Advances in Mathematics of Communications
Vol. 19, No. 2, April 2025, pp. 494-545
doi:10.3934/amc.2024009

FULLY DEVICE INDEPENDENT QUANTUM PRIVATE QUERY

Jyotirmoy Basak[✉*1] and Kaushik Chakraborty[✉2]

[1]Applied Statistics Unit, Indian Statistical Institute, Kolkata, India

[2]School of Informatics, The University of Edinburgh, UK

(Communicated by Delaram Kahrobaei)

Abstract. Quantum Private Query (QPQ) is an unconditional secure mistrustful cryptographic primitive which is assumed to be a probabilistic version of the Oblivious Transfer (OT) schemes or an imperfect version of Symmetric Private Information Retrieval (SPIR) schemes. Recently, Maitra et al. (Phys. Rev. A, 2017) identified that the B92 QKD-based QPQ scheme proposed by Yang et al. (Quant. Inf. Process., 2014) is vulnerable whenever the devices involved in that scheme are dubious and to improve the overall security, they suggested a semi-Device Independent (DI) proposal for that QPQ scheme by introducing a local test at the server's end. In this work, we overcome the limitation of the Maitra et al. proposal by removing trustworthiness from all the (involved) devices, and suggest a full DI proposal for the Yang et al. scheme, exploiting a proper self-testing mechanism of observables along with the local version of the tilted CHSH game. We compare the performance of our proposal with a recent full DI-QPQ scheme (arxiv 1901.03042) and discuss their relative advantages. Additionally, we present a DI proposal for a modified version of the Yang et al. scheme, enabling the client to retrieve maximum raw key bits during the oblivious key generation phase. We evaluate the security of all our proposals with a formal analysis.

1. **Introduction.** The field of quantum cryptography has expanded greatly since Bennett and Brassard [8] first proposed it in 1984. One area that has received significant attention is Quantum Private Query (QPQ). It's a two-party cryptographic primitive where a client queries a database held by a server to get the values of the data bits corresponding to the queried indices while allowing only a small amount of information about the unintended data bits to the client and the client's privacy is ensured in a cheat-sensitive way where it is assumed that the server will not cheat if a non-zero probability exists of being caught cheating.

The idea of QPQ originates from the idea of Private Information Retrieval (PIR) and Oblivious Transfer (OT) schemes. However, QPQ is closely related to a variant of the PIR scheme known as Symmetric PIR (SPIR). In PIR, the client wants to know some bits from the database holds by the server and the server sends the complete database to the client so that the client can retrieve the required bits. Although the server can not learn anything about the queries of the client, the client can retrieve many more bits from the database other than the required bits.

2020 *Mathematics Subject Classification.* Primary: 94A60; Secondary: 81P94.

Key words and phrases. Quantum private query, device independence, tilted CHSH, self-test, projective measurement, POVM.

*Corresponding author: Jyotirmoy Basak.

This inefficiency of the PIR is mitigated in the SPIR primitive where database security is also taken into account along with the client's privacy. The functionality of the SPIR scheme is similar to the 1 out of N OT where the client wants to know some data bits such that both the user privacy and the data privacy is maintained.

Functionally, QPQ, SPIR, and OT are related but have some differences. SPIR is similar to OT but it is not possible to design an unconditional secure OT scheme in either quantum or classical settings. However, unconditional secure SPIR schemes can be designed in a distributed database setting [14] with non-communication assumptions. QPQ is similar to 1 out of N OT or SPIR but with a weaker security requirement that allows the malicious user to learn a small amount of additional information about the database, and expects the server not to cheat if a non-zero probability exists of being caught. Because of this weaker security requirement, QPQ is considered a probabilistic version of 1 out of N OT or an imperfect version of SPIR schemes and this enables the design of unconditional secure QPQ schemes in a single database setting.

The history of Quantum Private Query (QPQ) started with the proposal by Giovannetti et al. [15], followed by [16] and [28]. However, all those initial schemes incorporated the idea of quantum memories and none of those were possible to implement practically. The first practical QPQ protocol was proposed by Jakobi et al. [19], which was based on Quantum Key Distribution (QKD) [32]. In 2012, Gao et al. [13] proposed a flexible generalization of this QPQ scheme. Rao et al. [29] suggested two more efficient modifications to the protocol's classical post-processing in the same year. Zhang et al. [36] proposed a counterfactual QKD [27] based QPQ scheme in 2013, and Yang et al. proposed a flexible QPQ scheme [35] based on B92 QKD [7] in 2014.

It is already discussed that designing QPQ protocols emerged as a response to the difficulties encountered in developing unconditionally secure single-server SPIR schemes that enforce a cheat sensitivity in the adversarial model assuming that if there is a non-zero probability of being caught cheating then the server will not cheat. In reality, the desired primitive is as follows.

- The malicious client's knowledge of additional data bits is limited to a small fraction beyond what is intended to know by her. The server's goal is to minimize dishonest clients' knowledge of extra information about the database.
- While being honest, the server can only gain limited information about the client's query indices. Jakobi et al. [19] demonstrated that a dishonest server can't obtain both conclusiveness information and the values of the raw key bits recorded by the client during the oblivious key generation phase. If the server attempts to retrieve more information about the client's query indices, there is a risk of providing false information about the intended data bits to the client, which would damage the server's reputation as a database owner. Thus, in the QPQ primitive, it is assumed that the server will not cheat if there exists a non-zero probability of being caught cheating.

Very recently, Maitra et al. [25] identified that the security of the existing QPQ schemes is based on the trustful assumptions over the devices (source as well as measurement devices) involved in that particular protocol. As an example, they consider the Yang et al. [35] QPQ scheme and showed that the client can retrieve more data bits than the intended one if the source device (which supplies the shared states) does not work accordingly. To overcome this security loophole and remove the trustful assumptions over the devices, they suggested a Device Independent (DI)

version of the QPQ scheme [35] in [25]. They introduced a local testing phase at the server side in [25] which certifies the measurement devices at the server side and the state generation device. However, this test does not certify the measurement devices on the client's side. So, their proposal in [25] is a semi-DI version of the QPQ scheme [35]. Although this limitation is mentioned in [5], to the best of our knowledge, the procedure for proper DI certification of the QPQ scheme [35] is not mentioned anywhere.

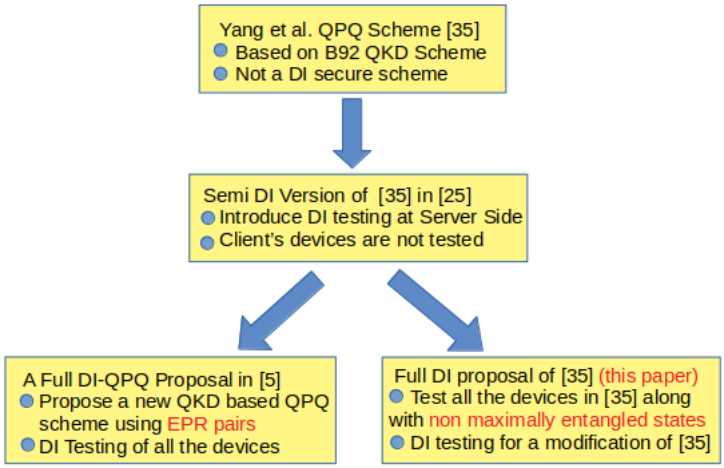

FIGURE 1. Evolution of QPQ schemes in the DI scenario. The highlighted portion (in red) is the main difference between the QPQ scheme [5] and the QPQ scheme [35]

In this direction, here we overcome the limitations in [25] and propose a full DI version of the Yang et al. QPQ scheme [35]. Our proposed scheme exploits the proper self-testing mechanism of the observables involved in [35] along with a local version of the tilted CHSH test to certify all the measurement devices. We also compare the performance of this proposed full DI version of the QPQ scheme [35] with the performance of the full DI-QPQ scheme mentioned in [5] and discuss the relative advantages of both these protocols. We further come up with a DI proposal for a modified version of the QPQ scheme [35] where the client can retrieve the maximum number of raw key bits at her end. In opposition to current DI-QPQ approaches, here in this modified proposal, we replace the usual projective measurement at the client's side with the optimal POVM measurement to retrieve the maximum number of shared raw key bits. A flow diagram involving the evolution of the QPQ scheme [35] in the DI scenario is shown in Figure 1.

1.1. **Revisiting the QPQ scheme [35] and its DI version in [25].** In this section, we first revisit the QPQ protocol mentioned in [35] and then restate the DI version of this QPQ scheme introduced in [25]. In [35], the authors proposed a QKD-based QPQ scheme exploiting the idea of B92 QKD protocol. Their proposed QPQ scheme is composed of mainly two phases namely the key generation phase and the private query phase.

In the key generation phase of the QPQ scheme [35], the server Bob and the client Alice share non-maximally entangled states between them which are of the form $\frac{1}{\sqrt{2}}(|0\rangle|\phi_0\rangle + |1\rangle|\phi_1\rangle)$ where $|\phi_0\rangle = \left(\cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle\right)$ and $|\phi_1\rangle = \left(\cos\frac{\theta}{2}|0\rangle - \sin\frac{\theta}{2}|1\rangle\right)$ (at the beginning of the protocol, the exact value of this θ is decided by the server Bob to the third party based on the number of raw key bits that Bob wants Alice to know after the key generation phase). Bob initially receives states from the third party and subsequently sends the second particle of each state to Alice. Upon receiving the particles, Alice declares instances where she receives them correctly, discarding those instances where reception is incorrect. For each successfully shared state, Bob measures his particle in the $|0\rangle, |1\rangle$ basis, while Alice randomly measures her particle in either the $|\phi_0\rangle, |\phi_0^\perp\rangle$ basis or the $|\phi_1\rangle, |\phi_1^\perp\rangle$ basis. Post-measurement, Bob assigns the i-th raw key bit as 0 if the corresponding measurement outcome is $|0\rangle$ and 1 (i.e., for outcome $|1\rangle$) otherwise.

Analysis of the states used in the QPQ scheme [35] reveals that, for a given instance, if Bob observes the outcome $|0\rangle$ ($|1\rangle$) at his end, the corresponding state at Alice's side collapses to $|\phi_0\rangle$ ($|\phi_1\rangle$). Since $|\phi_0\rangle$ and $|\phi_1\rangle$ are non-orthogonal, Alice can conclusively determine Bob's raw key bits only for instances where she observes outcomes $|\phi_0^\perp\rangle$ or $|\phi_1^\perp\rangle$. When Alice observes $|\phi_0^\perp\rangle$ for an instance, she deduces that the state at her side is surely $|\phi_1\rangle$ and the corresponding raw key bit at Bob's side is 1. Similarly, if she observes $|\phi_1^\perp\rangle$, she concludes that the state at her side is surely $|\phi_0\rangle$ and the corresponding raw key bit at Bob's side is 0. After measurement, Alice and Bob engage in classical post-processing on their raw key bits, reducing Alice's knowledge of the final key to one bit. Consequently, after this key generation phase, Bob possesses the entire key, whereas Alice knows only a subset of bits (ideally, just one bit after post-processing)

In the private query phase, if Alice knows the j-th bit of the final key and wants to retrieve the bit indexed by i of the database then she declares the integer $s = (j - i)$ publicly. Bob then shifts his key by s bits, encrypts the database with this shifted key using the one-time pad, and sends it to Alice. Alice decrypts the j-th bit and gets the required element of the database.

It is already mentioned in [35] that by following the specified strategy, Alice can conclusively retrieve only $\frac{\sin^2\theta}{2}$ (on average) fraction of bits of the entire raw key obtained by Bob. This guarantees the security of the proposed QPQ scheme because although Alice gets the whole encrypted database, she can not retrieve all the database bits because of her partial knowledge about the raw key as well as the final key.

However, it was shown in [25] that if the dishonest Alice colludes with the third party and supplies the states of the form $(\alpha|0\rangle|\phi_0\rangle + \beta|1\rangle|\phi_1\rangle)$ where $|\alpha|^2 = \left(\frac{1}{2} + \epsilon\right)$ and $|\beta|^2 = \left(\frac{1}{2} - \epsilon\right)$ then the dishonest Alice can retrieve additional $2\epsilon^2 \sin^2\theta$ fraction of bits of the entire raw key. For this reason, to overcome the security loophole (a schematic diagram of different phases of the DI-QPQ scheme [25] is shown in the left picture of Figure 3), a DI version of the QPQ scheme [35] was proposed in [25].

In the DI scheme proposed in [25], the server Bob performs a tilted version of the original CHSH test locally to certify the devices. Although this local test certifies the states and Bob's measurement devices (as depicted in [25, Theorem 1 & Theorem 2]), this local test actually fails to provide any certification about Alice's measurement devices (we have demonstrated an attack in Section 2 which shows this limitation of the DI-QPQ proposal [25]) as those devices aren't involved in the local test mentioned in [25]. This implies that the scheme mentioned in [25]

is a semi-DI version of the Yang et al. [35] QPQ scheme. Here we overcome this limitation of the scheme [25] and propose a full DI version of the Yang et al. [35] QPQ scheme.

1.2. **Preliminaries.** A *quantum state* in the qubit system can be represented as a unit (column) vector in the $\mathbb{C}^2$ plane, spanned by the basis states $|0\rangle$ and $|1\rangle$ represented as follows.

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Any general one qubit state $|\psi\rangle$ can be written as a *superposition* of the basis states in the following way.

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{1}$$

The *conjugate transpose* of a quantum state $|\psi\rangle$ is denoted as $\langle\psi| = (|\psi\rangle)^\dagger$ and is represented as a row vector as follows.

$$\langle\psi| = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}^\dagger = \begin{bmatrix} \alpha^\dagger & \beta^\dagger \end{bmatrix}$$

The *inner product* between two quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$ is denoted as $\langle\psi_1\,|\psi_2\rangle$ and is represented as follows.

$$\langle\psi_1\,|\psi_2\rangle = \begin{bmatrix} \alpha_1^\dagger & \beta_1^\dagger \end{bmatrix} \cdot \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} = \alpha_1^\dagger\alpha_2 + \beta_1^\dagger\beta_2$$

The *outer product* between two quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$ is denoted as $|\psi_1\rangle\,\langle\psi_2|$ and is represented as follows.

$$|\psi_1\rangle\,\langle\psi_2| = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \cdot \begin{bmatrix} \alpha_2^\dagger & \beta_2^\dagger \end{bmatrix} = \begin{bmatrix} \alpha_1\alpha_2^\dagger & \alpha_1\beta_2^\dagger \\ \beta_1\alpha_2^\dagger & \beta_1\beta_2^\dagger \end{bmatrix}$$

For the system of *pure state* (i.e., the system having all the states of the same form say like the one mentioned in equation 1), the state of that corresponding system can also be represented (along with the representation mentioned in equation 1) as $\rho = |\psi_1\rangle\,\langle\psi_1|$ which is known as the *density matrix representation*.

If a quantum system is in a state $\{|\psi_i\rangle\}_{1\le i\le n}$ with probability p_i then the state of that system is called a *mixed state*. A mixed state can only be written in the form of a density matrix or a density operator which is a positive semidefinite operator having unit trace and is represented in the following form.

$$\rho = \sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|\,.$$

Note that, a quantum state ρ is pure if it satisfies $\mathrm{Tr}[\rho^2] = 1$, and is mixed if it satisfies $\mathrm{Tr}[\rho^2] < 1$.

Quantum measurement is described by a collection $\{M_m\}$ of measurement operators that act on the state space of the system being measured. The measurement operators must satisfy the completeness condition i.e.,

$$\sum_m M_m^\dagger M_m = \mathbb{I}.$$

Here, 'm' refers to the measurement outcome generated after the experiment. If a quantum system $|\psi\rangle$ is measured then after measurement, the probability $\Pr(m)$ of occurring the result m is given by,

$$\Pr(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle .$$

After the measurement, the state of the system will be,

$$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}.$$

For example, one can consider the state $|\psi\rangle$ as mentioned in equation 1. If this state $|\psi\rangle$ is measured in $\{|0\rangle, |1\rangle\}$ basis, then the measurement outcome will be $|0\rangle$ with probability $|\alpha|^2$ and $|1\rangle$ with probability $|\beta|^2$.

There are basically two types of measurement, *projective measurement* and *POVM*. A measurement is called *projective measurement* if the measurement operators $\Pi_m = M_m^\dagger M_m$ are orthogonal projectors i.e., they satisfy the property $\Pi_m^2 = \Pi_m$ and they sum up to the identity matrix. This measurement has the property that performing the same measurement again immediately after the one yields the same result with probability 1.

