Quantum Secure Privacy Preserving Technique to Obtain the Intersection of Two Datasets for Contact Tracing

Sumit Kumar Debnath[1], Vikas Srivastava[2], Tapaswini Mohanty[3], Nibedita Kundu[4,*] and Kouichi Sakurai[5]

[1,2,3]Department of Mathematics, National Institute of Technology Jamshedpur, Jamshedpur-831014, India
[4]Department of Mathematics, The LNM Institute of Information Technology, Jaipur-302031, India
[5]Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka- 8190395, Japan

Abstract

Contact tracing has emerged as a powerful and effective measure to curb the spread of contagious diseases. It is a robust tool, but on the downside, it possesses a risk of privacy violations as contact tracing requires gathering a lot of personal information. So there is a need for a cryptographic primitive that obfuscate the personal data of the user. Taking everything into account, private set intersection seems to be the natural choice to address the problem. Nearly all of the existing PSI protocols are relying on the number theoretic assumption based hard problems. However, these problems are not secure in quantum domain. As a consequence, it becomes essential to designing PSI that can resist quantum attack and provide long-term security. One may apply quantum cryptography to develop such PSI protocol. This paper deals with the design of PSI using quantum cryptography (QC), where the security depends on the principles of basic quantum mechanics. Our scheme achieves long-term security and remains secure against quantum attacks due to the use of QC. As opposed to the existing quantum PSI protocols, the communication and computation costs of our scheme are independent of the size of universal set. In particular, the proposed protocol achieves optimal communication and computation costs in the domain of quantum PSI. Moreover, we require only single photon quantum resources and simple single-particle projective measurements, unlike most of the existing quantum PSI protocols.

Keywords: Contact Tracing; Private Set Intersection; Quantum Communication; Quantum Computation; Long-Term Security

1 Introduction

The Covid-19 pandemic has brought to the table an unprecedented task for researchers worldwide to find solutions to curb the virus's spread. Contact tracing has emerged as an important mitigation tool for health authorities to fight the outbreaks of highly contagious diseases like Covid-19, Ebola, and SARS. Contact tracing is a powerful countermeasure that can be utilized to control the spread of infection.

*Corresponding author. Email address: nknkundu@gmail.com

In contact tracing, a centralized credible authority identifies and assesses people exposed to a contagious disease to put a break over the further transmission of the disease in the community.

It requires recording and measuring physical contiguity among individuals to ensure that when one person reports infection, it is possible to mark all other persons who have been close to that person during an appropriate period in the past and warn them about the potential exposure.

On the one hand, it helps to alert the users who have been in close contact with the infected persons, and in the meantime, it also allows health authorities to take proper actions like isolating the contacts and doing early treatment to break the chains of transmission.

When done systematically and ethically, contact tracing transpires as a robust tool in the fight against pandemic and contagious diseases. However, it has its own downsides as contact tracing requires gathering personal information of citizens, such as geographical location or contact numbers which raises ethical issues and serious privacy violations. It has motivated the cryptographic community to come up with novel ideas for private contact tracing.

If contact tracing is employed without proper deliberation and careful examination, it can turn out to be a surveillance tool that can damage and compromise the user's privacy.

The goal should be that computations over private data is kept as low as possible. It also needs to be ensured that a corrupt client cannot learn more private information than what is intended to be revealed.

This gives a natural impulse for a cryptographic primitive that would allow the obfuscation of private information of individuals. Cryptography techniques like *private set intersections (PSI)* can help alleviate privacy concerns. PSI is a cryptographic technique that allows parties to exchange datasets (contact numbers, geographical locations etc.) without revealing the actual data to each other.

Involved parties can compute the intersection securely without any concern of privacy violations. Private Set Intersection(PSI) will allow a person to determine if the data they gathered matches the dataset of diagnosed patients without revealing their private information to the server. Our protocol enables the user to check if their contact history dataset matches the dataset of diagnosed patients without disclosing their personal information to the server. In addition to that client would then only learn the intersection, nothing more than that. As we can see, PSI provides a powerful solution when it comes to private contact tracing and address the issue of privacy violation and data leakage efficiently.

In the domain of classical cryptography, several constructions of PSI have been proposed [1–3, 7, 10, 13, 15–18, 20, 25–29]. Most of them are depending on the number theoretic assumption based hard problems. However, due to Shor's algorithm [24], these hard problems do not remain secure in quantum domain.

Moreover, the issue of long-term security is not resolved by the existing classical cryptosystem. As a consequence, existing classical cryptosystem becomes a threat not only for present but also for future in several real life applications. For instance, sensitive information (such as electronic health records, government documents) should be stored securely for a long period. However, there is a possibility of loosing long-term security during the transmission of classical encrypted version of those data through public channel. This is due to the fact that during the transmission, adversary may save encrypted data and it will wait for the development of efficient classical algorithms or quantum computers. Once efficient classical algorithms

or quantum computers come into market, the adversary can decrypt the data which were saved in encrypted form during transmission. In other words, confidentiality of sensitive information may have a very limited lifespan. Thereby, suitable replacement is required to overcome these security threats.

One may think post-quantum cryptography (PQC) as an appropriate candidate since it provides quantum security with respect to the existing classical or quantum algorithms. However, it does not provide long-term security as in future it may be possible to develop some quantum or classical algorithms for breaking the hard problems of PQC.

