ if the result of measurement is $|1\rangle$. Due to

$$|\psi\rangle = \frac{|\phi_0\rangle + |\phi_1\rangle}{\sqrt{2}}|+\rangle + \frac{|\phi_0\rangle - |\phi_1\rangle}{\sqrt{2}}|-\rangle, \tag{14}$$

when Bob chooses the X basis, the state of system a will become $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$ if the result of measurement is $|+\rangle$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$ if the result of statement is $|-\rangle$. The probabilities of various results are as below

$$P(|0\rangle) = \frac{k_0}{2}, \tag{15}$$

$$P(|1\rangle) = \frac{k_1}{2}, \tag{16}$$

$$P(|+\rangle) = \frac{1+\lambda+\lambda^*}{4}, \tag{17}$$

$$P(|-\rangle) = \frac{1-\lambda-\lambda^*}{4}. \tag{18}$$

Then, if Bob announces 0, this indicates that Bob's measurement result is either $|0\rangle$ or $|-\rangle$, then the state of system a is $|\phi_0\rangle/\sqrt{k_0}$ or $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$. Therefore, Alice can get Bob's encoded bit by distinguishing between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$. If Bob announces 1, Alice needs to distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$.

When Bob announces 0, the prior probability of $|\phi_0\rangle/\sqrt{k_0}$ is $P_1 = 2k_0/(2k_0+1-\lambda-\lambda^*)$ and the prior probability of $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$ is $P_2 = (1-\lambda-\lambda^*)/(2k_0+1-\lambda-\lambda^*)$. Thus, the process of Alice determining Bob's encoded bit is transformed into a problem of distinguishing between two states with prior probabilities. The highest probability that Alice can distinguish between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$ is

$$\begin{aligned} P_3 &= 1 - 2\sqrt{P_1P_2}\left|\frac{\langle\phi_0|}{\sqrt{k_0}}\frac{|\phi_0\rangle - |\phi_1\rangle}{\sqrt{1-\lambda-\lambda^*}}\right| \\ &= 1 - \frac{2\sqrt{2}}{2k_0+1-\lambda-\lambda^*}|\langle\phi_0|\phi_0\rangle - \langle\phi_1|\phi_1\rangle| \\ &= 1 - \frac{2\sqrt{2}}{2k_0+1-\lambda-\lambda^*}|k_0-\lambda|. \end{aligned} \tag{19}$$

By the same token, when Bob announces 1, the highest probability that Alice can distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$ is

$$P_4 = 1 - \frac{2\sqrt{2}}{2k_1+1+\lambda+\lambda^*}|k_1+\lambda^*|. \tag{20}$$

Thus, combining the probabilities of the two cases mentioned above, we can infer that the probability that Alice can infer one of the correct bits of the key is

$$\begin{aligned} P_5 &= (P(|0\rangle) + P(|-\rangle)) \\ &\quad \times P_3 + (P(|1\rangle) + P(|+\rangle)) \times P_4 \\ &= \left(\frac{k_0}{2} + \frac{1-\lambda-\lambda^*}{4}\right) \times \left(1 - \frac{2\sqrt{2}}{2k_0+1-\lambda-\lambda^*}|k_0-\lambda|\right) \\ &\quad + \left(\frac{k_1}{2} + \frac{1+\lambda+\lambda^*}{4}\right) \times \left(1 - \frac{2\sqrt{2}}{2k_1+1+\lambda+\lambda^*}|k_1+\lambda^*|\right) \\ &= 1 - \frac{1}{\sqrt{2}}(|k_0-\lambda| + |k_1+\lambda^*|) \\ &\leqslant 1 - \frac{1}{\sqrt{2}}|k_0-\lambda+k_1+\lambda^*| \\ &= 1 - \frac{1}{\sqrt{2}} \approx 0.29. \end{aligned} \tag{21}$$

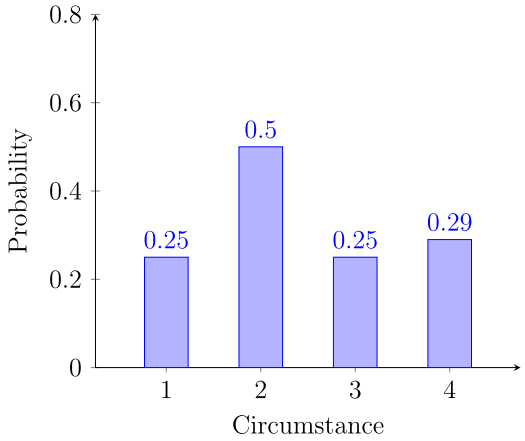

Figure 1. The probabilities that Alice can infer one of the correct bits of the key under different circumstances.