If the post-measurement state is not of particular interest, then one can perform a more efficient measurement known as *Positive-Operator-Valued Measurement (POVM)*. This measurement is described by a set of positive semi-definite hermitian matrices that sum to the identity matrix i.e., $\{E_m\}$ such that $\sum_m E_m = \mathbb{I}$ where the index m denotes the measurement outcome. If a pure quantum state $|\psi\rangle$ is measured then for this measurement, the probability of getting the measurement outcome m is given by,

$$\Pr(m) = \langle\psi| E_m |\psi\rangle .$$

A significant distinction lies in the fact that the elements of a POVM are not necessarily orthogonal. Consequently, the count of POVM elements can exceed the dimension of the corresponding Hilbert space they act on. Conversely, for a projective measurement, the number of elements is constrained to be at most equal to the dimension of the Hilbert space.

1.2.1. *Notations used in our scheme.* Here, we enlist different notations that are used in our schemes.

- $\mathcal{K}$: The total number of states needed, assumed to be large.
- $|\psi\rangle_{\mathcal{B}_i\mathcal{A}_i}$: The ith shared state with the first qubit belonging to Bob ($\mathcal{B}_i$) and the second to Alice ($\mathcal{A}_i$).
- $\rho_{\mathcal{B}_i\mathcal{A}_i}$: The density matrix of the ith state.
- $\rho_{\mathcal{A}_i}$ and $\rho_{\mathcal{B}_i}$: The reduced density matrices for Alice and Bob, respectively, of the ith state.
- X: The database held by Bob, with N bits.
- R and $R_{\mathcal{A}}$: The entire raw key at Bob's and Alice's sides, respectively, each with kN bits.
- F and $F_{\mathcal{A}}$: The entire final key at Bob's and Alice's sides, respectively, each with N bits.
- R_i and $R_{\mathcal{A}_i}$: The ith raw key bit of Bob and Alice, respectively.
- F_i and $F_{\mathcal{A}_i}$: The ith final key bit of Bob and Alice, respectively.
- $\mathcal{I}_l$: The set of query indices made by Alice.
- D: Alice's POVM device in a modified proposal.

- A and B: Alice and Bob's measurement outcomes, respectively.
- $\mathcal{A}$ and $\mathcal{A}^*$: Alice with honest or dishonest behavior.
- $\mathcal{B}$ and $\mathcal{B}^*$: Bob with honest or dishonest behavior.
- $|\phi_0\rangle = \cos\frac{\theta}{2}|0\rangle + \sin\frac{\theta}{2}|1\rangle$
- $|\phi_1\rangle = \cos\frac{\theta}{2}|0\rangle - \sin\frac{\theta}{2}|1\rangle$.

1.2.2. *Adversarial model.* In a QPQ primitive, none of the parties trust each other, resulting in different security goals for each party. Protocol Correctness refers to the security of the entire protocol, while Data Privacy protects the security of the server (Bob) and User Privacy protects the security of the client (Alice). This work revisits the security definitions previously defined in [5] for the QPQ primitive.

Definition 1.1. Correctness of the protocol:

In case of honest implementation, after the *key generation phase*, Alice is highly likely to correctly retrieve the expected number of data bits through a single query. This means that if Alice is aware of X data bits and is expected to know Y data bits (according to the scheme), then following the *key generation phase*,

$$\Pr(|X - Y| \leq \delta_t \wedge \text{the scheme doesn't abort}) \geq P_c \tag{2}$$

where Bob tolerates a deviation of δ_t and the probability of X being within the range of $[Y - \delta_t, Y + \delta_t]$ is referred to as P_c, which should ideally be high.

Definition 1.2. Robustness of the protocol:

For honest implementation of our schemes, the likelihood of Alice knowing none of the final key bits (or data bits) and the protocol needing to restart after the *key generation phase* is low. More formally,

$$\Pr(\text{the parties abort the scheme in honest scenario}) \leq P_a \tag{3}$$

where the likelihood that no final key bits are known to Alice and the protocol terminates is represented by P_a, which ideally should be low.

Definition 1.3. Privacy of the database owner:

The privacy of the database owner is protected in a QPQ scheme if, in a single query, the dishonest Alice ($\mathcal{A}^*$) can only retrieve (on average) at most τ fraction of bits from the N-bit database X, where τ $(0 < \tau < 1)$ is very small compared to N, or if the scheme aborts with a high probability in the long run. If the number of bits extracted (on average) by dishonest Alice in a query is denoted as $D_{\mathcal{A}^*}$, then according to the above definition,

$$E_R(D_{\mathcal{A}^*}) \leq \tau N \tag{4}$$

where the expectation is calculated over the random coins R utilized in the protocol.

Definition 1.4. Privacy of the user:

A QPQ scheme ensures user privacy if either dishonest server Bob ($\mathcal{B}^*$) can correctly guess, on average, at most a small fraction δ of indices from $\mathcal{I}_l$ (the query index set of Alice) or the scheme terminates with high likelyhood in the long run. If the number of correctly guessed indices by dishonest Bob is denoted by $\mathcal{I}_{\mathcal{B}^*}$ then according to the above definition,

$$E_{R'}(\mathcal{I}_{\mathcal{B}^*}) \leq \delta l \tag{5}$$

where the expectation is based on the random coins R' used in the protocol.

1.3. **Our contribution:** In this paper, we focus on the Yang et al. [35] QPQ scheme that provides privacy for both the user and the database owner in a classical database search. While Maitra et al. [25] proposed a semi-DI version of that scheme for improved security, we present a further improvement with a full DI version. Our main contributions in this work are the proposals of different testing phases (for the QPQ scheme [35] and a modification of it) and the rigorous mathematical analysis that justify the incorporation of those testing phases towards the certification of the functionality of all the involved devices. All the proof techniques and the mathematical tools employed in our proposals to certify the full DI security are firmly rooted in standard linear algebra principles. The main operational descriptions of our contributions along with their corresponding mathematical tools and proof techniques can be summarized as follows.

1. In the DI proposal [25], the server Bob locally performs a variation of the actual CHSH test (where he measures the qubits using some randomly chosen orthogonal projectors instead of the actual orthogonal projectors defined for the CHSH test) which only involves the entangled states and his own measurement devices. Here, we show that the local test mentioned in [25] fails to preserve the data privacy of the database (as the client's measurement devices are not tested in [25]), and the client Alice can retrieve some additional raw key bits (as well as the data bits in a single database query) if she performs an optimal POVM measurement instead of the actual projective measurement mentioned in [35].
   - *Mathematical tools behind this contribution* : In Section 2, we show that for the semi DI-QPQ proposal [25], if Alice performs a particular measurement at her end that is composed of a specific collection of positive semi-definite Hermitian matrices (i.e., the optimal POVM operator that provides optimal distinction between the two specified non-orthogonal states), differing from the set of the particular orthogonal projectors stated in [35], then she can pass the testing phase in [25] and can retrieve some additional raw key bits (that may result in the retrieval of some additional data bits in a single database query) which violates the data privacy.
2. We propose a full DI version of the QPQ scheme [35] by exploiting a local version of the tilted CHSH test (mentioned in [1, 4]) at both the server and the client's side (in the *source device and Bob's measurement device verification phase* of our scheme) along with the self-testing of orthogonal projectors (mentioned in [20]) at the client Alice's side (in the *Alice's measurement device verification phase* of our proposal). Additionally, we perform a comparative evaluation between our proposed full DI version and the full DI-QPQ scheme introduced in [5], considering different parameters and discussing the relative advantages of both these schemes.
   - *Mathematical tools behind this contribution* : In Appendix B, we show how our mentioned TiltedCHSH test (in the full DI proposal of the QPQ scheme [35]) utilizes the sum-of-squares decompositions for the Bell operators to achieve tight upper bounds on the maximum quantum value of the corresponding tilted CHSH expression and certify the actual partially entangled state mentioned in [35] and the Hermitian and dichotomic operators (i.e., the observables) at the server Bob's side. For certifying

the set of orthogonal projectors at Alice's side, we adopt a commutation-based measure technique from [20] and discuss (in Appendix C) how our mentioned OBStestAlice authenticate the anticommuting observables at Alice's side from the observed Bell violation, dealing with rank-deficient reduced density matrices.

3. We further come up with a full DI proposal for a modified version of the QPQ scheme [35] where the client Alice can retrieve the maximum number of raw key bits at her end during the oblivious key generation phase by performing the optimal POVM measurement.
   - *Mathematical tools behind this contribution* : In our modified proposal, we exploit the proper self-testing mechanism of a particular class of positive semi-definite Hermitian matrices (i.e., POVM devices) along with the local version of the tilted CHSH test (mentioned in [1, 4]) and the self-testing of orthogonal projectors (mentioned in [20]) to certify all the devices. For the certification of non-maximally entangled states and orthogonal projectors at Bob's side, we follow the same approach as in our first proposal. However, for certifying the specific optimal set of positive semi-definite Hermitian matrices on Alice's side, we come up with a new distinct technique. We treat each matrix (i.e., every POVM element) from the set of positive semi-definite Hermitian matrices as a generalized form of the Bloch vector, and based on defined constraints and conditions, we solve relevant equations to demonstrate (in Appendix D) that if Alice attains the intended value of the security parameter defined in our POVMtestAlice algorithm, then it serves as authentication for her measurement device.

2. **An attack on the DI-QPQ scheme** [25]**.** In the DI-QPQ scheme [25], the server Bob first selects some entangled states (from the set of states that will be used for the QPQ scheme [35]) and performs a tilted version of the local CHSH test to certify the states and the measurement devices involved in the QPQ scheme [35]. However, this local test does not certify Alice's measurement devices as it only involves the entangled states and Bob's measurement devices. This implies that if Alice performs some other measurement at her side instead of the actual projective measurement (mentioned in [35]), then the local test mentioned in [25] can not detect that fraudulent device.

In the QPQ scheme [35], it is shown that if the involved devices behave accurately according to the proposed scheme, then Alice can guess (on average) approximately $\frac{\sin^2\theta}{2}$ fraction of bits from the entire raw key. Now suppose, for the QPQ scheme [35], Alice measures her qubits using the POVM $D = \{D_0, D_1, D_2\}$ instead of performing the projective measurements in $\{|\phi_0\rangle, |\phi_0^{\perp}\rangle\}$ or $\{|\phi_1\rangle, |\phi_1^{\perp}\rangle\}$ basis randomly where

$$\begin{aligned} D_0 &\equiv \frac{(\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle)(\sin\frac{\theta}{2}\langle 0| + \cos\frac{\theta}{2}\langle 1|)}{(1+\cos\theta)} \\ D_1 &\equiv \frac{(\sin\frac{\theta}{2}|0\rangle - \cos\frac{\theta}{2}|1\rangle)(\sin\frac{\theta}{2}\langle 0| - \cos\frac{\theta}{2}\langle 1|)}{(1+\cos\theta)} \\ D_2 &\equiv I - D_0 - D_1 \end{aligned}$$

In this case, Alice can successfully pass the local CHSH test (mentioned in [25]) if Bob's measurement devices measure correctly in all the bases mentioned in algorithm 1 of [25] and the states are of the actual form. So, Alice and Bob proceed further for the QPQ scheme where Alice measures her qubits using the POVM $D = \{D_0, D_1, D_2\}$.

Now, in this case, whenever Alice gets the outcome D_0, she concludes that Bob's measurement outcome for that instance is $|0\rangle$ and the raw key bit at Bob's side is 0. Similarly, whenever Alice gets the outcome D_1, she concludes that Bob's measurement outcome for that instance is $|1\rangle$ and the raw key bit at Bob's end is 1. However, if Alice gets the outcome D_2, she remains inconclusive about the value of the raw key bit at Bob's side.

We now calculate the success probability of Alice in guessing the raw key bits correctly. Let us assume that $\Pr(D_j||\phi_i\rangle)$ denotes the probability of getting the result D_j whenever the state at Alice's side is $|\phi_i\rangle$ i.e.,

$$\Pr(D_j||\phi_i\rangle) = \langle\phi_i| D_j |\phi_i\rangle$$

This implies that whenever the state at Alice's side is $|\phi_0\rangle$, the success probabilities are

$$\begin{aligned}
\Pr(D_0||\phi_0\rangle) &= \langle\phi_0| D_0 |\phi_0\rangle \\
&= (1-\cos\theta) \\
\Pr(D_1||\phi_0\rangle) &= \langle\phi_0| D_1 |\phi_0\rangle \\
&= 0 \\
\Pr(D_2||\phi_0\rangle) &= \langle\phi_0| D_2 |\phi_0\rangle \\
&= \cos\theta
\end{aligned}$$

Similarly, one can calculate the success probabilities whenever the state at Alice's side is $|\phi_1\rangle$. We formalize all the conditional probabilities in the following table (i.e., Table 1).

Cond. Probability of Alice			
Bob \ Alice	A=D_0	A=D_1	A=D_2
B=$\|\phi_0\rangle$	$1-\cos\theta$	0	$\cos\theta$
B=$\|\phi_1\rangle$	0	$1-\cos\theta$	$\cos\theta$

TABLE 1. Probabilities of Different POVM Outcomes

According to this strategy, whenever Alice gets the outcome $D_0(D_1)$, she concludes that the raw key bit at Bob's side is 0(1). Thus, the success probability of Alice in guessing Bob's i-th raw key bit can be written as

$$\begin{aligned}
& \Pr(R_i = R_{\mathcal{A}_i}) \\
= \; & \Pr(R_i = 0, R_{\mathcal{A}_i} = 0) + \Pr(R_i = 1, R_{\mathcal{A}_i} = 1) \\
= \; & (1-\cos\theta).
\end{aligned}$$

One can check that this success probability (i.e., (1-cos θ)) is always greater than or equal to the actual success probability $\frac{\sin^2\theta}{2}$ that Alice should obtain in case of honest implementation of the QPQ scheme [35]. The superiority of Alice's success probability using the mentioned POVM over the success probability achieved with the actual projective measurement in [35] is illustrated in Figure 2 for different values of θ.

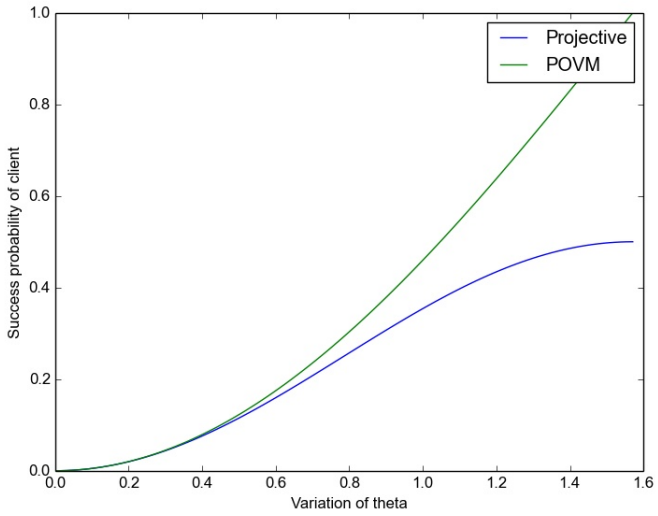

FIGURE 2. Comparison between the success probabilities of getting a raw key bit using projective and POVM measurements

In Theorem 1 and Theorem 2 of the paper [25], it is already discussed how the local test mentioned in [25] certifies the involved devices. As this local test does not involve Alice's measurement device, this will not be able to certify the functionality of the measurement devices of Alice.

That means whenever the shared states are of the form $\frac{1}{\sqrt{2}}(|0\rangle\,|\phi_0\rangle+|1\rangle\,|\phi_1\rangle)$ and Bob's measurement devices measure correctly in all the bases specified in the local test in [25], Alice can successfully pass the testing phase in [25] using any measurement device (other than the actual projective measurement device specified in [35]) at her side. If Alice utilizes the POVM device described in the aforementioned example, she can obtain more raw key bits at her side. Consequently, this enables her to retrieve more data bits in a single query, thereby compromising the database privacy of the protocol [35].

Thus, our proposed attack reveals the vulnerability of the DI proposal in [25] showing that the proposal in [25] actually fails to preserve the data privacy of the QPQ scheme [35]. That's why, a full DI version of the Yang et al. [35] QPQ scheme is necessary to guarantee the database security as well as the user privacy.

3. Full DI proposal for the QPQ scheme [35].

In this section, we describe our proposal for certifying all the devices involved in the QPQ scheme [35]. We split up this entire section into three subsections. In the first subsection, we mention our assumptions that are required for the security of our proposed scheme. In the next subsection, we introduce different steps of our proposed scheme, and in the last subsection, we mention the security related issues of our proposal.

3.1. Assumptions for the full DI scheme:

The assumptions that are required for the security of our proposed full DI version of the QPQ scheme [35] are more or

less the same as the assumptions for the full DI-QPQ proposal in [5] that can be summarized as follows.

1. This proposal assumes that the devices, including the measurement devices and state generation device, involved follow the principles of quantum mechanics and generate outcomes as per the Born rule.
2. In this proposal, we follow a similar assumption as in [9], which is based on the bounded-quantum-storage-model and computational assumptions. Our assumption is that the measurement devices and the state generation devices involved in the scheme are described by a tensor product of Hilbert spaces, one for each device. This means that the devices behave the same in all trials and each use is independent of previous uses. It implies that the statistics for all rounds are independent and identically distributed. Additionally, the inputs chosen by the honest party for each round are also assumed to be random and independent.