On the contrary, by the laws of quantum physics, quantum cryptography (QC) provides long-term security against an eavesdropper with unlimited computational power and remains secure against quantum computer. Hence, it is essential to employ QC in the construction of privacy preserving protocols, particularly in PSI protocols.

Shi et al. [22] developed first quantum PSI protocol. In the following, Cheng et al. [5] showed that client's query can be manipulated by a dishonest server in the work of [22] and hence fairness is not preserved in [22]. Although a fully trusted passive third party (TP) is involved in [22] to achieve fairness, still the existence of fully trusted TP is impossible in real life. Later, set member decision protocol of [23] was extended by Maitra [19] to develop a quantum PSI. Recently, [6] designed a quantum PSI by using the asymmetric key distribution of [9].

Our Contribution: In this work, we concentrate ourselves in the design and analysis of unconditionally secure two-party quantum private set intersection protocol QuPSI. We then investigate the application of the proposed QuPSI in the context of contact tracing. In order to design the proposed scheme, we have used the asymmetric key distribution of [9] and Bloom filter [4] as the building blocks. The design of QuPSI protocol allows two entities, each holding a dataset, to evaluate the intersection of datasets privately without disclosing the actual data. Our scheme mitigate privacy concerns and provides a robust solution. Security of QuPSI relies on the basic principles of quantum mechanics. As a consequence, it remains secure against well-known quantum attacks and achieves long-term security, in contrast to the classical PSI protocols. In this paper, we put forward a cryptographic building block QuPSI for privacy-preserving contact tracing. Our scheme, QuPSI, by design, ensures that both entities' private information- healthy individuals and diagnosed patients remain secure.

In contrast to [5, 6, 19, 22], the communication cost and computation cost of our scheme does not depend on universal set's size. Particularly, our proposed protocol attains optimal communication and computation cost in the context of quantum PSI. The works [5, 19, 22] need measurement in higher dimensional Hilbert space, "multi-particle entangled states" as quantum resources, and "complicated oracle operators". However, with the existing technologies implementation of these oracle operators and preparation of these resources are not feasible. As a consequence, only theoretical approaches have been provided by [5, 19, 22] in the context of quantum PSI. On the other hand, similar to [6], our proposed quantum PSI is feasible and practical since simple single-particle projective measurements and single photons are used in it. Moreover, unlike [5, 19, 22], multiple execution of set intersection functionality is possible in our scheme by performing only one time quantum computation and quantum communication, provided the parameter m remains same.

2 Bloom Filter [4]

Bloom filter is a data structure representing a set $Z = \{z_1, ..., z_d\}$ of d elements by an array of size n and uses λ independent hash functions $\mathsf{Hash} = (H_1, \ldots, H_\lambda))$ with $H_i : \{0,1\}^* \rightarrow \{1, \ldots, n\}$ for inserting or checking the membership of elements into that array. We represent the Bloom filter for Z by $\mathsf{Bloom}_Z \in \{0,1\}^n$ and i-th entry in it by $\mathsf{Bloom}_Z[i]$. The following three operations are performed in a Bloom filter:

– *Initialization*: Put 0 at all the entries of an array of size n, which we call as empty Bloom filter.

– *Add*: To add an element $z \in Z$ into Bloom filter, compute $H_1(z), \ldots, H_\lambda(z)$ and put 1 to the positions $\{H_1(z), \ldots, H_\lambda(z)\}$ of the Bloom filter. Repeat the process for each $z \in Z$ to obtain Bloom_Z.

– *Check*: To check the presence of an element $\overline{z}$ in Z, compute $H_1(\overline{z}), \ldots, H_\lambda(\overline{z})$. If atleast one of $\mathsf{Bloom}_Z[H_1(\overline{z})], \ldots, \mathsf{Bloom}_Z[H_\lambda(\overline{z})]$ is 0 then $\overline{z} \notin Z$; else, $\overline{z}$ provably belongs to Z.

Bloom filter attains *false positive* since an element that has not been added in the filter can falsely pass the check test. While, it never allows false negative since an element that has been added in the filter will always pass the check test.

3 Proposed Quantum PSI

In this section, we describe the construction of the proposed quantum PSI between a client C and server S. The asymmetric key distribution of [9] and Bloom filter [4] have been utilized as the building blocks of our construction.