That is, Alice can infer one of the bits of the key with a probability of less than 0.29, which is similar to that in the J protocol [16]. Obviously, the improved protocol has a higher database security than the original protocol. The probabilities of Alice inferring one of the correct bits of the key under various circumstances are depicted in figure 1.

Circumstance 1 represents the normal operating condition in the original protocol; circumstance 2 depicts the scenario under attack in the original protocol; circumstance 3 represents the normal operating condition in the improved protocol; and circumstance 4 depicts the scenario under attack in the improved protocol.

The analysis of user privacy is the same as the original protocol, so it will not be repeated here.

5.2. Security analysis to the second improved version

The preliminary preparations and the probabilities of Bob's different measurement results are the same as in the security analysis to the first improved version, so they will not be repeated here.

If Bob announces$\{|0\rangle, |+\rangle\}$, this indicates that Bob's measurement result is either $|0\rangle$ or $|+\rangle$, then the state of system a is $|\phi_0\rangle/\sqrt{k_0}$ or $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$. Therefore, Alice can get Bob's encoded bit by distinguishing between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$. If Bob announces $\{|0\rangle, |-\rangle\}$, Alice needs to distinguish between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$. If Bob announces $\{|1\rangle, |+\rangle\}$, Alice needs to distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$. If Bob announces $\{|1\rangle, |-\rangle\}$, Alice needs to distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$.

Based on the security analysis above, we can know that the highest probability that Alice can distinguish between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$ is P_3 and the highest probability that Alice can distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$ is P_4.

When Bob announces $\{|0\rangle, |+\rangle\}$, the prior probability of $|\phi_0\rangle/\sqrt{k_0}$ is

$$P_6 = \frac{2k_0}{(2k_0+1+\lambda+\lambda^*)}, \tag{22}$$

and the prior probability of $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$ is

$$P_7 = \frac{(1+\lambda+\lambda^*)}{(2k_0+1+\lambda+\lambda^*)}. \tag{23}$$

Therefore, the highest probability that Alice can distinguish between $|\phi_0\rangle/\sqrt{k_0}$ and $(|\phi_0\rangle + |\phi_1\rangle)/\sqrt{1+\lambda+\lambda^*}$ is

$$\begin{aligned} P_8 &= 1 - 2\sqrt{P_6P_7}\left|\frac{\langle\phi_0|}{\sqrt{k_0}}\frac{|\phi_0\rangle+|\phi_1\rangle}{\sqrt{1+\lambda+\lambda^*}}\right| \\ &= 1 - \frac{2\sqrt{2}}{2k_0+1+\lambda+\lambda^*}|\langle\phi_0|\phi_0\rangle + \langle\phi_1|\phi_1\rangle| \\ &= 1 - \frac{2\sqrt{2}}{2k_0+1+\lambda+\lambda^*}|k_0+\lambda|. \end{aligned} \tag{24}$$

By the same token, when Bob announces $\{|1\rangle, |-\rangle\}$, the highest probability that Alice can distinguish between $|\phi_1\rangle/\sqrt{k_1}$ and $(|\phi_0\rangle - |\phi_1\rangle)/\sqrt{1-\lambda-\lambda^*}$ is

$$P_9 = 1 - \frac{2\sqrt{2}}{2k_1+1-\lambda-\lambda^*}|k_1-\lambda^*|. \tag{25}$$