   Note. In order to detect any fraudulent behavior by a dishonest party in QPQ, it is reasonable to assume that the input choices made by the honest party are independent and identically distributed. More general scenarios without this assumption may be explored, but are not covered in this work.
3. The only way for an honest party to interact with unknown devices is by making input queries and receiving the outcomes. Before the protocol starts, a dishonest party is allowed to manipulate any devices, but after the protocol starts, he is restricted from manipulating any devices at the honest party's end or opening up his own devices. Then he must interact with the devices in the same way as the honest party. It is also assumed that the dishonest party operates his data in an independent and identically distributed manner.
4. The scenario assumes a distrustful primitive where each party wants to gather as much information as possible while leaking as little of their own information as they can. In each testing phase, the party who wants to detect cheating must act honestly, as if both parties act deceitfully the fraudulent activities will remain undetected.

   For local tests, there is no communication between the laboratories as the honest party randomly selects the input bits for his own device. However, in distributed tests, the honest party selects the input bits for both the parties and communicates the input while the dishonest party generates the corresponding output bits and announces those outputs. The honest party is also assumed to have the ability to shield their devices to prevent any information leakage.

   Note. In distributed tests, the dishonest party may not measure their devices according to the input bits selected by the honest party. Our proposal includes a *device-independent security analysis* that explains how the honest party can detect this dishonest behavior in the corresponding testing phases.
5. In self-tests, the device generating the input bits for one party must be independent and uncorrelated (classically or quantumly) with the devices of the other party. The inputs must be selected freely without any influence from the other systems involved in the protocol.
6. For the QKD-based QPQ schemes, it is already shown in [19] that if the server attempts to retrieve more information about a client's query indices, then there is a risk of providing false information about the intended data bits to the client, which would damage the server's reputation as a database

owner. That's why for the QPQ schemes, it is assumed that the server will not cheat if there exist a non-zero probability of being caught cheating. For this proposal, the server can cheat without being detected because of the underlying computational hiding commitment scheme. But here, we assume that the server has limitations on his computational resources and he is a polynomial time adversary i.e., the server will try at most polynomial time to guess a committed value of the client.

Note. For the QKD-based QPQ schemes, the size of the final key is equal to the size of the database which is usually very large, and the number of raw key bits is even more than that (usually some integer multiple of the number of final key bits). In this situation, it is impractical that the server spends more than the polynomial time to retrieve a raw key bit. For this reason, the polynomial time assumption seems justified here.

3.2. **Proposed full DI version of the scheme** [35]**:** Depending on the functionality, our entire protocol is divided into four phases. The first phase is termed as *Source Device and Bob's Measurement Device Verification Phase*. This phase certifies that the states are of the specified form and Bob's device measures correctly on the specified basis. In this phase, Bob first receives all the states (that will be used for the protocol) from a third party (need not be a trusted one and may collude with the dishonest party) and shares those states with Alice. After that, they check the functionality of the devices in two subphases where at first Bob acts as a referee, chooses some samples randomly, and performs a tilted version of the original CHSH test locally to certify the states and his devices. In the next subphase, Alice also does the same that Bob did in the previous subphase and certifies the states.

After the certification of this source device and Bob's measurement device, they proceed to *Alice's Measurement Device Verification Phase*. This phase certifies the measurement bases of Alice specified in [35]. In this phase, Alice and Bob consider the remaining shared states and perform some measurements assuming their devices as unknown boxes. Then from the outcomes, Alice concludes about the functionality of her measurement device for those specified bases. After successful completion of these two testing phases, Bob and Alice conclude that the states given to them are of the specified form and their measurement devices measure correctly in the bases specified in [35].

The next phase is termed *Key Generation Phase* where Bob generates a key and Alice knows some bits of that key and Bob can not guess the known indices of Alice. The last phase is termed as *private query phase* where Bob encrypts the entire database using the key generated in the previous phase and sends it to Alice. Alice then decrypts the intended bits of the database using her partial knowledge about the final key bits.

Our scheme consists of several steps, which are described below. It should be noted that channel noise is not considered in this description, so it is assumed that all operations are error-free.

Source Device and Bob's Measurement Device Verification Phase

1. Bob starts with $\mathcal{K}$ (we assume here that $\mathcal{K}$ is asymptotically large) number of states (say $|\psi\rangle_{\mathcal{BA}}$) provided by the third party and shares those states with Alice in such a way that the first particle of each state corresponds to Bob and the second particle corresponds to Alice.

Algorithm 1: TiltedCHSH($\mathcal{S}, \mathcal{P}$)

- For every $i \in \mathcal{S}$, $\mathcal{P}$ does the following-
  1. If $y_i = 0$, $\mathcal{P}$'s device applies the measurement operator B_0^0 or B_1^0 randomly on the i-th state's first qubit and generates the output bits $b_i = 0$ and $b_i = 1$ respectively.
  2. If $y_i = 1$, $\mathcal{P}$'s device applies the measurement operator B_0^1 or B_1^1 randomly on the i-th state's first qubit and generates the output bits $b_i = 0$ and $b_i = 1$ respectively.
  3. Similarly, if $x_i = 0$, $\mathcal{P}$'s device applies the measurement operator $A_0^{'0}$ or $A_1^{'0}$ randomly on the i-th state's second qubit and generates the output bits $a_i = 0$ and $a_i = 1$ respectively.
  4. If $x_i = 1$, $\mathcal{P}$'s device applies the measurement operator $A_0^{'1}$ or $A_1^{'1}$ randomly on the i-th state's second qubit and generates the output bits $a_i = 0$ and $a_i = 1$ respectively.
- From these inputs and outputs, the following quantity is estimated by $\mathcal{P}$,

$$\beta_{\mathcal{B}} = \alpha_{\mathcal{B}} \sum_{a\in\{0,1\}} (-1)^a \langle\psi_{\mathcal{BA}}|\mathbb{I} \otimes A_a^{'0}|\psi_{\mathcal{BA}}\rangle + \sum_{x,y,a,b\in\{0,1\}} (-1)^{d_{xyab}} \langle\psi_{\mathcal{BA}}|B_b^y \otimes A_a^{'x}|\psi_{\mathcal{BA}}\rangle$$

  where $\alpha_{\mathcal{B}} = \frac{2}{\sqrt{1+2\tan^2\theta}}$ (for the same θ chosen for the states) and d_{xyab} is defined as follows,

$$d_{xyab} := \begin{cases} 0 & \text{If } xy = a \oplus b \\ 1 & \text{otherwise.} \end{cases}$$

- If $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$ (for the θ chosen for the states) then $\mathcal{P}$ continues with the protocol, otherwise $\mathcal{P}$ aborts the protocol.

  (In the case of honest implementation, this exact desired value of $\beta_{\mathcal{B}}$ and also the exact values of other security parameters in other mentioned algorithms can be obtained for asymptotically large number of samples. However, in practice, with finite number of samples, it is nearly always impossible to exactly match with the desired value of the estimated statistic. Hence, a small deviation from the desired value is allowed in practice. A discussion regarding the variation of the deviation range with the sample size is mentioned in Appendix A. However, how the existing security definitions will vary with the noise parameter, is out of the scope of this present work and we will try to explore this issue in our future works.)

2. Bob chooses $\frac{\gamma_1\mathcal{K}}{2}$ instances randomly from these $\mathcal{K}$ shared states (in practice, how Bob and Alice choose the specific value of γ_1 from the set $[0, 1]$ is mentioned in Appendix A), declares those instances publicly and constructs the set $\Gamma^{\mathcal{B}}_{\text{CHSH}}$ with these chosen instances.
3. For all the instances in $\Gamma^{\mathcal{B}}_{\text{CHSH}}$, Alice sends her qubits to Bob.
4. For the instances in $\Gamma^{\mathcal{B}}_{\text{CHSH}}$, Bob plays the role of the referee as well as the two players and plays TiltedCHSH game.

Algorithm 2: OBStestAlice($\mathcal{S}$)

- Bob has already measured his share of every i-th state of the remaining instances for inputs $y_i = 0$ and $y_i = 1$, and obtained outputs $b_i = 0$ or $b_i = 1$.
- Similarly, Alice has already measured her share of every i-th state of the remaining instances for inputs $x_i = 0$ and $x_i = 1$, obtained outputs $a_i = 0$ or $a_i = 1$, and sent the commitments of those a_i values to Bob.
- For every $i \in \mathcal{S}$, Bob and Alice do the following-
  1. Alice reveals the commitments of a_i values only for the instances chosen in the set $\mathcal{S}$.
  2. Bob then estimates the following quantity from the declared outcomes,
  $$\beta_{\mathcal{A}} = \frac{1}{4} \sum_{x,y,a,b \in \{0,1\}} (-1)^{d'_{xyab}} \alpha_{\mathcal{A}}^{1 \oplus y} \langle \psi_{\mathcal{BA}} | B_b^y \otimes A_a^x | \psi_{\mathcal{BA}} \rangle$$
  where $\alpha_{\mathcal{A}} = \cot\theta$ (for the same θ chosen for the shared states) and d'_{xyab} is as follows,
  $$d'_{xyab} := \begin{cases} 0 & \text{If } xy = a \oplus b \\ 1 & \text{otherwise.} \end{cases}$$
  3. If $\beta_{\mathcal{A}} = \frac{1}{2\sin\theta}$ (for the θ chosen for the shared states) then Bob continue with the protocol, otherwise Bob abort the protocol.

5. For every i-th sample in $\Gamma^{\mathcal{B}}_{\text{CHSH}}$, Bob randomly generates input bits x_i and y_i for his two measurement devices (these devices act as separate parties without any communication), with $x_i, y_i \in \{0, 1\}$.
6. Bob performs TiltedCHSH($\Gamma^{\mathcal{B}}_{\text{CHSH}}$, Bob), according to the procedure outlined in algorithm 1 for the set $\Gamma^{\mathcal{B}}_{\text{CHSH}}$.
7. If Bob passes this TiltedCHSH($\Gamma^{\mathcal{B}}_{\text{CHSH}}$, Bob) test then both Alice and Bob proceed further, otherwise they abort.
8. From the rest $\left(\mathcal{K} - \frac{\gamma_1 \mathcal{K}}{2}\right)$ shared states, Alice randomly chooses $\frac{\gamma_1 \mathcal{K}}{2}$ (in practice, how Bob and Alice choose the specific value of γ_1 from the set $[0, 1]$ is mentioned in Appendix A) instances, declares those instances publicly and constructs the set $\Gamma^{\mathcal{A}}_{\text{CHSH}}$ with these chosen instances.
9. For all the instances in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Bob sends his qubits to Alice.
10. For these instances in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice plays the role of the referee as well as the two players and plays TiltedCHSH game.
11. For every i-th sample in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice randomly generates input bits x_i and y_i for her two measurement devices (these devices act as separate parties without any communication), with $x_i, y_i \in \{0, 1\}$.
12. Alice performs TiltedCHSH($\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice), according to the procedure outlined in algorithm 1 for the set $\Gamma^{\mathcal{A}}_{\text{CHSH}}$.
13. If Alice passes the TiltedCHSH($\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice) test then both Alice and Bob proceed to the next phase where Alice self-tests her measurement device.

Alice's Measurement Device Verification Phase:

1. Alice and Bob consider the rest $(\mathcal{K} - \gamma_1 \mathcal{K})$ states and do the following-

- For every i-th state, Bob randomly generates an input bit $x_i \in_R 0, 1$ for Alice's device and publicly declares all $(\mathcal{K} - \gamma_1\mathcal{K})$ x_i values. After all x_i values are declared, Alice acknowledges receipt to Bob.
- Bob further generates another random bit $y_i \in_R \{0, 1\}$ for every i-th state, as the input of his device.
- If $y_i = 0$, Bob applies measurement operator B_0^0 or B_1^0 randomly on his share of the i-th state and generates the output bit $b_i = 0$ and $b_i = 1$ respectively.
- If $y_i = 1$, Bob applies measurement operator B_0^1 or B_1^1 randomly (here $B_0^1 = B_0^0$ and $B_1^1 = B_1^0$) on his share of the i-th state and generates the output bit $b_i = 0$ and $b_i = 1$ respectively.
- Similarly, if $x_i = 0$, Alice applies measurement operator A_0^0 or A_1^0 randomly on her share of the i-th state and generates the output bit $a_i = 0$ and $a_i = 1$ respectively.
- If $x_i = 1$, Alice applies measurement operator A_0^1 or A_1^1 randomly on her share of the i-th state and generates the output bit $a_i = 0$ and $a_i = 1$ respectively.
- Alice encodes all her a_i values using a *computational hiding perfect binding* commitment scheme (Computationally hiding statistically binding commitment schemes are easy to design from a pseudo-random generator and one-way permutation [26, 30, 31]. As these schemes are perfectly binding, Alice can't cheat at all. For the hiding part, we assume that Bob has limitations on his computational resources and he is a polynomial adversary. That means, we assume that Bob can try at most polynomial time to guess a committed bit value. In the multi-client scenario, it is also possible to use some relativistic bit commitment schemes [24, 10, 12]. However, these are outside the scope of this work.) and send those commitments of a_i values to Bob (The inclusion of a commitment scheme is crucial in this context because here Alice performs a non-optimal projective measurement at her end. This introduces the possibility that she might perform the exact projective measurement during the testing phases and later switch to the optimal POVM measurement discussed in Section 2 for the instances used in the private query phase. To eliminate this possibility, bit commitment is required as it prevents Alice from postponing measurements for any of her particles and ensures that Alice measures all her particles using the actual projective measurement).

2. Bob then chooses $\gamma_2(\mathcal{K} - \gamma_1\mathcal{K})$ instances randomly from these rest $(\mathcal{K} - \gamma_1\mathcal{K})$ instances (in practice, how Bob chooses the specific value of γ_2 form the set $[0, 1]$ is mentioned in Appendix A), constructs a set Γ_{obs} with those chosen instances and declares those instances publicly.
3. Alice reveals the commitments of a_i values for all the instances chosen in the set Γ_{obs}.
4. Bob then performs OBStestAlice(Γ_{obs}), by following the procedure mentioned in algorithm 2, for the set Γ_{obs}.
5. If Alice passes the OBStestAlice(Γ_{obs}) then Bob and Alice proceed to the next phase of the protocol where they generate the raw key bits at their end.

Key Generation Phase:

- Alice and Bob consider the rest $(\mathcal{K} - |\Gamma_{\text{CHSH}}| - |\Gamma_{\text{obs}}|)$ samples and construct a set Γ_{QPQ} with those instances where $|\Gamma_{\text{QPQ}}| = kN$.
- For $1 \leq i \leq (|\Gamma_{\text{QPQ}}|)$, Bob and Alice do the following-
  - If Bob's measurement device generates the outcome $b_i = 0 (b_i = 1)$ for the i-th shared state, Bob considers $R_i = 0 (R_i = 1)$.
  - Alice already knows the a_i values for all these instances. If Alice's measurement device receives the input $x_i = 1$ $(x_i = 0)$ and generates the outcome $a_i = 1$ for her share of the i-th state, Alice considers $R_{\mathcal{A}_i} = 0 (R_{\mathcal{A}_i} = 1)$.
  - If Alice's measurement device receives the input $x_i = 0$ or $x_i = 1$ and generates the outcome $a_i = 0$ for her share of the i-th state, Alice remains inconclusive about the value of the raw key bit indexed by i.

Private Query Phase: Alice and Bob perform the following steps (as mentioned in [35]) for the rest $|\Gamma_{\text{QPQ}}|$ samples.

- Alice and Bob share a kN bit raw key after the *key generation phase*, with Bob having full knowledge of the raw key and Alice knowing some unknown bits (corresponding indices unknown to Bob).
- The raw key is divided into k substrings of length N and a bitwise XOR operation is performed to produce the N bit final key.
- If Alice wants to retrieve the bit indexed by j of the database and knows only the i-th bit F_i of Bob's final key F, she declares the shift number $s = (i - j)$.
- Bob shifts his key F by s positions and encrypts the database using one-time pad.
- The encrypted database can be retrieved by Alice as the j-th database bit is encrypted with F_i (the final key bit indexed by i) known to her.

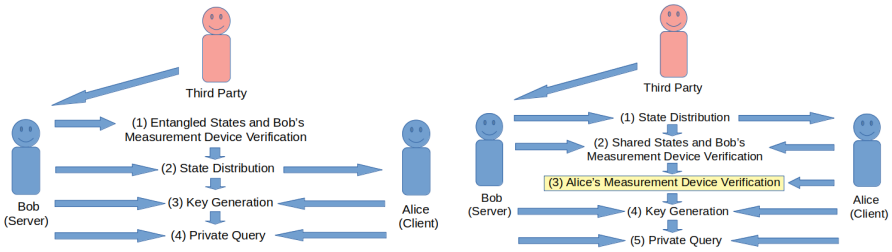

FIGURE 3. Schematic diagrams of the semi DI proposal in [25] (left figure) and our full DI proposal of the QPQ scheme [35] (right figure)

A schematic diagram of the full device independent proposal for the QPQ scheme [35] is shown in the right subfigure of Figure 3.