A high level overview: We use the asymmetric key distribution of [9] for distributing only few bits of the key to a party, the whole key to another party. Let us consider that the client C with the private set $X = \{x_1, \ldots, x_v\} \subset \{0,1\}^*$ and the server S with the private set $Y = \{y_1, \ldots, y_w\} \subset \{0,1\}^*$ involve into the QuPSI protocol, which involves three algorithms: (i) $\mathsf{QuPSI.Raw\ Key}$, (ii) $\mathsf{QuPSI.Asymmetric\ Key}$, (iii) $\mathsf{QuPSI.Set\ Intersection}$. Also, let $(n, \mathsf{Hash} = (H_1, \ldots, H_\lambda))$ be Bloom filter parameters with $H_i : \{0,1\}^* \rightarrow \{1, \ldots, n\}$ and $\{p_1, \ldots, p_m\}$ be the position set with non-zero entries in Bloom_X. Initially, the algorithm $\mathsf{QuPSI.Raw\ Key}$ enables S to learn an n-bit raw key $\mathsf{KE}_1 = \{k_1, \ldots, k_n\}$ and C to learn only m bits $T_1 = \{t_1, \ldots, t_m\} = \{k_{u_1}, \ldots, k_{u_m}\}$ of KE_1 such that S does not have any knowledge about the position u_i of t_i in KE_1. At later stage, $\mathsf{QuPSI.Asymmetric\ Key}$ is run for allowing S to generate $\mathsf{KE}_2 = \{\overline{k}_1, \ldots, \overline{k}_n\}$ from KE_1 and C to learn $T_2 = \{\overline{k}_{p_1}, \ldots, \overline{k}_{p_m}\}$ by using a permutation π over the set $\{1, \ldots, n\}$ such that the set $\{u_1, \ldots, u_m\}$ is mapped to the set $\{p_1, \ldots, p_m\}$. In the following, during $\mathsf{QuPSI.Set\ Intersection}$, S sends $B = \mathsf{KE}_2 \oplus \mathsf{Bloom}_Y = \{\overline{k}_1 \oplus \mathsf{Bloom}_Y[1], \ldots, \overline{k}_n \oplus \mathsf{Bloom}_Y[n]\} = \{b_1, \ldots, b_n\}$ to C. On receiving B, C evaluates $A = \{\overline{k}_{p_1} \oplus b_{p_1}, \ldots, \overline{k}_{p_m} \oplus b_{p_m}\} = \{a_1, \ldots, a_m\}$. It allows C to learn the entries of $\{a_1, \ldots, a_m\} = \{\mathsf{Bloom}_Y[p_1], \ldots, \mathsf{Bloom}_Y[p_m]\}$ of Bloom_Y. Then C forms a resulting Bloom filter Bloom of length n by inserting 1 at all such p_i's where a_i's are 1 and 0's at the remaining entries. In the following, C outputs a set χ as $X \cap Y$, where χ is the collection of all such elements of X for which the set membership test of Bloom is satisfied. The proposed QuPSI is discussed below in detail.

Protocol 1. QuPSI

QuPSI.Raw Key*:*

1. *S and C run the asymmetric key distribution of [9] in order to share an $n+\kappa$-bit key $\mathsf{KE} = \{r_1, \ldots, r_{n+\kappa}\}$ such that S receives the whole KE and C learns only $m+\kappa$ bits T of KE. In order to do that, on needs to set $\alpha = \sin^{-1}\left(\sqrt{\frac{2(m+\kappa)}{n+\kappa}}\right)$ in [9]. On the other hand, the choice $\kappa = (n-4m)/3$ in [14] enables the same kind of key distribution.*
2. *C chooses κ bits randomly from T and requests S to reveal the corresponding bits from KE. Then C compares its own selected part with S's corresponding part. If C's κ bits are equal to S's κ bits then they move to the next step; otherwise, C aborts the execution by stating that S is dishonest.*
3. *S deletes the compared κ bits from KE and updates the positions of the rest n bits in KE to obtain the updated key $\mathsf{KE}_1 = \{k_1, \ldots, k_n\}$. For instance, $\kappa = 2$, and bits of 2nd and 5th positions i.e., r_2 and r_5 are deleted. Then $k_1 = r_1$, $k_2 = r_3$, $k_3 = r_4$, $k_i = r_{i+2}$ for $i = 4, \ldots, n$.*
4. *C deletes the compared κ bits from T and updates the positions of the rest m bits in T for obtaining the updated key $T_1 = \{t_1, \ldots, t_m\} = \{k_{u_1}, \ldots, k_{u_m}\}$, where u_i is the position of $t_i = k_{u_i}$ in KE_1. For instance, if $m = 5$ and $T = \{r_2, r_3, r_5, r_9, r_{12}, r_{15}, r_{20}\}$ then $t_1 = r_3 = k_2$, $t_2 = r_9 = k_7$,$t_3 = r_{12} = k_{10}$, $t_4 = r_{15} = k_{13}$ and $t_5 = r_{20} = k_{18}$, i.e., $u_1 = 2$, $u_2 = 7$, $u_3 = 10$, $u_4 = 13$, $u_5 = 18$.*

QuPSI.Asymmetric Key*:*

1. *Let $\{p_1, \ldots, p_m\}$ be the positions with non-zero entries in Bloom_X. A random permutation π over the set $\{1, \ldots, n\}$ is selected by C such that the set $\{u_1, \ldots, u_m\}$ is mapped to the set $\{p_1, \ldots, p_m\}$. C sends π to S and obtains the updated key $T_2 = \{\overline{k}_{p_1}, \ldots, \overline{k}_{p_m}\}$ from $T_1 = \{k_{u_1}, \ldots, k_{u_m}\}$ by applying π over the position set $\{u_1, \ldots, u_m\}$. Note that $\overline{k}_{p_i} = \overline{k}_{\pi(u_j)} = k_{u_j}$ if $p_i = \pi(u_j)$.*
2. *S, on receiving π from C, obtains the updated key $\mathsf{KE}_2 = (\overline{k}_1, \ldots, \overline{k}_n)$ from $\mathsf{KE}_1 = \{k_1, \ldots, k_n\}$ by applying π over the position set $\{1, \ldots, n\}$. Note that $\overline{k}_i = \overline{k}_{\pi(j)} = k_j$ if $i = \pi(j)$.*