Thus, combining the probabilities of the four cases mentioned above, we can infer that the probability that Alice can infer one of the correct bits of the key is

$$\begin{aligned} P_{10} &= \tfrac{1}{2}[(P(|0\rangle) + P(|-\rangle)) \\ &\quad \times P_3 + (P(|1\rangle) + P(|+\rangle)) \times P_4 \\ &\quad + (P(|0\rangle) + P(|+\rangle)) \\ &\quad \times P_8 + (P(|1\rangle) + P(|-\rangle)) \times P_9] \\ &= \tfrac{1}{2}\Bigg[\left(\tfrac{k_0}{2} + \tfrac{1-\lambda-\lambda^*}{4}\right) \times \left(1 - \tfrac{2\sqrt{2}}{2k_0+1-\lambda-\lambda^*}|k_0-\lambda|\right) \\ &\quad + \left(\tfrac{k_1}{2} + \tfrac{1+\lambda+\lambda^*}{4}\right) \times \left(1 - \tfrac{2\sqrt{2}}{2k_1+1+\lambda+\lambda^*}|k_1+\lambda^*|\right) \\ &\quad + \left(\tfrac{k_0}{2} + \tfrac{1+\lambda+\lambda^*}{4}\right) \times \left(1 - \tfrac{2\sqrt{2}}{2k_1+1+\lambda+\lambda^*}|k_0+\lambda|\right) \\ &\quad + \left(\tfrac{k_1}{2} + \tfrac{1-\lambda-\lambda^*}{4}\right) \times \left(1 - \tfrac{2\sqrt{2}}{2k_1+1-\lambda-\lambda^*}|k_1-\lambda^*|\right)\Bigg] \\ &= 1 - \tfrac{1}{2\sqrt{2}}(|k_0-\lambda| + |k_1+\lambda^*| + |k_0+\lambda| + |k_1-\lambda^*|) \\ &\leqslant 1 - \tfrac{1}{2\sqrt{2}}|k_0-\lambda+k_1+\lambda^*+k_0+\lambda+k_1-\lambda^*| \\ &= 1 - \tfrac{1}{\sqrt{2}} \approx 0.29\,, \end{aligned} \tag{26}$$

which is the same as P_5.

Thus, the second improved version has the same level of the database security as the first improved version.

6. Conclusion

In this paper, we presented an unambiguous state discrimination attack method to the ZBLSY protocol, utilizing which Alice could infer one of the correct bits of the key with

successful probability $1/2$ in the ZBLSY protocol. It suggests that the advantage of the user is greater in the ZBLSY protocol. To solve this problem, we proposed three improved protocols. Based on our analysis, the probability that the dishonest user could obtain the key bit decreased to approximately 0.29 in the first and second improved versions. This results in stronger database security, making the protocol more robust against potential attacks.

The improved protocols' resistance to internal attacks from the database and its ability to ensure user privacy even in the presence of entanglement attacks demonstrate its practicality and potential for real-world applications. It offers a promising solution for secure data retrieval and privacy-preserving information access, particularly in scenarios where classical encryption schemes are susceptible to quantum attacks.

We have analyzed the security of the ZBLSY protocol under ideal conditions here; in practice, it is also necessary to consider the security vulnerabilities introduced by the imperfections of the devices [29]. The relevant research can also refer to the literatures [30–32].

Acknowledgments

This work was supported by the National Key R&D Program of China (Grant No. 2022YFC3801700), the National Natural Science Foundation of China (Grant No. 62472052), and Xinjiang Production and Construction Corps Key Laboratory of Computing Intelligence and Network Information Security (Grant No. CZ002702-3).