3.3. **Analysis of our scheme:** Here, we examine the performance of the proposed full DI version of the QPQ scheme [35]. First, we determine the values of relevant parameters. Then, we evaluate the DI security of our proposed scheme. Finally, we assess the security of the database and user in our proposal.

Note that here we present all our analyses considering the asymptotic scenario. In reality, the values of different parameters (derived here) may deviate from their derived value depending on the chosen sample size.

3.3.1. *Estimation of parameters for private query phase:* Here, we perform parameter estimation for maintaining both user and data privacy. In this scheme, after the *key generation phase*, Bob has kN many raw key bits such that Bob knows all the bits but Alice knows only some of those bits. In the *private query phase* , both Bob and Alice cut their raw keys in some particular positions to prepare N substrings of length k such that $k = \frac{|\Gamma_{\text{QPQ}}|}{N}$ where $|\Gamma_{\text{QPQ}}|$ denotes the total number of samples at the *private query phase* and N denotes the number of database bits.

Estimation of θ for improved security : Similar to the QPQ scheme [5], here also the server Bob wants the client Alice to retrieve only one data bit in a single query for database security.

In the QPQ scheme of [35], Alice and Bob share kN raw key bits, with Alice able to retrieve on average $\left(\frac{\sin^2\theta}{2}\right)$ of them. The expected value of the number of raw key bits known by Alice after the *key generation phase* (denoted as n_r here) can be calculated as follows,

$$E[n_r] \approx \left(\frac{\sin^2\theta}{2}\right) kN \tag{6}$$

Alice's probability of correctly guessing a final key bit is roughly $P_f \approx \left(\frac{\sin^2\theta}{2}\right)^k$ since she must correctly guess all k corresponding raw key bits, which are XORed to form the final key bit.

Here, the number of final key bits known by Alice, n_f (let's say), is a binomial random variable with N total bits and a success probability of $P_f = \left(\frac{\sin^2\theta}{2}\right)^k$. So, the expected number of final key bits known by Alice after the *key generation phase* is,

$$E[n_f] = P_f N \approx \left(\frac{\sin^2\theta}{2}\right)^k N \tag{7}$$

In the scheme, dishonest Alice needs to perform correct basis measurements (as specified in [35]) to successfully complete DI testing phases. That means, if the protocol does not abort, the maximum probability of dishonest Alice in guessing R_i (the raw key bit indexed by i) correctly will be atmost $\frac{\sin^2\theta}{2}$ i.e.,

$$\Pr[R_{\mathcal{A}_i^*} = R_i] \leq \frac{\sin^2\theta}{2} \tag{8}$$

where $\mathcal{A}_i^*$ denotes dishonest Alice's subsystem corresponding to the i-th shared state.

It is clear that after Bob's measurement, Alice's states are independent and we assume that the measurement devices at dishonest Alice's side are also independent and memoryless. So, the guessing probability of dishonest Alice for F_i (i.e., the final key bit indexed by i) will be upper bounded by $\left(\frac{\sin^2\theta}{2}\right)^k$ i.e.,

$$\Pr[F_{\mathcal{A}_i^*} = F_i] = P_f \leq \left(\frac{\sin^2\theta}{2}\right)^k \tag{9}$$

Based on the equations 7 and 9, it can be seen that the maximum expected number of final key bits that a dishonest Alice can correctly guess, assuming the protocol does not abort, will be limited to a maximum of,

$$E[F_{\mathcal{A}^*}] \leq \left(\frac{\sin^2\theta}{2}\right)^k N \tag{10}$$

In our scheme, the expected number of data bits correctly guessed by dishonest Alice in a single query is also limited to $\left(\frac{\sin^2\theta}{2}\right)^k N$ as the database is encrypted by XORing with the final key and correctly guessing a final key bit implies correctly guessing a corresponding database bit, provided the protocol does not abort. This implies that,

$$E[D_{\mathcal{A}^*}] \leq \left(\frac{\sin^2\theta}{2}\right)^k N \tag{11}$$

In our scheme, for the protocol to continue, Alice must know at least one final key bit, while Bob wants Alice to know less than two final key bits. Thus, the following condition must be met in the non-abort scenario.

$$1 \leq E[n_f] < 2$$

This implies that,

$$1 \leq \left(\frac{\sin^2\theta}{2}\right)^k N < 2$$

$$\frac{1}{N} \leq \left(\frac{\sin^2\theta}{2}\right)^k < \frac{2}{N} \tag{12}$$

All these results boil down to the following conclusion,

Corollary 3.1. *To ensure that the client Alice only knows less than two final key bits and the proposal doesn't terminate, the server Bob must select the values of θ and the parameter k such that,*

$$\frac{1}{N} \leq \left(\frac{\sin^2\theta}{2}\right)^k < \frac{2}{N}$$

Estimation of P_a and P_c for improved security :

Here, we first determine the likelihood that the protocol will not terminate in an honest scenario. Then using the derived bound on the value of $\sin^2\theta$, we can obtain a lower bound on the value of P_c from the Chernoff-Hoeffding inequality [17] (we estimate the value of P_c using Chernoff-Hoeffding inequality because we consider here the i.i.d scenario).

In our scheme, the probability of Alice not correctly guessing a final key bit is calculated as $\left[1 - \left(\frac{\sin^2\theta}{2}\right)^k\right]$ based on the success probability of Alice in guessing a final key bit, which is $\left(\frac{\sin^2\theta}{2}\right)^k$.

So, the probability that Alice does not know any of the N final key bits is approximately,

$$\left[1 - \left(\frac{\sin^2\theta}{2}\right)^k\right]^N \approx e^{-\left(\frac{\sin^2\theta}{2}\right)^k N} \tag{13}$$

That means the following bound on P_a can be obtained for our proposed scheme.

$$P_a \leq e^{-\left(\frac{\sin^2\theta}{2}\right)^k N} \tag{14}$$

If Bob sets θ such that $\left(\frac{\sin^2\theta}{2}\right)^k = \frac{1}{N}$, then equation 14 gives us the following result according to the relation in equation 12.

$$\boxed{P_a \leq e^{-1}} \tag{15}$$

That means this scheme offers a small P_a value. So, the likelyhood of the proposal not aborting in the honest scenario (i.e. Alice knowing at least one final key bit) is

$$\Pr(\text{protocol doesn't terminate in honest scenario}) \geq \left[1 - e^{-1}\right] \tag{16}$$

So, our proposed scheme has a high probability of not aborting in the honest scenario. We now refer to the Chernoff-Hoeffding inequality in [17].

Proposition 3.2. *(Chernoff-Hoeffding Inequality) Let* $X = \frac{1}{m}\sum_{1\leq i\leq m} X_i$ *be the average of* m *independent random variables* $X_1, X_2, \cdots, X_m$ *with values* $(0, 1)$*, and let* $\mathbb{E}[X] = \frac{1}{m}\sum_{1\leq i\leq m}\mathbb{E}[X_i]$ *be the expected value of* X*. Then for any* $\delta_{CH} > 0$*, we have* $\Pr\left[|X - \mathbb{E}[X]| \geq \delta_{CH}\right] \leq \exp(-2\delta_{CH}^2 m)$.

To derive the bound on P_c, we consider $X_i = 1$ when Alice knows the value of the final key bit indexed by i (or its corresponding data bit) in a non-abort scenario (meaning all raw key bits related to the final key bit indexed by i give either $\left|\phi_0^{\perp}\right\rangle$ or $\left|\phi_1^{\perp}\right\rangle$ as an outcome at Alice's side). If the final key has N many bits, the random variable X is defined as $X = \sum_{i=1}^{N} X_i$.

We have already determined that in the scenario where the proposal doesn't terminate, the expected final key bits known to Alice is $\left(\frac{\sin^2\theta}{2}\right)^k N$ out of a total of N final key bits. To ensure that the number of known final key bits (X) falls within an error margin $\delta_{CH} = \epsilon\left(\frac{\sin^2\theta}{2}\right)^k N$ (where ϵ is a small constant that depends on the number of samples, one may refer to Appendix A for details), we use the Chernoff-Hoeffding inequality. This is because the final key bits are independent and the measurement devices at Alice's end are also independent and memoryless. The calculations of X and $E[X]$ are based on the non-abort scenario. So, we can write the following from the Chernoff-Hoeffding inequality in proposition 3.2.

$$\begin{aligned}&\Pr\left[|X - \mathbb{E}[X]| < \delta_{CH} \wedge \text{protocol doesn't terminate}\right]\\ &\geq 1 - \exp(-2\delta_{CH}^2 m)\end{aligned} \tag{17}$$

After the *key generation phase*, Bob has N final key bits and we want Alice's known final key bits to fall within the range of $[p - \epsilon p, p + \epsilon p]$, where $p = \left(\frac{\sin^2\theta}{2}\right)^k N$ and $\delta_{CH} = \epsilon\left(\frac{\sin^2\theta}{2}\right)^k N$ is the accepted deviation. Plugging in δ_{CH} and $m = N$ into equation 17 gives,

$$\boxed{\begin{aligned}&\Pr\left[|X - \mathbb{E}[X]| < \delta_{CH} \wedge \text{protocol doesn't terminate}\right]\\ &\geq 1 - \exp(-2\delta_{CH}^2 N)\\ &\text{where } \delta_{CH} = \epsilon\left(\frac{\sin^2\theta}{2}\right)^k N\end{aligned}} \tag{18}$$

In equation 12, the following bound is already derived on $\left(\frac{\sin^2\theta}{2}\right)^k$.

$$\frac{1}{N} \leq \left(\frac{\sin^2\theta}{2}\right)^k < \frac{2}{N}$$

By setting $\left(\frac{\sin^2\theta}{2}\right)^k = \frac{1}{N}$ in equation 18, we obtain,

$$\boxed{\begin{aligned}&\Pr\left[|X - \mathbb{E}[X]| < \epsilon \wedge \text{protocol doesn't terminate}\right]\\ &\geq 1 - \exp(-2\epsilon^2 N)\end{aligned}}$$

If the scheme is implemented honestly, the following lower bound of the parameter P_c can be obtained from the definition 1.1 as guessing a final key bit correctly means correctly guessing the corresponding data bit.

$$\boxed{P_c \geq [1 - \exp(-2\epsilon^2 N)]} \tag{19}$$

As in practice, N is large, this probability will be significant. That means, in case of honest implementation of our proposed scheme, the probability that Alice knows the expected number of data bits (with atmost ϵ deviation from the expected number) and the protocol does not abort is high.

The bound on δ_{CH} can be obtained from equation 12 as $\delta_{CH} = \epsilon\left(\frac{\sin^2\theta}{2}\right)^k N$.

$$\epsilon \leq \delta_{CH} < 2\epsilon \tag{20}$$

From this, it's clear that ϵ must satisfy the constraint $2\epsilon \leq 1$, resulting in an upper bound of $\epsilon \leq \frac{1}{2}$. Now we move to discuss the security concerns of our proposed scheme.

3.3.2. *Security in device independent scenario:* In this work, we propose a full DI version of the QPQ scheme [35]. The correctness of this scheme is already mentioned in [25]. Hence, we mention here only the security related issues of our full DI proposal. Based on the results obtained from Theorem 3.3 and Theorem 3.4, here we conclude about the DI security of the QPQ scheme [35].

Theorem 3.3. *(DI testing of shared states and Bob's measurement devices) In the TiltedCHSH test of the source device and Bob's measurement device verification phase, either the devices achieve* $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$ *for both Alice and Bob (i.e., the states provided by the third party are identical with the actual states as mentioned in the QPQ scheme* [35] *and Bob's measurement device measures correctly in the* $\{|0\rangle, |1\rangle\}$ *basis) or the scheme is likely to abort with high probability (as the limit approaches infinity).*

The proof of this theorem exactly follows from the results mentioned in [1] and [4]. We present an outline of this proof in Appendix B.

So, Theorem 3.3 guarantees that either the states shared between Alice and Bob are of the specified form and Bob's measurement device measures correctly in $\{|0\rangle, |1\rangle\}$ basis or the scheme terminates with high likelyhood (as the limit approaches infinity). The next DI testing is done in *Alice's measurement device verification phase*. This phase basically guarantees the functionality of Alice's measurement device. Alice and Bob lead to this phase whenever they successfully pass the first DI testing phase. In this phase, Alice measures in $\{|\phi_0\rangle, |\phi_0^\perp\rangle\}$ or $\{|\phi_1\rangle, |\phi_1^\perp\rangle\}$

basis randomly whereas Bob measures in $\{|0\rangle, |1\rangle\}$ basis. From the measurement outcome, they estimate the value of a parameter $\beta_{\mathcal{A}}$ and check whether this value is equal to $\frac{1}{2\sin\theta}$. Theorem 3.4 guarantees that either Alice's devices measure correctly in the specified basis, resulting in $\beta_{\mathcal{A}} = \frac{1}{2\sin\theta}$, or the protocol will abort with high probability as the limit approaches infinity.

Theorem 3.4 (DI testing of Alice's measurement devices)**.** *In OBStestAlice, either Alice's measurement devices achieve the value of the parameter* $\beta_{\mathcal{A}} = \frac{1}{2\sin\theta}$ *(i.e., her devices correctly measure in* $\{|\phi_0\rangle, |\phi_0^{\perp}\rangle\}$ *and* $\{|\phi_1\rangle, |\phi_1^{\perp}\rangle\}$ *basis) or the protocol terminates with a high likelihood of failure (as the limit approaches infinity).*

The proof of this theorem is explained in detail in Appendix C and follows the same method outlined in [20] for certifying non-maximally incompatible observables.

Note. Here, we claim that if Alice and Bob successfully pass both the TiltedCHSH test and the OBStestAlice mentioned in our full DI proposal, then in the QPQ scheme [35], neither of Alice and Bob can retrieve any additional information in the noiseless scenario. Now, suppose that our claim is wrong i.e., Alice and Bob can pass all the tests mentioned in our scheme and later Alice can retrieve more data bits (than what she intends to know) in a single query or Bob can guess the query indices of Alice with a more certain probability (than his intended probability).

Similar to the analysis in [5], here also we discuss this issue in the context of a particular form of non-i.i.d. attack, where a specific number of states are independently corrupted (more general attacks are also possible but these are outside the scope of this work). In this context, we will show that if some of the corrupted states are included during the testing phases, then there is some probability of being caught as the limit approaches infinity.

At the beginning of our scheme, the untrusted third party provides all the states to the server Bob and then Bob shares those states with Alice. As in the *source device and Bob's measurement device verification phase*, both the parties choose the states randomly from the shared instances for the local tests at their end, the dishonest party can not guess beforehand the shared instances that the honest party will choose at his end for the local test. According to our assumption, the dishonest party can not manipulate the honest party's device once the protocol starts. So, to successfully pass the TiltedCHSH test at the honest party's end, the shared states must be of the actual form as specified in [35]. Similarly, in the TiltedCHSH test performed at Bob's side, the honest Bob must measure the states in the specified basis (to detect the corrupted states) which also certifies the specific measurement bases of Bob. This implies that the *source device and Bob's measurement device verification phase* certifies all the states provided by the untrusted third party and also certifies the measurement device of Bob for the standard basis.

We now explain these things more formally. Let us suppose that initially, the untrusted third party colludes with either the dishonest Alice or the dishonest Bob and shares either $\mathcal{K}_{\mathcal{A}}$ corrupted states in favour of Alice (let us denote this type of states as $\mathcal{A}$-type) or $\mathcal{K}_{\mathcal{B}}$ corrupted states in favour of Bob (let us denote this type of states as $\mathcal{B}$-type) among $\mathcal{K}$ shared states. So, while choosing randomly for the TiltedCHSH test at honest Bob's end, the probability that a chosen state is of $\mathcal{A}$-type is $\frac{\mathcal{K}_{\mathcal{A}}}{\mathcal{K}}$. Similarly, for the TiltedCHSH test at honest Alice's end, the probability that a chosen state is of $\mathcal{B}$-type is $\frac{\mathcal{K}_{\mathcal{B}}}{\mathcal{K}}$. Let us further assume that for the $\mathcal{A}$-type states, the value of the parameter $\beta_{\mathcal{B}}$ is $\beta'_{\mathcal{A}}$ (where $\beta'_{\mathcal{A}} = \beta_{\mathcal{B}} + \epsilon_{\mathcal{A}}$ such

that $\epsilon_{\mathcal{A}} > 0$) and for the $\mathcal{B}$-type states, the value of the parameter $\beta_{\mathcal{B}}$ is $\beta'_{\mathcal{B}}$ (where $\beta'_{\mathcal{B}} = \beta_{\mathcal{B}} + \epsilon_{\mathcal{B}}$ such that $\epsilon_{\mathcal{B}} > 0$).