QuPSI.Set Intersection*:*

1. *S computes Bloom_Y and evaluates $B = \mathsf{KE}_2 \oplus \mathsf{Bloom}_Y = \{\overline{k}_1 \oplus \mathsf{Bloom}_Y[1], \ldots, \overline{k}_n \oplus \mathsf{Bloom}_Y[n]\} = \{b_1, \ldots, b_n\}$ and sends B to C.*
2. *On receiving $B = \{b_1, \ldots, b_n\}$, the client C evaluates $A = \{\overline{k}_{p_1} \oplus b_{p_1}, \ldots, \overline{k}_{p_m} \oplus b_{p_m}\} = \{a_1, \ldots, a_m\}$. Note that it enables C to obtain the entries of $\{a_1, \ldots, a_m\} = \{\mathsf{Bloom}_Y[p_1], \ldots, \mathsf{Bloom}_Y[p_m]\}$ of Bloom_Y. In the following, C performs the following steps:*
   (a) *selects an empty set χ,*
   (b) *forms a resulting Bloom filter $\mathsf{Bloom} = \{g_1, \ldots, g_n\}$ of length n by setting $g_i = 1$ if $a_i = \mathsf{Bloom}_Y[p_i] = 1$; otherwise, sets $g_i = 0$,*
   (c) *for $i = 1, \ldots, v$, if x_i passes the set membership test of Bloom then includes x_i into χ,*
   (d) *outputs the resulting set χ as the desired intersection $X \cap Y$.*

4 Security Analysis

A PSI protocol, with functionality $F_{PSI} : (X, Y) \rightarrow (X \cap Y, \perp)$, should attain the following security properties:

1. **Correctness:** On completion of the protocol, the exact intersection i.e., $X \cap Y$ (possibly empty) should be the output of C.

2. **Client's privacy:** At the end of the protocol, C should obtain only $X \cap Y$, not beyond that.

3. **Server's privacy:** On completion of the interaction, S should not obtain anything.

Now, we discuss each of the aforementioned security requirements for our QuPSI.

Correctness: We need to prove that χ provides the desired intersection $X \cap Y$ for showing the the correctness of QuPSI. The set χ is obtained by collecting all such elements of X for which the set membership test of Bloom is satisfied. The resulting Bloom filter Bloom is formed by inserting 1 at all such p_i's where $a_i = \mathsf{Bloom}_Y[p_i] = 1$'s are 1 and 0's at the remaining entries. Moreover, $\{p_1, \ldots, p_m\}$ represents the positions with non-zero entries in Bloom_X i.e., $\mathsf{Bloom}_X[j] = 1$ for $j \in \{p_1, \ldots, p_m\}$ and $\mathsf{Bloom}_X[j] = 0$ for $j \in \{1, \ldots, n\} \setminus \{p_1, \ldots, p_m\}$. Thus $\mathsf{Bloom}[i] = 1$ if and only if $\mathsf{Bloom}_X[i] = 1$ and $\mathsf{Bloom}_Y[i] = 1$. An element $x \in X$ satisfies the set membership test for Bloom implies $\mathsf{Bloom}[H_i(x)] = 1$ for $i = 1, \ldots, q$. Furthermore, $\mathsf{Bloom}[H_i(x)] = 1$ implies $\mathsf{Bloom}_X[H_i(x)] = 1$ and $\mathsf{Bloom}_Y[H_i(x)] = 1$ for all $i = 1, \ldots, q$. Therefore, x satisfies the set membership test of Bloom_Y which implies $x \in Y$ except with negligible probability ϵ (false positive rate of the Bloom filter Bloom_Y). Hence, we may conclude that $x \in X$ satisfies the set membership test for Bloom implies $x \in X \cap Y$ except with negligible probability ϵ. In other words, χ determines the $X \cap Y$.

Client C's privacy: The permutation π is the only classical message which is sent to S from C. The server S would not be able to obtain any information about Bloom_X from π since it does not know $\{u_1, \ldots, u_m\}$. We now show that the probability of determining the position set $\{u_1, \ldots, u_m\}$ of C's part by a dishonest S is negligible (atmost $\frac{1}{2^\kappa}$). Initially, a dishonest server S can prepare a state of two qubits $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ for applying entangle-measure attack, where the first qubit is received by C and the second qubit is obtained by S. Then S can measure the state in its register to get information associated to the C's measurement's conclusiveness. Nevertheless, this attack brings bit errors as mentioned in [9, 14]. This is because if knowledge associated to the conclusiveness of C's bits is obtained by S then it will never be able to obtain the knowledge about C's recorded bit values. In other words, it is not possible for S to simultaneously get both the correct bit value and the address of C's correct measurement basis. As a consequence, S is unable to simultaneously determine the correctly measured bit value r_i of C and its position i. In the proposed QuPSI, the honesty of S is checked by C by comparing κ measurement bit values with the associated bits declared by S. Thereby, the honest test will be passed by a dishonest S atmost with negligible probability $\frac{1}{2^\kappa}$. Consequently, it can be concluded that C's privacy is preserved.