References

[1] Wang P, Zhang R, Jiang G H and Sun Z W 2022 Computationally secure quantum oblivious transfer *Adv. Quantum. Technol.* **5** 2100125
[2] Liu M M, Hu Y P, Juliane K and Johannes B 2017 Quantum security analysis of a lattice-based oblivious transfer protocol *Front. Inform. Technol. Electron. Eng.* **18** 1348–69
[3] Liu M M and Hu Y P 2019 Universally composable oblivious transfer from ideal lattice *Front. Comput. Sci.* **13** 879–906
[4] Peng Z N *et al* 2023 On the security of fully homomorphic encryption for data privacy in Internet of Things *Concurr. Comput. Pract. Exp.* **35** e7330
[5] Grover L K 1996 A fast quantum mechanical algorithm for database search *STOC'96: Proceedings of the 28th Annual ACM Symposium on Theory of Computing* 212–9
[6] Song Y *et al* 2024 A quantum federated learning framework for classical clients *Sci. China. Phys. Mech.* **67** 250–311
[7] Shor P W 1994 Algorithms for quantum computation: discrete logarithms and factoring *Proceedings 35th Annual Symposium on Foundations of Computer Science* 124–34
[8] Xiao M and Lei S M 2021 Quantum private query with authentication *Quantum. Inf. Process.* **20** 1–13
[9] Yu F *et al* 2020 Security improvements of several basic quantum private query protocols with O(log N) communication complexity *Quantum. Inf. Process.* **807** 330–40
[10] Ravi A T, Chitra S *et al* 2015 Privacy preserving data mining *Rese. J. Appl. Sci. Eng. Tech.* **9** 616–21
[11] Gao F, Qin S J, Huang W and Wen Q Y 2019 Quantum private query: a new kind of practical quantum cryptographic protocol *Sci. China. Phys. Mech.* **62** 10–21
[12] De Martini F *et al* 2009 Experimental quantum private queries with linear optics *Phys. Rev.* A **80** 010302
[13] Xiao M and Zhao M J 2024 Multi-user quantum private query using Bell states *Quantum. Inf. Process.* **23** 81
[14] Yang Y G *et al* 2024 Error-tolerant measurement-device-independent quantum private queries of blocks *Int. J. Theor. Phys.* **63** 180
[15] Zhang X *et al* 2023 Security loophole and improvement of quantum private query protocol based on W state *Int. J. Theor. Phys.* **62** 162
[16] Jakobi M *et al* 2011 Practical private database queries based on a quantum-key-distribution protocol *Phys. Rev.* A **83** 022103
[17] Xiao M and Zhang D F 2019 Practical quantum private query with classical participants *Chin. Phys. Lett.* **36** 9–13
[18] Liu B, Gao F, Huang W and We Q Y 2015 QKD-based quantum private query without a failure probability *Sci. China. Phys. Mech.* **58** 12–7
[19] Yang Y G *et al* 2017 Robust QKD-based private database queries based on alternative sequences of single-qubit measurements *Sci. China. Phys. Mech.* **60** 1–11
[20] Liu L, Guo F Z and Wen Q Y 2021 Practical decoy-state quantum private queries against joint-measurement attack under weak coherent pulse sources *Quantum. Inf. Process.* **20** 1–17
[21] Cheng Y *et al* 2018 Quantum private query protocol based on EPR pairs *Chin. J. Elec.* **27** 256–62
[22] Cheng Y *et al* 2019 Practical two-way QKD-based quantum private query with better performance in user privacy *Int. J. Theor. Phys.* **58** 2069–80
[23] Wei C Y, Gao F, Wen Q Y and Wang T Y 2015 Practical quantum private query of blocks based on unbalanced-state Bennett–Brassard-1984 quantum-key-distribution protocol *Sci. Rep.* **4** 7537
[24] Wei C Y *et al* 2020 Error tolerance bound in QKD-based quantum private query *IEEE J. Sel. Area. Comm.* **38** 517–27
[25] Zhou Y H *et al* 2018 A quantum private query protocol for enhancing both user and database privacy *Commun. Theor. Phys.* **69** 31–6
[26] Yang Y G *et al* 2016 Quantum private query with perfect user privacy against a joint-measurement attack *Phys. Lett.* A **380** 4033–8
[27] Wei C Y, Cai X Q and Wang T Y 2024 Reexamination of the realtime protection for user privacy in practical quantum private query arXiv:2407.19147
[28] Yu F, Qiu D W, Situ H Z, Wang X M and Long S 2015 Enhancing user privacy in SARG04-based private database query protocols *Quantum. Inf. Process.* **14** 4201–10
[29] Qin L Z, Liu B, Gao F, Xu B J and Li Y 2024 Decoy-state quantum private query protocol with two-way communication *Physica* A **633** 070301
[30] Liu B *et al* 2022 Decoy-state method for quantum-key-distribution-based quantum private query *Sci. China. Phys. Mech.* **65** 240312
[31] Jiao Y F *et al* 2024 Analysis and protection to user privacy in quantum private query with non-ideal light source *Quantum. Inf. Process.* **23** 133
[32] Liu B *et al* 2024 Decoy-state quantum-key-distribution-based quantum private query with error tolerance bound *Phys. Rev.* A **109** 052442