Now, suppose that only Alice is dishonest and the third party supplies $\mathcal{K}_{\mathcal{A}}$ number of corrupted states (in favour of Alice) along with $(\mathcal{K} - \mathcal{K}_{\mathcal{A}})$ actual states. Then, in the local test at Bob's end, the probability that a chosen state is not of the $\mathcal{A}$-type is $\left(1 - \frac{\mathcal{K}_{\mathcal{A}}}{\mathcal{K}}\right)$. One can easily check that this probability is also same for a chosen state in the final QPQ phase. As, dishonest Alice's aim is to gain as much additional data bits as possible, she needs to choose the value of $\mathcal{K}_{\mathcal{A}}$ such that $(\mathcal{K} - \mathcal{K}_{\mathcal{A}}) = c$ where c is exponentially smaller than $\mathcal{K}$ (i.e., she will try to maximize the probability that a state chosen for the final QPQ phase is of the $\mathcal{A}$ type). Then, the probability that Bob will choose none of the corrupted states (i.e., the $\mathcal{A}$ type states) among his chosen $\frac{\gamma_1 \mathcal{K}}{2}$ states for the TiltedCHSH test at his end is,

$$\left(1 - \frac{\mathcal{K}_{\mathcal{A}}}{\mathcal{K}}\right)^{\frac{\gamma_1 \mathcal{K}}{2}} = \left(\frac{c}{\mathcal{K}}\right)^{\frac{\gamma_1 \mathcal{K}}{2}}$$

which is negligible in $\mathcal{K}$. Similarly, whenever Bob is dishonest, the same thing can be shown for the local TiltedCHSH test at Alice's end. This implies that if the third party colludes with the dishonest party and supplies corrupted states then the probability that none of those corrupted states are chosen for the local test at the honest party's end is negligible.

In our scheme, we consider the ideal scenario where there are no channel noise. So for dishonest Alice, to successfully pass the TiltedCHSH test at the honest Bob's end, the following relation must hold in the noiseless condition.

$$\begin{aligned} \frac{\mathcal{K}_{\mathcal{A}} \beta'_{\mathcal{A}}}{\mathcal{K}} + \frac{(\mathcal{K} - \mathcal{K}_{\mathcal{A}})\beta_{\mathcal{B}}}{\mathcal{K}} &= \beta_{\mathcal{B}} \\ \mathcal{K}_{\mathcal{A}} \beta'_{\mathcal{A}} + (\mathcal{K} - \mathcal{K}_{\mathcal{A}})\beta_{\mathcal{B}} &= \mathcal{K}\beta_{\mathcal{B}} \\ \mathcal{K}_{\mathcal{A}}(\beta'_{\mathcal{A}} - \beta_{\mathcal{B}}) &= 0 \end{aligned}$$

Now, replacing the values of $\beta'_{\mathcal{A}}$ from the relation $\beta'_{\mathcal{A}} = \beta_{\mathcal{B}} + \epsilon_{\mathcal{A}}$, one can get,

$$\mathcal{K}_{\mathcal{A}} \epsilon_{\mathcal{A}} = 0 \tag{21}$$

As the value of $\epsilon_{\mathcal{A}} > 0$, from this relation, one can easily conclude that in the noiseless scenario, the value of $\mathcal{K}_{\mathcal{A}}$ must be zero to successfully pass the local test at the honest Bob's end. Similarly, one can show that whenever Bob is dishonest, the value of $\mathcal{K}_{\mathcal{B}}$ must be zero to successfully pass the local test at the honest Alice's end. In practice, for finite number of samples, one can show that the values of $\mathcal{K}_{\mathcal{A}}$ and $\mathcal{K}_{\mathcal{B}}$ must be negligible to successfully pass the local test at the honest party's end.

In this proposal, we consider a scenario where the shared states are exchanged between the two parties before the start of the protocol, and the dishonest party cannot manipulate the honest party's device after the start of the protocol. As we focus on the *i.i.d.* case, it's clear from the proof of Theorem 3.3 in Appendix B that either the scheme terminates with high likelyhood (as the limit approaches infinity), or the TiltedCHSH test will certify that the shared states in the QPQ scheme [35] reach the desired value of the parameter $\beta_{\mathcal{B}}$.

Similarly, the TiltedCHSH test at the honest Bob's end also confirms that either Bob aborts the scheme with high probability (as the limit approaches infinity), or the TiltedCHSH test at his end certifies that his measurement devices achieve the intended value of the parameter $\beta_{\mathcal{B}}$.

The next DI testing is done in *Alice's measurement device verification phase* where Bob and Alice perform distributed test to certify Alice's projective measurement device. Here, one may think that if Bob is dishonest, then for the instances chosen in *Alice's measurement device verification phase*, he will measure in the actual measurement basis at his end to detect the fraudulent behaviour of Alice, and later for the instances to be used for the actual QPQ phase, he will measure in some different basis to guess the positions of Alice's known key bits.

From the result obtained in Lemma 3.10, it is clear that for Bob to guess Alice's query indices with more certainty, he must reveal more data bits to dishonest Alice in a single query. But doing so violates assumption 4, which states that neither Alice nor Bob leaks more information (from their side) to gain additional knowledge from the other party. Therefore, Bob should act honestly for all the instances in *Alice's measurement device verification phase* as well as in *key generation phase* to ensure the validity of Alice's measurement device, prevent dishonest Alice from obtaining any additional information, and also to maintain his reputation as a database owner. For our proposal, Bob has a chance to cheat because of the inclusion of the computational hiding perfect binding commitment scheme. However, we assume that Bob has limitations on his computational resources and he is a polynomial-time adversary. This assumption bounds Bob to guess a committed bit of Alice. It is also impractical that Bob spends more than the polynomial time to retrieve a particular raw key bit. That's why the computational hiding commitment scheme introduced in our scheme will not leak any additional information to Bob.

As Bob acts honestly for *Alice's measurement device verification phase* and chooses the input bits randomly for both the parties in OBStestAlice, there is no possibility that the inputs for OBStestAlice are chosen according to some dishonest distribution. As the focus of this proposal is on the *i.i.d.* scenario, it can be easily concluded (based on the proof of Theorem 3.4 in Appendix C) that either Alice and Bob will abort the scheme with high likelihood (as the limit approaches infinity), or OBStestAlice will confirm that Alice's measurement devices achieve the intended value of $\beta_{\mathcal{A}}$.

That means we can conclude the following from all these discussions.

Corollary 3.5. *Our DI scheme either terminates with high likelyhood (as the limit approaches infinity) or certifies that the devices in the QPQ scheme* [35] *achieve the desired values of* $\beta_{\mathcal{B}}$ *and* $\beta_{\mathcal{A}}$ *in the TiltedCHSH test and OBStestAlice respectively.*

Given the discussion above on some types of non-i.i.d. attack in our DI proposal, the statement in corollary 3.5 can probably be generalized to some non-i.i.d. cases, but it is outside the scope of this work.

3.3.3. *Security of database against dishonest alice:* Here we estimate the amount of raw key bits that dishonest Alice can guess in the *key generation phase*, and the probability of her retrieving more than the expected data bits in a single query. Dishonest Alice can guess additional raw key bits either from the loophole of the underlying bit commitment scheme or by manipulating the other devices and using an optimal measurement device at her side.

For the underlying computational hiding and perfect binding bit commitment scheme using a pseudo-random generator, the security of the database against dishonest Alice follows from Claim 3.1 in [26] which states that for any i-th committed bit a_i, Alice can fool Bob (i.e., Alice can successfully verify the commitment for a different bit other than the committed one) with probability at most 2^{-n} where n is

the security parameter which is chosen such that no feasible machine can break the underlying pseudorandom generator for seeds of length n. That means, dishonest Alice can't retrieve more raw key bits and if she tries to do so and commits the a_i values obtain from the optimal measurement then it will be detected by Bob during OBStestAlice. More precisely, the security of the entire bit commitment protocol follows from the result mentioned in [26, Theorem 3.1] which states the following.

Corollary 3.6. *If the underlying device G is a pseudorandom generator, then for all polynomials p and large enough security parameter n, the corresponding bit commitment protocol obeys the following.*

- *After commitment, no probabilistic polynomial-time Bob can guess any committed a_i value with probability greater than $\left(\frac{1}{2} + \frac{1}{p(n)}\right)$.*
- *Alice can reveal only the committed bit, except with probability less than 2^{-n}.*

In the case of the manipulation of the other devices and her device, the estimation follows from the DI results in corollary 3.5 which states that after the DI testing phases, either the scheme will abort with high probability (as the limit approaches to infinity) or the devices involved in [35] will meet the intended values of parameters $\beta_{\mathcal{A}}$ and $\beta_{\mathcal{B}}$ as indicated in our proposal.

Theorem 3.7. *In our scheme, in the absence of OBStestAlice, dishonest Alice can retrieve, at most, $\left(\frac{1}{2} + \frac{1}{2}\sin\theta\right)$ fraction of the entire raw key, inconclusively (i.e., the indices of the correctly guessed bits are unknown), during the key generation phase.*

The proof of this Theorem directly follows from the proof of Theorem 5 in [5]. The only difference here is that in this scheme, Alice needs to distinguish between the two non orthogonal quantum states $|\phi_0\rangle$ and $|\phi_1\rangle$ as compared to the two non orthogonal states $|0\rangle$ and $|0'\rangle$ (or $|1\rangle$ and $|1'\rangle$) in [5].

In our full DI proposal, dishonest Alice ($\mathcal{A}^*$) can not perform any other measurement other than the projective measurement mentioned in [35] because if she performs any other measurement at her side then it will be detected in *Alice's measurement device verification phase*. Because of this, we can get a bound on the number of raw key bits that dishonest Alice can retrieve (on average) in this full DI proposal of the QPQ scheme [35].

Lemma 3.8. *Either our protocol terminates with high probability in the long run, or dishonest Alice ($\mathcal{A}^*$) can retrieve (on average) $\frac{\sin^2\theta}{2}$ fraction of bits from the entire raw key after the key generation phase of our scheme.*

Proof. According to the QPQ scheme [35], after the measurements at the server Bob's side, the client Alice has kN independent non-orthogonal qubits at her end. For each of the instances, Alice now tries to distinguish between the non-orthogonal states $|\phi_0\rangle$ and $|\phi_1\rangle$.

From the QPQ scheme [35], it is clear that if Alice measures her qubits in $\{|\phi_0\rangle, |\phi_0^{\perp}\rangle\}$ and $\{|\phi_1\rangle, |\phi_1^{\perp}\rangle\}$ basis randomly, then Alice can guess a raw key bit with certainty whenever the outcome is either $|\phi_0^{\perp}\rangle$ or $|\phi_1^{\perp}\rangle$.

From the correctness of the QPQ scheme [35], it is clear that for each of the instances, the probability of getting the outcome $|\phi_0^{\perp}\rangle$ or $|\phi_1^{\perp}\rangle$ using projective measurement is $\frac{\sin^2\theta}{2}$.

Our DI proposal requires dishonest Alice to independently measure each of the kN qubits at her end in a specified basis to pass the testing phases. If she performs

random and independent projective measurements in the $|\phi_0\rangle, |\phi_0^{\perp}\rangle$ and $|\phi_1\rangle, |\phi_1^{\perp}\rangle$ basis, on average, she can retrieve $\left(\frac{\sin^2\theta}{2}\right) kN$ raw key bits correctly. This concludes the proof. □

In this DI proposal, the database contains N data bits. Although there is a chance that dishonest Alice can successfully pass all tests and learn more data bits than allowed through statistical fluctuations, the likelihood of this happening is low according to Corollary 3.5. Now, based on Definition 1.3 and equation 11, we can conclude the following.

Corollary 3.9. *In the case of dishonest Alice and honest Bob, either the proposed scheme will likely abort (as the limit approaches infinity) or dishonest Alice will, on average, be able to obtain τ fraction of bits from the entire final key, where*

$$\tau \leq \left(\frac{\sin^2\theta}{2}\right)^k \tag{22}$$

By using the upper bound from equation 12 in place of $\left(\frac{\sin^2\theta}{2}\right)^k$, we can obtain the following upper limit for the value of τ.

$$\boxed{\tau < \frac{2}{N}} \tag{23}$$

It shows that our full DI proposal results in τ being significantly smaller than N.

It is possible to validate the data privacy of our scheme in another way (other than the data privacy definition mentioned in Definition 1.3) showing that the probability with which dishonest Alice can successfully guess more than the expected number of final key (or equivalently data) bits (with a deviation more than the ϵ fraction from the expected number) such that the protocol doesn't terminate is low.

Like the discussion in Subsection C 1 (entitled "parameter estimation for private query phase"), here also we assume that the random variable X denotes the number of final key bits known to the dishonest Alice and $E[X]$ be the expected value in honest scenario. The probability $\Pr[|X - \mathbb{E}[X]| > \delta \wedge \text{protocol doesn't terminate}]$ can be shown to be negligible using the properties of basic probability theory. As we consider the i.i.d assumption in our proposal, there will be two different subcases- 1) all the devices attain the ideal TiltedCHSH value, or 2) all the devices do not attain the ideal TiltedCHSH value.

Note that $\Pr[|X - \mathbb{E}[X]| > \delta \wedge \text{protocol doesn't terminate}]$ is upper bounded by both $\Pr[|X - \mathbb{E}[X]| > \delta]$ and $\Pr[\text{protocol doesn't terminate}]$, according to the property of basic probability theory (which says $\Pr[A \wedge B] \leq \Pr[A]$ and $\Pr[A \wedge B] \leq \Pr[B]$).

Now for the first subcase, from the correctness result (i.e., the value of P_c for our scheme in equation 19) and the DI security statement in Theorem 3.4, one can easily conclude that $\Pr[|X - \mathbb{E}[X]| > \delta] \leq negl(N)$.

For the second subcase, by an analysis similar to the proof of Theorem 3.4, it can be concluded that $\Pr[\text{protocol doesn't terminate}] \leq negl(N)$. This implies that for both of these two subcases, $\Pr[|X - \mathbb{E}[X]| > \delta \wedge \text{protocol doesn't terminate}] \leq negl(N)$ (under the i.i.d. assumption).

3.3.4. *Security of user against dishonest bob:*

In this subsection, we determine the number of indices ($l_{\mathcal{B}^*}$) that dishonest Bob can accurately guess from $\mathcal{I}_l$ (the query index set of Alice). Additionally, we calculate the probability of Bob correctly guessing more indices than expected. Generally, for any QKD-based QPQ schemes, if Bob attempts to cheat, there is a risk of providing false information about the intended data bits to Alice, potentially harming his reputation as a database owner [19]. Therefore, for the QPQ primitive, Bob is assumed not to cheat if there is a non-zero probability of being caught. Our scheme provides Bob a chance to cheat without being detected due to the underlying bit commitment scheme. However, we assume that Bob is a polynomial-time adversary and has computational limitations. For this reason, even with the existence of a computational hiding bit commitment scheme, Bob cannot gain any information about Alice's committed bits. So, the calculation here is only based on the results of corollary 3.5, which states that either the scheme terminates with high likelihood or the devices in [35] achieve the desired values of $\beta_{\mathcal{A}}$ and $\beta_{\mathcal{B}}$ after the DI testing phases. Based on these results and those in [5], we can conclude the following.

Lemma 3.10. *Dishonest Bob can correctly guess a maximum of $\frac{l}{N}$ fraction of the indices from the query set $\mathcal{I}_l$ of Alice after l queries to the N-bit database (in the QPQ scheme [35]), i.e, for a particular index i,*

$$\Pr(\textit{Bob correctly guesses } i \in \mathcal{I}_l) \leq \frac{l}{N}$$

Proof. At the *key generation phase* of our proposal, Alice does not broadcast anything about her measurement outcome. So, dishonest Bob has no information about Alice's measurement outcomes and her known key bits. Now, Alice queries l many times to the database and retrieves l many data bits. After these l many queries, dishonest Bob will try to guess those query indices of Alice. As, Bob has no information about the known final key bits of Alice, he has to guess these l many indices (out of the N data bits) randomly.

So, for any i-th data bit, dishonest Bob can guess whether $i \in \mathcal{I}_l$ with probability atmost $\frac{l}{N}$. This completes the proof. □

This implies that Bob can guess whether a database index is in $\mathcal{I}_l$ (the query index set of Alice) with a probability of at most $\frac{l}{N}$. Assuming Alice only knows one data bit per query, if Bob guesses l bits, the expected number of correct guesses Bob can make from Alice's query set $\mathcal{I}_l$ will be,

$$\begin{aligned} E[\mathcal{I}_{\mathcal{B}}] &= \Pr(\text{Bob correctly guesses } i \in \mathcal{I}_l).l \\ &\leq \frac{l^2}{N} \end{aligned} \tag{24}$$

This DI-QPQ proposal includes tests to prevent Bob from discovering too much about $\mathcal{I}_l$ (the query index set of Alice), but due to statistical fluctuations, Bob still has a chance of passing the tests and obtaining more information than a negligible fraction of the indices. As the limit approaches infinity, Bob's likelyhood of passing all the tests becomes low according to Corollary 3.5. Furthermore, if Bob wants to increase the certainty of guessing a query index, he would need to allow Alice to know more data bits (as stated in the result of Lemma 3.10), which goes against assumption 4.