Server S's privacy: Note that if C wants to obtain any information about $Y \setminus X \cap Y$ then it requires to store more bit values in T. To perform this task, during the first step of Raw Key Storing phase [9, 14], C can store and use more effective measurements on the qubits sent by S to C. We now explore C's possible measurements.
Let us consider simple measurement of C. In this case, if S announces 0 for declaring that a qubit lies in $\{|0\rangle, |0'\rangle\}$ then C applies the optimal unambiguous state discrimination (USD) measurement [12, 21] to identify the qubit's state. Note that if $F(\phi_0, \phi_1)$ is the

fidelity between the two states (ϕ_0, ϕ_1) to be discriminated then the success probability of the USD measurement is bounded by $1 - F(\phi_0, \phi_1)$. According to [9], the probability is $\mathsf{Prob}_{USD} = 1 - \langle 0|0^{'}\rangle = 1 - \cos\alpha$. As a consequence, C's advantage is negligible for small α.
On the other hand, C can also perform joint measurement [14] on the γ qubits which contribute to the element of the final key. Here, C wants to obtain directly the final key's bit value, without distinguishing the raw key's individual bit values. To perform this, C can utilize the following two possible measurements:

1. *Helstrom's minimal error-probability measurement* for distinguishing two quantum states with highest information gain [8, 11]. If $D(\rho_0, \rho_1)$ is the trace distance between two equally likely quantum states ρ_0 and ρ_1 then the probability of distinguishing ϕ_0 from ϕ_1 and determining the right state is atmost $\mathsf{Prob}_D = \frac{1}{2} + \frac{1}{2}D(\phi_0, \phi_1)$. As a consequence, a final key bit can be determined with the probability at most $\mathsf{Prob}_D = \frac{1}{2} + \frac{1}{2}sin^{\gamma}\alpha$ which is closed to 1/2 for small α.

2. *USD measurement* is the other one, where the success probability of discriminating the two γ-qubit mixed states related to odd and even parity can be obtained. Note that the probability rapidly decreases with γ. Thus, following [9], we can conclude that for small α, the advantage of C is distinctly decreased.

Therefore, in order to preserve S's privacy, we need to select small α. This is to be observed that by choosing $n \gg v\lambda$, one can make α very small since $v\lambda > m$. Thereby, according to [9, 14], we may conclude that storing bit values more than $m+\kappa$ is impossible for C. Thus, the privacy of S is preserved.

5 Efficiency Analysis

In our scheme, quantum computation and quantum communication are only needed during the phase QuPSI.Raw Key. Moreover, the same raw key allows multiple execution of set intersection functionality for the private sets of C, where m remains same. Thereby, only one time quantum computation and quantum communication enables multiple execution of set intersection functionality. The round cost, communication cost and computation cost of the proposed scheme are discussed below.

Round cost: QuPSI.Raw Key needs 5 rounds (1 quantum and 4 classical), while each of QuPSI.Asymmetric Key and QuPSI.Set Intersection needs 1 classical round.

Communication cost: $O(n+\kappa)$ qubits and $O(n+\kappa)$ classical bits are required to be transmitted by S during QuPSI.Raw Key. In addition, S has to transmit n classical bits during QuPSI.Set Intersection. While, $O(n+\kappa)$ classical bits during QuPSI.Raw Key and a permutation π are required to be transmitted during QuPSI.Asymmetric Key.

Computation cost: $O(n+\kappa)$ projective measurements are required to be performed by C in 2- dimensional Hilbert space during QuPSI.Raw Key. In addition, it requires to perform one permutation π and m XOR operations in QuPSI.Asymmetric Key and QuPSI.Set Intersection respectively. While, computation of the permutation π and n XOR operations are required to be performed by S in QuPSI.Asymmetric Key and QuPSI.Set Intersection respectively.

Table 1 : Comparison summary of quantum PSI protocols

Protocol	Ours	[6]	[22]	[5]	[19]
Quantum resource	SP	SP	MPES	MPES	MPES
Complicated Oracle operators	not required	not required	required	required	required
Dimension of the Hilbert Space	2	2	N	N	N
Simple single-particle projective measurements	yes	yes	no	no	no
Multiple execution of set intersection functionality with only one time quantum communication and quantum computation	yes	yes	no	no	no
Intersection cardinality revealed to server	no	no	no	no	yes
Communication	$O(n+\kappa)$-qubit $O(n+\kappa)$-bit	$O(N+\kappa)$-qubit $O(N+\kappa)$-bit	$O(v \log N)$-qubit	$O(v \log N)$-qubit	$O(v+l)\log N$-qubit $O(\beta(\log l + v \log N))$-bit
Computation	$O(n+\kappa)$	$O(N+\kappa)$	$O(v)$	$O(v)$	$O(N+l)$
Round complexity in set intersection phase	1	1	2	3	4

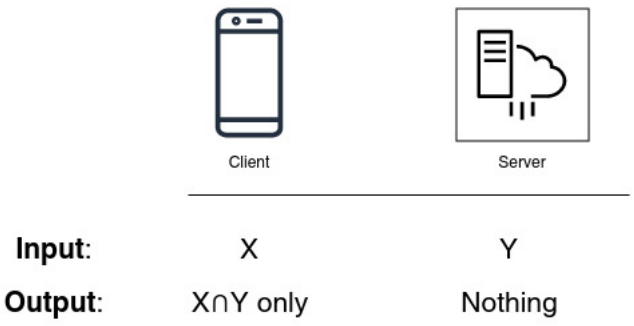
SP = single photons, MPES = multi-particle entangled states, l, κ = security parameters, N = the universal set's size, v = size of C's set, n= size of Bloom filter=$O(w)$, β = cardinality of the intersection

Table 1 represents a comparative summary of QuPSI with the existing quantum PSI protocols. The most attractive feature our scheme is that none of the communication and computation cost depends on universal set's size N, unlike the existing quantum PSI protocols. In other words, the communication cost and computation cost of our scheme are much less than those of the existing schemes since $n \ll N$.

6 Application to Contact Tracing

As a key building block for contact tracing, we presented the design of QuPSI protocol using Quantum Cryptography(QC) that enables two parties, each holding a dataset, to determine the intersection of datasets privately without revealing the actual data. Our protocol provides a robust and efficient solution to privacy vulnerabilities.

To illustrate, consider two datasets, X and Y. Let X be a dataset comprising a person's physical contact history (recorded using a cell phone). Let Y be the dataset that keeps the IDs of the diagnosed patients for the disease. Now our protocol enables the user to check if their contact history dataset matches the dataset of diagnosed patients without disclosing their personal information to the server. In addition to that client would then only learn the intersection, nothing more than that.



Using these intersections, we can discover if a person and a diagnosed patients were in close proximity during an appropriate period in the past.

Making use of QuPSI in the background as the cryptographic building block, we can see if the dataset of local contact histories of the individuals and database keeping the identities of the diagnosed patients have any match. QuPSI, by design, ensures that both parties' private information- healthy individuals and diagnosed patients remain secure. It also assures that the identities of the diagnosed individuals are concealed from others.

Therefore our PSI protocol QuPSI can be used to form an efficient, high-performance contact tracing system that provides strong privacy guarantees.

7 Conclusion

In this paper, we utilized the asymmetric key distribution of [9] and Bloom filter to design an unconditionally secure two-party quantum private set intersection protocol QuPSI. The basic principles of quantum mechanics ensure the security of QuPSI. Thereby, it is resistant against well-known quantum attacks and attains long-term security, unlike the classical PSI protocols. In this work, we put forward a cryptographic building block QuPSI for privacy-preserving contact tracing. The QuPSI ensures that both parties' private information- healthy individuals and diagnosed patients remain secure. In the proposed scheme, quantum resources are single photons and simple single-particle projective measurements are needed similar to [6]. Thus, our design is more feasible to implement with the present technology than [5, 19, 22]. On a more positive note, the communication cost and computation cost of our scheme are not dependent on universal set's size, unlike [5, 6, 19, 22]. In particular, our scheme attains optimal communication and computation cost in the context of quantum PSI. Moreover, only one time quantum computation and quantum communication enables multiple computation of set intersection functionality. Extending our two-party quantum PSI to multi-party would be an interesting direction of future work.

References

[1] Aydin Abadi, Sotirios Terzis, and Changyu Dong. O-psi: delegated private set intersection on outsourced datasets. In *IFIP International Information Security Conference*, pages 3–17. Springer, 2015.

[2] Aydin Abadi, Sotirios Terzis, and Changyu Dong. Vd-psi: verifiable delegated private set intersection on outsourced private datasets. *Finanacial Cryptography and Data Security*, 2016.

[3] Aydin Abadi, Sotirios Terzis, Roberto Metere, and Changyu Dong. Efficient delegated private set intersection on outsourced private datasets. *IEEE Transactions on Dependable and Secure Computing*, 2017.

[4] Burton H Bloom. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13(7):422–426, 1970.

[5] Xiaogang Cheng, Ren Guo, and Yonghong Chen. Cryptanalysis and improvement of a quantum private set intersection protocol. *Quantum Information Processing*, 16(2):37, 2017.

[6] Sumit Kumar Debnath, Kunal Dey, Nibedita Kundu, and Tanmay Choudhury. Feasible private set intersection in quantum domain. *Quantum Information Processing*, 20(1):1–11, 2021.

[7] M. J Freedman, K. Nissim, and B. Pinkas. Efficient private matching and set intersection. In *Advances in Cryptology-EUROCRYPT 2004*, pages 1–19. Springer, 2004.

[8] Christopher A Fuchs. Distinguishability and accessible information in quantum theory. *arXiv preprint quant-ph/9601020*, 1996.

[9] Fei Gao, Bin Liu, Qiao-Yan Wen, and Hui Chen. Flexible quantum private queries based on quantum key distribution. *Optics express*, 20(16):17411–17420, 2012.

[10] Carmit Hazay and Muthuramakrishnan Venkitasubramaniam. Scalable multi-party private set-intersection. In *IACR International Workshop on Public Key Cryptography*, pages 175–203. Springer, 2017.

[11] Carl W Helstrom and Carl W Helstrom. *Quantum detection and estimation theory*, volume 3. Academic press New York, 1976.