Comparing the expression in definition 1.4 with equation 24 provides the following upper bound for δ in our proposal.

Corollary 3.11. *The DI-QPQ proposal will either abort with high likelyhood (as the limit approaches infinity), or dishonest Bob will be able to correctly guess, on average, δ fraction from $\mathcal{I}_l$ (the query index set of Alice) where,*

$$\delta \leq \left(\frac{l}{N}\right) \tag{25}$$

In practice, the number of data bits in the database, N, is significantly larger than the size of $\mathcal{I}_l$ (i.e., l), with N approximately equal to l^n for some positive integer n. Using this information and equation 25, the following upper bound on the value of δ can be obtained.

$$\boxed{\delta \leq \frac{1}{l^{(n-1)}}} \tag{26}$$

This equation shows that the value of δ is smaller than l in our DI-QPQ proposal.

3.4. **Comparison with the QPQ scheme** [5]**.** Recently, a DI scheme was proposed in [5] addressing the same problem of Quantum Private Query. Here we mention a comparative study between the full DI proposal of the QPQ scheme [35] mentioned in this paper and the full DI-QPQ scheme mentioned in [5].

- **Total number of samples:** In the DI-QPQ scheme mentioned in [5], there are total 6 phases namely entanglement distribution phase, source device verification phase, DI testing phase for Bob's measurement device, DI testing phase for Alice's measurement device, key establishment phase and private query phase. On the other hand, in our proposed DI version of Yang et al. [35] QPQ scheme, there are total 4 phases namely source device and Bob's measurement device verification phase, Alice's measurement device verification phase, key generation phase and private query phase. For consistency and simplicity of comparison, here we consider that each of the protocols starts with N number of samples (i.e., states) and whenever Alice and Bob choose some samples for testing purpose, they choose γ fraction of instances all the time (i.e., for the scheme [5], here we consider $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$). Here we show that if Alice and Bob starts with same number of initial states (i.e., N) for both the protocols and choose γ fraction of samples for all the testing phases (how the specific value of γ is chosen from the set $[0, 1]$ is mentioned in Appendix A), then Alice and Bob can use more number of samples in the private query phase for this proposed full DI version of the QPQ scheme [35] as compared to the number of samples used in private query phase for the DI-QPQ scheme [5].

  So, for the DI-QPQ protocol mentioned in [5], considering $\mathcal{K} = N$ and $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$, Alice and Bob first chooses γN samples for their local CHSHtest which certifies the given states. Next in OBStest, each of Alice and Bob independently chooses $\frac{\gamma}{2}(N - \gamma N)$ samples randomly from the rest $(N - \gamma N)$ states to certify Bob's measurement device. So, the total number of samples used in the OBStest is $\gamma(N - \gamma N)$. Next in the DI testing phase for Alice's measurement device, Alice chooses γ fraction of samples randomly from the rest $(1 - \gamma)(N - \gamma N)$ samples to certify her measurement device. Atlast, the rest $[(1 - \gamma)(N - \gamma N) - \gamma(1 - \gamma)(N - \gamma N)] = (1 - \gamma)^3 N$ samples are used for private query phase. This implies that in the DI-QPQ scheme [5], the server Bob can generate a raw key of length $(1-\gamma)^3 N$ bits using N number

of states. So, if we consider that in the QPQ scheme [5], Alice and Bob use F_{old} fraction of initial samples for private query phase then $F_{old} = (1-\gamma)^3$.

Similarly, for the full DI version of the QPQ scheme [35] mentioned here, considering $\mathcal{K} = N$ and $\gamma_1 = \gamma_2 = \gamma$, Bob and Alice first chooses γN samples randomly for their local TiltedCHSH test which certifies the given states and Bob's measurement device. Each of Alice and Bob then chooses $\frac{\gamma(N-\gamma N)}{2}$ samples randomly from the rest $(N-\gamma N)$ states for OBStestAlice which certifies Alice's measurement device. Atlast, the rest $(N-\gamma N)-\gamma(N-\gamma N) = (1-\gamma)^2 N$ samples are used for key generation. This implies that in the full DI version of the QPQ scheme [35] mentioned here, the server Bob can generate $(1-\gamma)^2 N$ raw key bits using N number of states. So, if we consider that in this scheme, Alice and Bob use F_{new} fraction of initial samples for private query phase then $F_{new} = (1-\gamma)^2$.

A comparative study between the number of samples used for raw key generation in two different protocols for different values of γ is shown in Figure 4. From this figure, it is clear that for any value of γ (where $\gamma \in (0,1)$), the size of the raw key generated in the proposed full DI version of the QPQ scheme [35] is always greater than the size of the raw key generated in the DI-QPQ scheme [5]. This implies that to generate a raw key of a particular size, the DI-QPQ scheme mentioned in [5] requires more number of initial samples as compared to the full DI version of the QPQ scheme [35] mentioned here. So, in terms of the total number of samples, this full DI version of the QPQ scheme [35] is more efficient as compared to the DI-QPQ scheme [5].

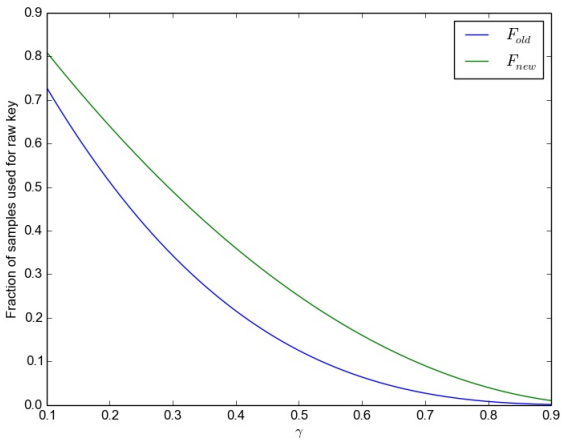

FIGURE 4. Comparison between the fraction of samples used for raw key generation in two different protocols for different values of γ

- **Projective measurement Vs. POVM:** The DI-QPQ schemes in this paper and in [5] are QKD-based, using non-orthogonal state distinction for key generation. In [5], the client uses POVM measurements to distinguish non-orthogonal states, while this paper's full DI version of the Yang et al. QPQ scheme uses projective measurements.

Although it is well-known that POVM measurements can be implemented as a projective measurements in the higher dimension, it requires additional gates as compared to the projective measurements in actual dimension. This implies that the implementation of POVM measurement is complicated as compared to the projective measurement. So, from the viewpoint of practical implementation of the measurement devices, the full DI version of Yang et al. scheme mentioned here is more efficient as compared to the DI-QPQ scheme mentioned in [5].

- It is well-known that the maximally entangled states are easy to prepare as compared to the non-maximally entangled states. The DI-QPQ scheme mentioned in [5] uses maximally entangled states whereas the full DI version of the QPQ scheme [35] mentioned here uses non-maximally entangled states. So, from the viewpoint of practical implementation of the source device, the DI-QPQ scheme mentioned in [5] is more efficient as compared to the DI-QPQ schemes mentioned in this work.

4. **Full DI proposal for a modified version of the QPQ scheme** [35]**.** From the analysis of section 2, it is clear that for the QPQ scheme [35], the client Alice can retrieve more number of database bits in a single query, if she performs optimal POVM measurement at her side instead of the projective measurements mentioned in [35]. In this direction, here we propose a full DI protocol for a modified version of [35] where the client Alice can retrieve optimal number of raw key bits at her end.

We divide this entire section into two subsections. In the first subsection, we propose different steps of our modified DI-QPQ scheme and in the last subsection, we mention the security related issues of this modified proposal. The assumptions for this modified DI-QPQ scheme is also same as the assumptions of our previous full DI-QPQ proposal (which are mentioned in subsection III A).

4.1. **Modified full DI protocol.** Like the previous DI proposal, here also we divide the entire protocol into four phases based on the functionality. The first phase which certifies the state generation device and Bob's measurement device is termed *Source Device and Bob's Measurement Device Verification Phase*. The next phase certifies the measurement devices for the client Alice and is termed *Alice's Measurement Device Verification Phase*. After successful completion of these two testing phases, Bob and Alice conclude that the states given to them are of the specified form and their measurement devices measure correctly in the specified bases (here 'specified' refers to the state and measurement bases mentioned in this modified QPQ proposal). After these testing phases, Bob and Alice proceed to the *Key Generation Phase* where Bob generates a key and Alice knows some bits of that key such that Bob can not know anything about Alice's known key bits. At last, they proceed to the *private query phase* where Bob encrypts the entire database using the key generated in the key generation phase and sends it to Alice. Alice then decrypts the intended bits of the database using her partial knowledge about the final key bits.

Now we describe different steps of our entire protocol. Note that like our previous scheme, here also we consider that there is no channel noise i.e., all the operations are perfect.

Algorithm 3: KeyGenAlice($\mathcal{S}$)

- For each index $i \in \mathcal{S}$, Alice performs the following steps-
  1. Alice uses the measurement device $D = \{D_0, D_1, D_2\}$ to measure her qubit of the shared state indexed by i.
  2. Alice concludes the raw key bit indexed by i as 0(1) if she gets the measurement outcome $D_0(D_1)$ for the shared state indexed by i.
  3. Alice remains uncertain about the raw key bit indexed by i if she gets the measurement outcome D_2 for the shared state indexed by i.

Algorithm 4: POVMtestAlice($\mathcal{S}$)

- For each index $i \in \mathcal{S}$, Bob and Alice perform the following steps-
  1. Bob first declares the value of R_i (i.e., the raw key bit indexed by i).
  2. Whenever $R_i = 0$ ($R_i = 1$), Alice considers that the state at her side is ρ_0 (ρ_1).
- Alice then computes the parameter

$$\Omega = \sum_{R_i, R_{\mathcal{A}_i} \in \{0,1\}} (-1)^{R_i \oplus R_{\mathcal{A}_i}} \mathrm{Tr}[D_{R_{\mathcal{A}_i}} \rho_{R_i}]$$

  where $D_{R_{\mathcal{A}_i}}$ is Alice's measurement outcome in KeyGenAlice() for the i-th instance.
- If for the set $\mathcal{S}$,

$$\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$$

  then they continue with the protocol, otherwise they abort.

Source Device and Bob's Measurement Device Verification Phase:

1. Bob starts with $\mathcal{K}$ (we assume here that $\mathcal{K}$ is asymptotically large) number of states (say $|\psi\rangle_{\mathcal{BA}}$) provided by the third party and shares those states with Alice in such a way that the first particle of each state corresponds to Bob and the second particle corresponds to Alice.
2. Bob chooses $\frac{\gamma_1 \mathcal{K}}{2}$ instances randomly from these $\mathcal{K}$ shared states (in practice, how Bob and Alice choose the specific value of γ_1 from the set $[0, 1]$ is mentioned in Appendix A), declares those instances publicly and constructs the set $\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$ with these chosen instances.
3. For all the instances in $\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$, Alice sends her qubits to Bob.
4. For the instances in $\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$, Bob plays the role of the referee as well as the two players and plays TiltedCHSH game.
5. For every i-th sample in $\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$, Bob randomly generates input bits x_i and y_i for his two measurement devices (these devices act as separate parties without any communication), with $x_i, y_i \in 0, 1$.
6. Bob performs TiltedCHSH($\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$, Bob), according to the procedure outlined in algorithm 1 for the set $\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$.
7. If Bob passes this TiltedCHSH($\Gamma^{\mathcal{B}}_{\mathrm{CHSH}}$, Bob) test then both Alice and Bob proceed further, otherwise they abort.

8. From the rest $\left(\mathcal{K} - \frac{\gamma_1 \mathcal{K}}{2}\right)$ shared states, Alice randomly chooses $\frac{\gamma_1 \mathcal{K}}{2}$ (in practice, how Bob and Alice choose the specific value of γ_1 from the set $[0, 1]$ is mentioned in Appendix A) instances, declares those instances publicly and constructs the set $\Gamma^{\mathcal{A}}_{\text{CHSH}}$ with these chosen instances.
9. For all the instances in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Bob sends his qubits to Alice.
10. For these instances in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice plays the role of the referee as well as the two players and plays TiltedCHSH game.
11. For every i-th sample in $\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice randomly generates input bits x_i and y_i for her two measurement devices (these devices act as separate parties without any communication), with $x_i, y_i \in 0, 1$.
12. Alice performs TiltedCHSH($\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice), according to the procedure outlined in algorithm 1 for the set $\Gamma^{\mathcal{A}}_{\text{CHSH}}$.
13. If Alice passes the TiltedCHSH($\Gamma^{\mathcal{A}}_{\text{CHSH}}$, Alice) test then both Alice and Bob proceed to the next phase where Alice self-tests her measurement device.

Alice's Measurement Device Verification Phase:

- Alice and Bob consider the rest $(\mathcal{K} - \gamma_1 \mathcal{K})$ samples and construct a set Γ_{test}
- For $1 \leq i \leq |\Gamma_{\text{test}}|$, Bob does the following-
  - Bob applies measurement operator B^0_0 or B^0_1 randomly on his particle of the shared state indexed by i and generates the output bit $b_i = 0$ and $b_i = 1$ respectively.
  - If the outcome of Bob's device for the shared state indexed by i is $b_i = 0$, Bob considers the raw key bit indexed by i as $R_i = 0$.
  - If the outcome of Bob's device for the shared state indexed by i is $b_i = 1$, Bob considers the raw key bit indexed by i as $R_i = 1$.
- Alice chooses $\gamma_2 |\Gamma_{\text{test}}|$ instances (in practice, how Alice chooses the specific value of γ_2 from the set $[0, 1]$ is mentioned in Appendix A) randomly from these $|\Gamma_{\text{test}}|$ states, constructs a set Γ_{POVM} with those samples and declares those instances (Note that no commitment scheme is required here like our previous proposal as in this modified scheme, Alice is performing optimal individual measurements at her end. So, Alice can't retrieve any additional bits in the *key generation phase* by performing any other measurements. Alice can at most perform joint measurements to retrieve the final key bits instead of the individual raw key bits. However, these optimal joint measurements are already shown to be inconclusive [19, 5] and are of no use to Alice).
- Alice first performs KeyGenAlice(Γ_{POVM}), according to the procedure introduced in algorithm 3 for the set Γ_{POVM}.
- Bob and Alice then perform POVMtestAlice(Γ_{POVM}), according to the procedure introduced in algorithm 4 for the same set Γ_{POVM}.
- If Alice and Bob pass the POVMtestAlice(Γ_{POVM}) then they proceed to the next phase of the protocol where they generate the shared key.

Key generation phase:

- Alice and Bob consider the rest $(|\Gamma_{\text{test}}| - |\Gamma_{\text{POVM}}|)$ samples, construct a set Γ_{QPQ} with those instances and do the following-
  1. Alice performs KeyGenAlice(Γ_{QPQ}), as mentioned in algorithm 3 for the set Γ_{QPQ}.
  2. Bob already generates the raw key bits for each of the instances in Γ_{QPQ}.

Private query phase:

- Alice and Bob then use classical methods to process the raw key and move to the private query phase described in [35] (detailed procedure is already mentioned in the previous proposal of this paper).

A visual illustration of different steps of this full device-independent proposal for a modification of the QPQ scheme [35] is depicted in Figure 5.

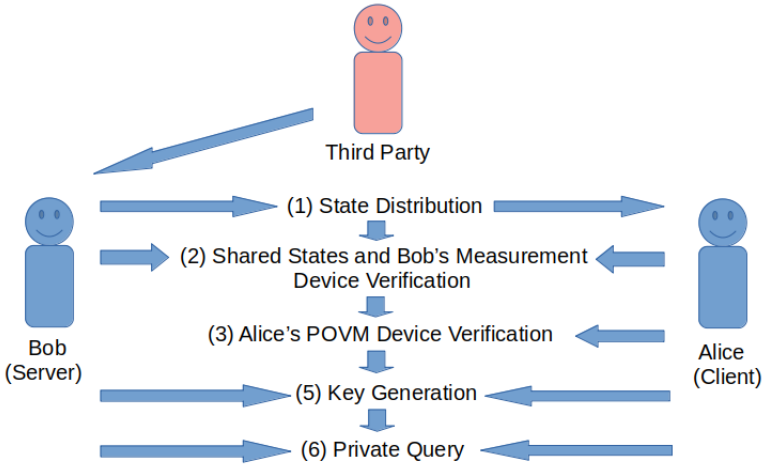

FIGURE 5. Visual representation of our modified DI-QPQ scheme

4.2. **Analysis of the modified scheme:** Here, we address the functionality of this proposal. At first, we prove its correctness, and next, we discuss the security aspects.

4.2.1. *Correctness of our modified scheme:* First, we prove the correctness of this modified scheme.