[12] Ulrike Herzog and János A Bergou. Optimum unambiguous discrimination of two mixed quantum states. *Physical Review A*, 71(5):050301, 2005.

[13] Roi Inbar, Eran Omri, and Benny Pinkas. Efficient scalable multiparty private set-intersection via garbled bloom filters. In *International Conference on Security and Cryptography for Networks*, pages 235–252. Springer, 2018.

[14] Markus Jakobi, Christoph Simon, Nicolas Gisin, Jean-Daniel Bancal, Cyril Branciard, Nino Walenta, and Hugo Zbinden. Practical private database queries based on a quantum-key-distribution protocol. *Physical Review A*, 83(2):022301, 2011.

[15] Seny Kamara, Payman Mohassel, Mariana Raykova, and Saeed Sadeghian. Scaling private set intersection to billion-element sets. In *International Conference on Financial Cryptography and Data Security*, pages 195–215. Springer, 2014.

[16] Alireza Kavousi, Javad Mohajeri, and Mahmoud Salmasizadeh. Improved secure efficient delegated private set intersection. *arXiv preprint arXiv:2004.03976*, 2020.

[17] Vladimir Kolesnikov, Naor Matania, Benny Pinkas, Mike Rosulek, and Ni Trieu. Practical multi-party private set intersection from symmetric-key techniques. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1257–1272. ACM, 2017.

[18] Fenghua Li, Ben Niu, Yanchao Wang, et al. Server-aided private set intersection based on reputation. *Information Sciences*, 2016.

[19] Arpita Maitra. Quantum secure two-party computation for set intersection with rational players. *Quantum Information Processing*, 17(8):197, 2018.

[20] Atsuko Miyaji and Shohei Nishida. A scalable multiparty private set intersection. In *International Conference on Network and System Security*, pages 376–385. Springer, 2015.

[21] Philippe Raynal. Unambiguous state discrimination of two density matrices in quantum information theory. *arXiv preprint quant-ph/0611133*, 2006.

[22] Run-hua Shi, Yi Mu, Hong Zhong, Jie Cui, and Shun Zhang. An efficient quantum scheme for private set intersection. *Quantum Information Processing*, 15(1):363–371, 2016.

[23] Run-hua Shi, Yi Mu, Hong Zhong, and Shun Zhang. Quantum oblivious set-member decision protocol. *Physical Review A*, 92(2):022309, 2015.

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

[25] Qiang Wang, Fucai Zhou, Jian Xu, and Su Peng. Tag-based verifiable delegated set intersection over outsourced private datasets. *IEEE Transactions on Cloud Computing*, 2020.

[26] Xiaoyuan Yang, Xiaoshuang Luo, Xu An Wang, and Shuaiwei Zhang. Improved outsourced private set intersection protocol based on polynomial interpolation. *Concurrency and Computation: Practice and Experience*, 30(1):e4329, 2018.

[27] En ZHANG and Ganggang JIN. Cloud outsourcing multiparty private set intersection protocol based on homomorphic encryption and bloom filter. *Journal of Computer Applications*, (8):20, 2018.

[28] En Zhang, Fenghua Li, Ben Niu, and Yanchao Wang. Server-aided private set intersection based on reputation. *Information Sciences*, 387:180–194, 2017.

[29] En Zhang, Feng-Hao Liu, Qiqi Lai, Ganggang Jin, and Yu Li. Efficient multi-party private set intersection against malicious adversaries. In *Proceedings of the 2019 ACM SIGSAC Conference on Cloud Computing Security Workshop*, pages 93–104, 2019.

A Toy Example

Let $X = \{Alice,\ Bob,\ Hary\}$, $Y = \{Hary,\ Jones,\ Jack,\ Bob,\ Henry\}$, $n = 18$, $\lambda = 2$, $H = \{H_1, H_2\}$. Also let $H_1(Alice) = 3$, $H_2(Alice) = 7$, $H_1(Bob) = 8$, $H_2(Bob) = 1$, $H_1(Hary) = 7$, $H_2(Hary) = 16$, $H_1(Jones) = 15$, $H_2(Jones) = 16$, $H_1(Jack) = 10$, $H_2(Jack) = 12$, $H_1(Henry) = 18$ and $H_2(Henry) = 1$. Then $\mathsf{Bloom}_X = [101000110000000100]$, $\mathsf{Bloom}_Y = [100000110101001101]$, the number of non-zero entries in Bloom_X is $m = 5$ and the associated positions are $p_1 = 1$, $p_2 = 3$, $p_3 = 7$, $p_4 = 8$, $p_5 = 16$.

QuPSI.Raw Key:

1. Set $\alpha = \sin^{-1}(\sqrt{\frac{7}{10}})$ in asymmetric key distribution of [9]. Then S obtains $n + \kappa = 20$ bit key $\mathsf{KE} = \{r_1, \ldots, r_{20}\}$ and C obtains $m + \kappa = 9$ bits (say $T = \{r_2, r_3, r_5, r_9, r_{12}, r_{15}, r_{20}\}$) of KE, where $r_i \in \{0, 1\}$ for $i = 1, \ldots, 20$.
2. Let C chooses bits of 2nd and 5th positions of KE i.e., r_2 and r_5 for comparison with S's part. After the deletion of r_2 and r_5 the updated part of S becomes $\mathsf{KE}_1 = \{k_1, \ldots, k_{18}\}$ and the updated part of C becomes $T_1 = \{k_2, k_7, k_{10}, k_{13}, k_{18}\}$, where $k_1 = r_1$, $k_2 = r_3$, $k_3 = r_4$, $k_i = r_{i+2}$ for $i = 4, \ldots, 18$. Note that $\{u_1, u_2, u_3, u_4, u_5\} = \{2, 7, 10, 13, 18\}$.

QuPSI.Asymmetric Key:

1. Suppose C selects the permutation π over the set $\{1,\ldots,20\}$ such that $\pi(2) = 3, \pi(7) = 7, \pi(10) = 1, \pi(13) = 16, \pi(18) = 8$ and $\pi(j) = i$ for $j \in \{1,\ldots,18\} \setminus \{2,7,10,13,18\}$ and $i \in \{1,\ldots,18\} \setminus \{1,3,7,8,16\}$ in some order i.e.,$\{u_1,u_2,u_3,u_4,u_5\} = \{2,7,10,13,18\}$ is mapped to $\{p_1,p_2,p_3,p_4,p_5\} = \{1,3,7,8,16\}$. Then the updated part of C becomes $T_2 = \{\overline{k}_1 = k_{10}, \overline{k}_3 = k_2, \overline{k}_7 = k_7, \overline{k}_8 = k_{18}, \overline{k}_{16} = k_{13}\}$.
2. S, on receiving π from C, obtains the updated key $\mathsf{KE}_2 = (\overline{k}_1,\ldots,\overline{k}_{18})$ from $\mathsf{KE}_1 = \{k_1,\ldots,k_{18}\}$ by applying π over the position set $\{1,\ldots,18\}$. Note that $\overline{k}_i = \overline{k}_{\pi(j)} = k_j$ if $i = \pi(j)$. Thereby, $\overline{k}_1 = k_{10}, \overline{k}_3 = k_2, \overline{k}_7 = k_7, \overline{k}_8 = k_{18}, \overline{k}_{16} = k_{13}$.

QuPSI.Set Intersection:

1. S evaluates $B = \mathsf{KE}_2 \oplus \mathsf{Bloom}_Y = \{\overline{k}_1 \oplus \mathsf{Bloom}_Y[1],\ldots,\overline{k}_{18} \oplus \mathsf{Bloom}_Y[18]\} = \{b_1,\ldots,b_{18}\}$ and sends B to C.
2. On receiving $B = \{b_1,\ldots,b_{18}\}$, the client C evaluates $A = \{\overline{k}_1 \oplus b_1, \overline{k}_3 \oplus b_3, \overline{k}_7 \oplus b_7, \overline{k}_8 \oplus b_8, \overline{k}_{16} \oplus b_{16}\} = \{a_1,\ldots,a_5\}$. Note that $a_1 = \overline{k}_1 \oplus b_1 = \overline{k}_1 \oplus \overline{k}_1 \oplus \mathsf{Bloom}_Y[1] = \mathsf{Bloom}_Y[1] = 1$, $a_2 = \overline{k}_3 \oplus b_3 = \overline{k}_3 \oplus \overline{k}_3 \oplus \mathsf{Bloom}_Y[3] = \mathsf{Bloom}_Y[3] = 0$, $a_3 = \overline{k}_7 \oplus b_7 = \overline{k}_7 \oplus \overline{k}_7 \oplus \mathsf{Bloom}_Y[7] = \mathsf{Bloom}_Y[7] = 1$, $a_4 = \overline{k}_8 \oplus b_8 = \overline{k}_8 \oplus \overline{k}_8 \oplus \mathsf{Bloom}_Y[8] = \mathsf{Bloom}_Y[8] = 1$, $a_5 = \overline{k}_{16} \oplus b_{16} = \overline{k}_{16} \oplus \overline{k}_{16} \oplus \mathsf{Bloom}_Y[16] = \mathsf{Bloom}_Y[16] = 1$.
   In the following, C performs the following steps:
   (a) forms a resulting Bloom filter $\mathsf{Bloom} = [100001100000000100] = \{g_1,\ldots,g_{18}\}$ by setting $g_i = 1$ if $a_i = \mathsf{Bloom}_Y[p_i] = 1$; otherwise, sets $g_i = 0$,
   (b) computes $H_1(Alice) = 3$, $H_2(Alice) = 7$, $H_1(Bob) = 8$, $H_2(Bob) = 1$, $H_1(Hary) = 7$, $H_2(Hary) = 16$. Note that $\mathsf{Bloom}[H_1(Alice)] = 0$, $\mathsf{Bloom}[H_2(Alice)] = 1$, $\mathsf{Bloom}[H_1(Bob)] = 1$, $\mathsf{Bloom}[H_2(Bob)] = 1$, $\mathsf{Bloom}[H_1(Hary)] = 1$, $\mathsf{Bloom}[H_2(Hary)] = 1$
   (c) outputs $\{Bob,\ Hary\}$ as the desired intersection $X \cap Y$, since $Alice$ does not satisfy the set membership test of Bloom, while each of Bob and $Hary$ satisfies that test.