Theorem 4.1. *If the modified proposal is implemented honestly, then after the key generation phase, Alice is able to retrieve only* $(1-\cos\theta)$ *fraction of the entire raw key.*

Proof. After the *key generation phase*, Bob and Alice share $|\Gamma_{\mathrm{QPQ}}|$ raw key bits. These raw key bits were generated from $|\Gamma_{\mathrm{QPQ}}|$ copies of shared entangled states which are of the form

$$\frac{1}{\sqrt{2}}(|0\rangle\,|\phi_0\rangle + |1\rangle\,|\phi_1\rangle)$$

where, $|\phi_0\rangle = \cos\frac{\theta}{2}\,|0\rangle + \sin\frac{\theta}{2}\,|1\rangle$ and $|\phi_1\rangle = \cos\frac{\theta}{2}\,|0\rangle - \sin\frac{\theta}{2}\,|1\rangle$. Here θ may vary from 0 to $\frac{\pi}{2}$.

Bob and Alice generate these $|\Gamma_{\mathrm{QPQ}}|$ many raw key bits as follows-

For each of the states in the set Γ_{QPQ}, Bob measures his qubits in $\{|0\rangle, |1\rangle\}$ basis. For any i-th instance, if Bob receives the outcome $|0\rangle$, he considers $R_i = 0$ and $R_i = 1$ otherwise (i.e., for outcome $|1\rangle$).

Now, Alice understands that after Bob's measurement, her qubits corresponding to each of the shared states collapse to either $|\phi_0\rangle$ or $|\phi_1\rangle$. However, to obtain the value of the raw key bit, Alice has to distinguish these two states conclusively. As,

$|\phi_0\rangle$ and $|\phi_1\rangle$ are non-orthogonal states (when $\theta \neq \frac{\pi}{2}$), Alice cannot distinguish these two states with certainty.

According to the strategy mentioned in this modified protocol, Alice chooses the POVM $\{D_0, D_1, D_2\}$ for measurement. After measurement, if Alice receives the outcome D_0 for i-th instance, she concludes that Bob's corresponding measurement outcome was $|0\rangle$. In such case, Alice concludes that $R_{\mathcal{A}_i} = 0$. Similarly, if Alice receives the outcome D_1 for i-th instance, she concludes that Bob's corresponding measurement outcome was $|1\rangle$. In such a case, Alice concludes that $R_{\mathcal{A}_i} = 1$. However, if the measurement outcome is D_2, then Alice remains inconclusive about the value of the raw key bit.

Now, we calculate the success probability of Alice in guessing each R_i correctly. Let us assume that $\Pr(D_j||\phi_i\rangle)$ denotes the corresponding success probability of getting the result D_j when the given state is $|\phi_i\rangle$ i.e.,

$$\Pr(D_j||\phi_i\rangle) = \langle\phi_i| D_j |\phi_i\rangle$$

We now calculate the corresponding success probabilities of getting different results for the states $|\phi_0\rangle$ and $|\phi_1\rangle$.

For $|\phi_0\rangle$, the success probabilities will be

$$\begin{aligned}
\Pr(D_0||\phi_0\rangle) &= \langle\phi_0| D_0 |\phi_0\rangle \\
&= (1 - \cos\theta) \\
\Pr(D_1||\phi_0\rangle) &= \langle\phi_0| D_1 |\phi_0\rangle \\
&= 0 \\
\Pr(D_2||\phi_0\rangle) &= \langle\phi_0| D_2 |\phi_0\rangle \\
&= \cos\theta
\end{aligned}$$

Similarly, for the state $|\phi_1\rangle$, the success probabilities will be

$$\begin{aligned}
\Pr(D_0||\phi_1\rangle) &= \langle\phi_1| D_0 |\phi_1\rangle \\
&= 0 \\
\Pr(D_1||\phi_1\rangle) &= \langle\phi_1| D_1 |\phi_1\rangle \\
&= (1 - \cos\theta) \\
\Pr(D_2||\phi_0\rangle) &= \langle\phi_1| D_2 |\phi_1\rangle \\
&= \cos\theta
\end{aligned}$$

We formalize all the conditional probabilities in table 1. Thus, the success probability of Alice in guessing R_i of Bob can be written as

$$\begin{aligned}
& \Pr(R_{\mathcal{A}_i} = R_i) \\
= \; & \Pr(R_{\mathcal{A}_i} = 0, R_i = 0) + \Pr(R_{\mathcal{A}_i} = 1, R_i = 1) \\
= \; & (1 - \cos\theta).
\end{aligned}$$

So, the success rate of Alice in guessing each bit of Bob's raw key in this modified proposal's key generation phase is $(1 - \cos\theta)$, meaning she can determine with certainty the positions of the correctly guessed bits and retrieve an average of $(1 - \cos\theta)$ fraction of bits from the entire raw key. □

4.2.2. *Estimation of parameters for private query phase:* Considering the honest implementation of this modified scheme, here we determine the values for different parameters to ensure both user privacy and data privacy.

Estimation of the parameter θ for security purpose: Like our previous full DI version of the QPQ scheme [35], here also the server Bob wants that for database security, the client Alice should not know more than one final key bit. In this modified proposal, the server Bob has a raw key with kN many bits and the client Alice can correctly guess each of those bits with probability around $(1 - \cos\theta)$. So, the expected number of raw key bits that Alice can know in $(1 - \cos\theta)kN$.

Then each of Alice and Bob XOR k number of raw key bits to generate every bit of the final key at their end. So, Alice can correctly guess every bit of Bob's final key with probability around $(1 - \cos\theta)^k$.

Now, if $F_{\mathcal{A}}$ denotes the number of final key bits known to Alice then we can conclude that the expected value of $F_{\mathcal{A}}$ will be,

$$E[F_{\mathcal{A}}] \approx (1 - \cos\theta)^k N \tag{27}$$

In this modified DI scheme, for dishonest Alice to pass DI testing phases, she must measure correctly for all instances. Moreover, it is known that the optimal probability in distinguishing two non orthogonal states is $(1 - \cos\theta)$, which means dishonest Alice's probability of correctly guessing a raw key bit and a final key bit without causing the protocol to terminate is capped at $(1-\cos\theta)$ and $(1-\cos\theta)^k$, respectively. That means, when the protocol doesn't terminate, the expected number of correctly guessed final key bits by dishonest Alice is at most limited by,

$$E[F_{\mathcal{A}^*}] \leq (1 - \cos\theta)^k N \tag{28}$$

Like the Yang et al [35] QPQ scheme, here also the database is encrypted with the final key by performing bitwise XOR. Hence, in non abort scenario, the expected maximum number of correctly guessed data bits by dishonest Alice in a single query is limited to $(1 - \cos\theta)^k N$. i.e.,

$$E[D_{\mathcal{A}^*}] \leq (1 - \cos\theta)^k N \tag{29}$$

Now, like the previous proposal, here also for the protocol to continue, Alice must know atleast one final key bit, while Bob wants Alice to know less than two final key bits i.e.,

$$1 \leq E[F_{\mathcal{A}}] < 2$$

This implies that,

$$1 \leq (1 - \cos\theta)^k N < 2$$

$$\frac{1}{N} \leq (1 - \cos\theta)^k < \frac{2}{N} \tag{30}$$

These results boil down to the following conclusion.

Corollary 4.2. *To ensure that the client Alice only knows less than two final key bits and the protocol doesn't terminate in this modified proposal, the server Bob must select the values of θ and the parameter k such that,*

$$\frac{1}{N} \leq (1 - \cos\theta)^k < \frac{2}{N}$$

Estimation of the security parameter P_a and P_c : Proceeding to the similar way as discussed in corollary 3.1, here we can assert that Alice can't guess any final key bit with probability

$$\begin{aligned}\Pr(\text{the protocol aborts}) &\approx [1-(1-\cos\theta)^k]^N \\ &\approx e^{-(1-\cos\theta)^k N} \end{aligned} \tag{31}$$

So, for the parameter P_a, we get the following upper bound for this modified scheme.

$$P_a \leq e^{-(1-\cos\theta)^k N} \tag{32}$$

If Bob sets θ so that $(1-\cos\theta)^k = \frac{1}{N}$, then equation 30 and 32 yield

$$\boxed{P_a \leq e^{-1}} \tag{33}$$

This implies that this modified proposal has a small P_a value. So, the probability of the protocol not aborting in the honest scenario is,

$$\begin{aligned}&\Pr(\text{protocol doesn't terminate in honest scenario}) \\ &\geq (1-e^{-1}) \end{aligned} \tag{34}$$

Hence, this modified proposal has a high probability of not aborting in the honest scenario.

Like the previous scheme, here also (proceeding to the similar way) one can achieve the below mentioned bound on P_c for this modified scheme.

$$\boxed{P_c \geq [1-\exp(-2\epsilon^2 N)]} \tag{35}$$

where $\epsilon \leq \frac{1}{2}$ for security purpose.

We now proceed to the security aspects of this modified proposal.

4.2.3. *Security in device independent scenario:* In this subsection, we discuss about the DI security of this modified QPQ proposal. Based on the results obtained from Theorem 4.3 and Theorem 4.4, here we conclude about the DI security of this modified QPQ scheme.

Theorem 4.3 (DI testing of shared states and Bob's measurement devices)**.** *In the TiltedCHSH test of the source device and Bob's measurement device verification phaseof our modified proposal, either the devices achieve $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$ for both Alice and Bob (i.e., the states provided by the third party are identical with the actual states and Bob's measurement device measures correctly in the $\{|0\rangle, |1\rangle\}$ basis) or the scheme is likely to abort with high probability (as the limit approaches infinity).*

Proof. This proof is same as the proof of theorem 3.3. □

So, Theorem 4.3 guarantees that either the states shared between Alice and Bob are of the specified form and Bob's measurement device measures correctly in $\{|0\rangle, |1\rangle\}$ basis or this modified scheme aborts with high likelyhood in the long run. The next testing for full DI certification is done in *Alice's measurement device verification phase.* This phase basically guarantees the functionality of Alice's POVM device. They lead to this phase whenever both of them successfully pass the first DI testing phase. In this phase, Alice performs the POVM measurement $D = \{D_0, D_1, D_2\}$ on the chosen states. From the measurement outcome, Alice computes the value of the parameter Ω and checks whether this value is equal to

$\frac{2\sin^2\theta}{(1+\cos\theta)}$. Theorem 4.4 guarantees that either Alice measures correctly using the measurement device $\{D_0, D_1, D_2\}$ (i.e., the devices achieve $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$) or this modified proposal terminates with high probability (as the limit approaches infinity).

Theorem 4.4 (DI Testing of Alice's POVM D)**.** *POVMtestAlice either results in a high probability of termination of this modified proposal (as the limit approaches infinity), or it guarantees that Alice's measurement devices attain* $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$*, meaning they are of this specified form (up to a local unitary),*

$$D_0 = \frac{1}{(1+\cos\theta)}(\left|\phi_1^{\perp}\right\rangle\left\langle\phi_1^{\perp}\right|)$$
$$D_1 = \frac{1}{(1+\cos\theta)}(\left|\phi_0^{\perp}\right\rangle\left\langle\phi_0^{\perp}\right|)$$
$$D_2 = \mathbb{I} - D_0 - D_1,$$

where $\left|\phi_1^{\perp}\right\rangle = \left(\sin\frac{\theta}{2}\left|0\right\rangle + \cos\frac{\theta}{2}\left|1\right\rangle\right)$ *and* $\left|\phi_0^{\perp}\right\rangle = \left(\sin\frac{\theta}{2}\left|0\right\rangle - \cos\frac{\theta}{2}\left|1\right\rangle\right)$.

The detailed proof of this theorem is mentioned in Appendix D. In the proof, we consider a general form of a single qubit three outcome POVM $\{D_0, D_1, D_2\}$ and show that if the input states are chosen randomly between $\left|\phi_0\right\rangle = (\cos\frac{\theta}{2}\left|0\right\rangle + \sin\frac{\theta}{2}\left|1\right\rangle)$ and $\left|\phi_1\right\rangle = (\cos\frac{\theta}{2}\left|0\right\rangle - \sin\frac{\theta}{2}\left|1\right\rangle)$ then either $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$ i.e., $\{D_0, D_1, D_2\}$ are of the specified form as mentioned in POVMtestAlice or this modified proposal terminates with high likelyhood(as the limit approaches infinity).

Note that in our proof, we have not imposed any dimension bound like the self-testing of POVM in a prepare and measure scenario in [34]. So, the devices that perform a Neumark dilation of this mentioned POVM (i.e., the equivalent larger projective measurement on both the original state and some ancilla system instead of the actual POVM measurement) could still achieve the intended value of Ω. But both of these operations produce the same output probabilities, which is sufficient for the purposes of this work.

Like the previous full DI proposal of the QPQ scheme [35], here also one can argue in a similar way that *this modified scheme either terminates with high probability (as the limit approaches infinity) or it certifies that the devices in this modified QPQ proposal achieve the desired values of the parameters* $\beta_{\mathcal{B}}$ *and* Ω *in the TiltedCHSH test and POVMtestAlice respectively.*

4.2.4. *Security of database against dishonest alice:* Here, we estimate the amount of raw key bits guessed by dishonest Alice during the *key generation phase* of this modified scheme. Similar to the result in Theorem 3.7, here also we can conclude the following-

Theorem 4.5. *For this modified DI-QPQ scheme, in the absence of POVMtestAlice, dishonest Alice can retrieve, at most,* $\left(\frac{1}{2} + \frac{1}{2}\sin\theta\right)$ *fraction of the entire raw key, inconclusively (i.e., the indices of the correctly guessed bits are unknown), during the key generation phase.*

The proof exactly follows from the proof of Theorem 5 in [5].

In this modified DI-QPQ proposal, Alice performs a particular POVM measurement to distinguish the non-orthogonal states at her end which is also the optimal measurement to distinguish that specified non-orthogonal states. Because of this

specific measurement, we can get a bound on the number of raw key bits guessed (on average) by dishonest Alice in this proposed scheme.

Lemma 4.6. *Either our modified protocol terminates with high probability in the long run, or dishonest Alice ($\mathcal{A}^*$) can retrieve (on average) $(1-\cos\theta)$ fraction from the entire raw key after the key generation phase of this modified scheme.*

The proof of this Lemma is based on the Theorem 4.1 which establishes the correctness of this modified scheme.

One can also note that this $(1-\cos\theta)$ is the optimal probability (this optimality is proven in [18]) of success in distinguishing two non-orthogonal states with certainty (which is the main objective of the client Alice here in this modfied proposal).

Now, proceeding to the similar way as mentioned in corollary 3.9, the bounds on τ and P_d can be achieved for this modified scheme. At first, equation 29 and definition 1.3 yield the following bound on τ.

Corollary 4.7. *In the case of dishonest Alice and honest Bob, either this modified proposal will likely abort (as the limit approaches infinity), or dishonest Alice will, on average, be able to obtain τ fraction of bits from the entire final key, where*

$$\tau \leq (1-\cos\theta)^k \tag{36}$$

By using the upper bound from equation 30 in place of $(1-\cos\theta)^k$, we can obtain the following upper limit for the value of τ.

$$\boxed{\tau < \frac{2}{N}} \tag{37}$$

It shows that this modified proposal results in τ being significantly smaller than N.

4.2.5. *Security of user against dishonest bob:*

In this subsection, we estimate the number indices that dishonest Bob can correctly guess from $\mathcal{I}_l$ (the query index set of Alice) after successfully passess the *key generation phase* of this modified scheme. Similar to the result in Lemma 3.10, here also we can conclude the following-

Lemma 4.8. *Dishonest Bob can correctly guess a maximum of $\frac{l}{N}$ fraction of the indices from the query set $\mathcal{I}_l$ of Alice after l queries to the N-bit database (in this modified proposal), i.e, for a particular index i,*

$$\Pr(\textit{Bob correctly guesses } i \in \mathcal{I}_l) \leq \frac{l}{N}$$

The proof of Lemma 4.8 is identical to the proof of Lemma 3.10.

Like the discussion in corollary 3.11, bounds on δ and P_u can also be obtained for this modified proposal.

Corollary 4.9. *In dishonest Bob and honest Alice scenario of this modified DI-QPQ proposal, the scheme will either abort with high likelyhood (as the limit approaches infinity), or dishonest Bob will be able to correctly guess, on average, δ fraction of indices from $\mathcal{I}_l$ (the query index set of Alice) where,*

$$\boxed{\delta \leq \frac{1}{l^{(n-1)}}} \tag{38}$$

where n is a positive integer such that $n > 1$. From this relation, one can conclude that δ is smaller than l for this modified proposal.

5. **Discussion and conclusion.** Like most of the initial quantum cryptography schemes, the security of the initial QPQ schemes also relies on the functionality of the involved devices. Later, it was shown that if those devices don't work accordingly then some information may leak to the adversary. Recently, Maitra et al. [25] first identified this loophole for the QPQ scheme [35] and suggested a tilted version of the local CHSH test on top of the QPQ scheme [35] to certify the functionality of the devices. However, their proposed local test at the server-side does not certify the functionality of the client's measurement device. Here in this present effort, we exploit the proper self-testing mechanism of observables along with the local version of the tilted CHSH test to certify the functionality of all the devices involved in the QPQ scheme [35]. We also compare the performance of this full DI proposal of [35] with the performance of our recent full DI-QPQ proposal in [5] and *discuss relative advantages of both these schemes*. We further propose a device-independent scheme for a modification of [35] where the client can retrieve the maximum conclusive raw key bits. Here, based on some assumptions, we have strengthened the security of the QPQ scheme [35] by providing a full DI proposal. Following the assumptions from the recent DI oblivious transfer proposal in [9], here we assume that the devices involved in our full DI proposals are independent and memoryless but it is not a very practical assumption. Recently, there are some results for multi-round protocols on bit commitment [3], oblivious transfer and bit commitment [11], weak string erasure [21] etc. without the i.i.d. assumption. Although in [11] and [21], there are bounded/noisy storage assumptions. There are also some results in the single-shot setting (where the i.i.d. assumption is irrelevant) on bit commitment and coin flipping [33], weak coin flipping [2], XOR oblivious transfer [23] etc. However, to the best of our knowledge, there is still no result on the DI scenario of the distrustful primitive QPQ without the *i.i.d* assumption. We also consider here the asymptotic scenario where no channel noise is there. Our future aim is to analyze the performance of these QPQ schemes considering the channel noise. We are also interested to remove the *i.i.d* assumption over the devices in our future works.

Acknowledgements. The authors would like to thank the anonymous reviewers for their thoughtful and valuable comments that helped in improving the technical as well as the editorial quality of this paper.

Appendix. Here we first revisit the technique of choosing the initial sample size for this proposed schemes from the discussion in [5]. We also mention the proofs of Theorem 3.3, Theorem 3.4 and Theorem 4.4 which confirm the device-independent security of the above mentioned protocols. Atfirst in appendix A, we revisit the method of choosing initial sample size and the values of γ's in different phases from the discsuuion in [5]. Next in appendix B, C and D, we restate Theorem 1, 2 and 8 respectively along with their proofs.

Appendix A: Choice of initial sample size in practice. Similar to the discussion in [5], here in this section, we discuss the strategy of choosing the initial sample size for the proposed DI-QPQ schemes in finite sample scenario. Although we have assumed here that the channels are noiseless, in practice, there will be some channel noise and the parties have to allow some deviation (from the actual values of the testing parameters because of finite sample size) in each testing phase to certify the devices.

It is well known that the number of samples required to distinguish two events having probabilities p and $p(1+\epsilon)$ (for small ϵ) is approximately $O(\frac{1}{p\epsilon^2})$. One may require approximately $\frac{64}{p\epsilon^2}$ samples to achieve a confidence of more than 99% in distinguishing these two events. Recently, a more involved expression of the sample size is derived in [6] using Chernoff-Hoeffding [17] bound (stated here in proposition 3.2).

We consider $X_i = 1$ whenever Bob and Alice win the i-th instance in the testing phases of our proposed schemes, and $X_i = 0$ otherwise. Now if we consider $\mathbb{E}[X_i] = p$ and want to estimate the success probability p within an error margin of ϵp and confidence $1-\eta$, then we can write (from the result mentioned in [6]) that the required sample size m_{req} will be,

$$m_{\text{req}} \geq \frac{1}{2\epsilon^2 p^2} \ln \frac{1}{\eta} \tag{39}$$

From this expression of m_{req}, the two parties Bob and Alice can estimate the expected number of samples that are required for a particular testing phase to certify a device with certain accuracy and confidence.

Now, to ensure that each of Alice and Bob get the expected number of samples in every testing phase (to conclude with chosen accuracy and confidence) of our proposed schemes, they can choose their total initial sample size (i.e., the value of $\mathcal{K}$) as follows-

- Atfirst, before the start of the protocol, Bob and Alice (based on the protocol description) calculate the minimum number of samples required (according to the expression in inequality 39) in each testing phase to conclude with chosen accuracy and confidence.
- Then they choose the value of k to calculate the total number of samples required for the *key generation phase*.
- Atlast, they sum up all the number of samples required in every phase along with the number of samples required in *private query phase* to calculate the total initial sample size.
- After getting the total sample size, Bob and Alice proceed to each of the testing phases (according to the protocol description), select the required number of samples randomly from the shared instances and check for those chosen samples whether the value of a predefined parameter lies within the interval $[V-\epsilon p, V+\epsilon p]$ where V is the actual value of the testing parameter in honest scenario for asymptotically large number of samples. If this is the case, then with accuracy ϵp and chosen confidence η, they conclude about the functionality of their devices.

As an example, here we demonstrate the method of choosing samples for the first phase namely *source device and Bob's measurement device verification phase*. Before the start of the protocol, Bob and Alice choose the accuracy and confidence parameter for this phase with which they want to certify the source device and Bob's measurement device and let n_1 be the required number of samples for this phase. Now, similar to this *source device and Bob's measurement device verification phase*, they calculate the required number of samples for the other phases also and from that calculate the value of $\mathcal{K}$ i.e., the total number of samples required initially.

Bob and Alice then calculate the value of γ_1 such that,

$$n_1 = \gamma_1 \mathcal{K}$$

After getting the value of γ_1, Bob first chooses $\frac{\gamma_1 \mathcal{K}}{2}$ number of samples randomly from the $\mathcal{K}$ shared states and performs the TiltedCHSH test locally to certify the states and his devices. Then, from the rest $\left(\mathcal{K} - \frac{\gamma_1 \mathcal{K}}{2}\right)$ samples, Alice randomly chooses $\frac{\gamma_1 \mathcal{K}}{2}$ number of samples and performs the TiltedCHSH test locally to certify the states. In this similar way, they choose the samples for the remaining testing phases.

Note that this is a specific way of choosing samples that we demonstrate here from the several other possibilities. It is needless to say that one may follow any other strategies for choosing samples in different testing phases.

Appendix B: Statement and proof of Theorem 3.3.

Statement of Theorem 3.3: In the TiltedCHSH test of the *source device and Bob's measurement device verification phase*, either the devices achieve $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$ for both Alice and Bob (i.e., the states provided by the third party are identical with the actual states as mentioned in the QPQ scheme [35] and Bob's measurement device measures correctly in the $\{|0\rangle, |1\rangle\}$ basis) or the scheme is likely to abort with high probability (as the limit approaches infinity).

Proof. Here we prove the result considering that the game is played at the party $\mathcal{P}$'s end (one can replace $\mathcal{P}$ with Alice or Bob for the specific instances). Suppose, the first measurement operators of $\mathcal{P}$ are $\{B_b^y\}_{y,b\in\{0,1\}}$, for the input y and the output b and the second measurement operators of $\mathcal{P}$ are $\{A_a^{'x}\}_{x,a\in\{0,1\}}$, for the input x and the output a. Here, $\mathcal{P}$'s observable corresponding to the input $y \in \{0,1\}$ is,

$$B_y = \sum_{b\in\{0,1\}} (-1)^b B_b^y. \tag{40}$$

Similarly, $\mathcal{P}$'s observable corresponding to the input $x \in \{0,1\}$ is,

$$A'_x = \sum_{a\in\{0,1\}} (-1)^a A_a^{'x}. \tag{41}$$

Note that, in the TiltedCHSH test, the fraction $\beta_{\mathcal{B}}$ is being computed as follows,

$$\beta_{\mathcal{B}} = \alpha_{\mathcal{B}} \sum_{a\in\{0,1\}} (-1)^a \langle\psi_{\mathcal{BA}}| \mathbb{I} \otimes A_a^{'0} |\psi_{\mathcal{BA}}\rangle \tag{42}$$

$$+ \sum_{x,y,a,b\in\{0,1\}} (-1)^{d_{xyab}} \langle\psi_{\mathcal{BA}}| B_b^y \otimes A_a^{'x} |\psi_{\mathcal{BA}}\rangle \tag{43}$$

$$= [\langle\psi_{\mathcal{BA}}| W_{\mathcal{B}}^1 |\psi_{\mathcal{BA}}\rangle + \langle\psi_{\mathcal{BA}}| W_{\mathcal{B}}^2 |\psi_{\mathcal{BA}}\rangle] \tag{44}$$

$$= \langle\psi_{\mathcal{BA}}| W_{\mathcal{B}} |\psi_{\mathcal{BA}}\rangle \tag{45}$$

where $W_{\mathcal{B}}^1 := \alpha_{\mathcal{B}} \sum_{a\in\{0,1\}} (-1)^a \mathbb{I} \otimes A_a^{'0}$, $W_{\mathcal{B}}^2 := \left(\sum_{x,y,a,b\in\{0,1\}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'x}\right)$ are the two operators corresponding to $\beta_{\mathcal{B}}$ of the TitedCHSH test and $W_{\mathcal{B}} := W_{\mathcal{B}}^1 + W_{\mathcal{B}}^2$. We can rewrite the expression of $W_{\mathcal{B}}^1$ in the following way,

$$\begin{aligned} W_{\mathcal{B}}^1 &= \alpha_{\mathcal{B}} \sum_{a\in\{0,1\}} (-1)^a \mathbb{I} \otimes A_a^{'0} \\ &= \alpha_{\mathcal{B}} (\mathbb{I} \otimes A_0^{'0} - \mathbb{I} \otimes A_1^{'0}) \\ &= \alpha_{\mathcal{B}} [\mathbb{I} \otimes (A_0^{'0} - A_1^{'0})] \end{aligned}$$

By substituting the value of $(A_0^{'0} - A_1^{'0})$ from the equation 41 on the right-hand side of the above expression we get,

$$W_{\mathcal{B}}^1 = \alpha_{\mathcal{B}}(\mathbb{I} \otimes A_0'). \tag{46}$$

Similarly, We can also rewrite the expression of $W_{\mathcal{B}}^2$ in following way,

$$\begin{aligned} W_{\mathcal{B}}^2 = &\left(\sum_{\substack{x=0 \\ y,a,b\in\{0,1\}}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'0} \right) + \\ &\left(\sum_{\substack{x=1 \\ y,a,b\in\{0,1\}}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'1} \right) \\ &= W_{\mathcal{B}}^{02} + W_{\mathcal{B}}^{12} \end{aligned} \tag{47}$$

where $W_{\mathcal{B}}^{02} := \left(\sum_{\substack{x=0 \\ y,a,b\in\{0,1\}}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'0} \right)$ and $W_{\mathcal{B}}^{12} := \left(\sum_{\substack{x=1 \\ y,a,b\in\{0,1\}}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'1} \right)$. Note that, we can simplify further the expression of $W_{\mathcal{B}}^{02}$ in the following way,

$$\begin{aligned} W_{\mathcal{B}}^{02} = &\sum_{\substack{x=0 \\ y,a,b\in\{0,1\}}} (-1)^{d_{xyab}} B_b^y \otimes A_a^{'0} \\ = &\sum_{\substack{x=0 \\ y,a,b\in\{0,1\} \\ a\oplus b=0}} B_b^y \otimes A_a^{'0} - \sum_{\substack{x=0 \\ y,a,b\in\{0,1\} \\ a\oplus b\neq 0}} B_b^y \otimes A_a^{'0} \\ &= (B_0^0 \otimes A_0^{'0} + B_0^0 \otimes A_0^{'0} + B_1^0 \otimes A_1^{'0} + B_1^1 \otimes A_1^{'0}) - \\ &(B_1^0 \otimes A_0^{'0} + B_1^1 \otimes A_0^{'0} + B_0^0 \otimes A_1^{'0} + B_0^1 \otimes A_1^{'0}) \\ &= [(B_0^0 - B_1^0) \otimes A_0^{'0} - (B_0^0 - B_1^0) \otimes A_0^{'0} + \\ &(B_0^1 - B_1^1) \otimes A_0^{'0} - (B_0^1 - B_1^1) \otimes A_1^{'0}] \\ &= [(B_0^0 - B_1^0) \otimes (A_0^{'0} - A_1^{'0}) + \\ &(B_0^1 - B_1^1) \otimes (A_0^{'0} - A_1^{'0})] \\ &= [(B_0^0 - B_1^0) + (B_0^1 - B_1^1)] \otimes (A_0^{'0} - A_1^{'0}). \end{aligned}$$

By substituting the values of $(A_0^{'0} - A_1^{'0})$, $(B_0^0 - B_1^0)$ and $(B_0^1 - B_1^1)$ from the equation 41 and the equation 40 on the right-hand side of the above expression we get,

$$W_{\mathcal{B}}^{02} = (B_0 + B_1) \otimes A_0'. \tag{48}$$

Using similar approach we get the following simplified version of the expression $W_{\mathcal{B}}^{12}$.

$$W_{\mathcal{B}}^{12} = (B_0 - B_1) \otimes A_1'. \tag{49}$$

By substituting the values of $W_{\mathcal{B}}^{02}$ and $W_{\mathcal{B}}^{12}$ from the equation 48 and the equation 49 to the equation 47 we get,

$$W_{\mathcal{B}}^2 = (B_0 + B_1) \otimes A_0' + (B_0 - B_1) \otimes A_0'. \tag{50}$$

So, the right-hand side of the TiltedCHSH operator $W_{\mathcal{B}}$ is of the form,

$$W_{\mathcal{B}} = \alpha_{\mathcal{B}}(\mathbb{I} \otimes A'_0) + (B_0 + B_1) \otimes A'_0 + (B_0 - B_1) \otimes A'_1 \tag{51}$$

Note that this TiltedCHSH operator is exactly of the same form as the Tilted CHSH operator mentioned in [1]. Also, the states mentioned in our protocol can be obtained from the non-maximally entangled states mentioned in [1] by just applying a local unitary (hadamard gate) on the first qubit of the states mentioned in [1]. So, by following the same strategy as mentioned in [1], we can derive the following upper bound on the value of $\beta_{\mathcal{B}}$.

$$\beta_{\mathcal{B}} \leq \frac{4}{\sqrt{1+\sin^2\theta}} \tag{52}$$

One can easily check that for the TiltedCHSH test, the observables of $\mathcal{P}$ are of the following form,

$$B_0 = \sigma_z \quad B_1 = \sigma_x \tag{53}$$

$$A'_0 = \cos\mu\sigma_z + \sin\mu\sigma_x \quad A'_1 = \cos\mu\sigma_z - \sin\mu\sigma_x \tag{54}$$

It is already mentioned in [4] that the maximum value of the Tilted CHSH operator (here $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$) certifies that the states are of the form $\cos\frac{\theta}{2}|00\rangle + \sin\frac{\theta}{2}|11\rangle$ and the observables of $\mathcal{P}$'s are of the same form as mentioned in our TiltedCHSH test. As the states shared in our scheme is just a local isometry of the states mentioned in [4], we can easily conclude from the results mentioned in [4] that the maximum value of $\beta_{\mathcal{B}}$ (i.e., $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$) certifies the states in our scheme along with the standard basis of Bob's measurement device. According to our DI proposal, whenever the devices don't achieve the value $\beta_{\mathcal{B}} = \frac{4}{\sqrt{1+\sin^2\theta}}$, the protocol aborts. This concludes the proof. □

Appendix C: Statement and proof of Theorem 3.4.

Statement of Theorem 3.4: In OBStestAlice, either Alice's measurement devices achieve the value of the parameter $\beta_{\mathcal{A}} = \frac{1}{2\sin\theta}$ (i.e., her devices correctly measure in $\{|\phi_0\rangle, |\phi_0^{\perp}\rangle\}$ and $\{|\phi_1\rangle, |\phi_1^{\perp}\rangle\}$ basis) or the protocol terminates with a high likelihood of failure (as the limit approaches infinity).

Proof. It is already mentioned in the proof of theorem 3.3 that Alice's measurement operators are $\{A_a^x\}_{x,a\in\{0,1\}}$, corresponding to the input x and output a and Bob's measurement operators are $\{B_b^y\}_{y,b\in\{0,1\}}$, corresponding to the input y and output b. So, Alice's observable, corresponding to the input $x \in \{0,1\}$ is,

$$A_x = \sum_{a\in\{0,1\}} (-1)^a A_a^x. \tag{55}$$

Similarly, Bob's observable corresponding to the input $y \in \{0,1\}$ is,

$$B_y = \sum_{b\in\{0,1\}} (-1)^b B_b^y. \tag{56}$$

Note that in the OBStestAlice, the fraction $\beta_{\mathcal{A}}$ is being computed as follows,

$$\beta_{\mathcal{A}} = \frac{1}{4} \sum_{x,y,a,b\in\{0,1\}} (-1)^{d'_{xyab}} \alpha_{\mathcal{A}}^{1\oplus y} \langle\psi| B_b^y \otimes A_a^x |\psi\rangle \tag{57}$$

$$= \frac{1}{4}\langle\psi|W_{\mathcal{A}}|\psi\rangle, \tag{58}$$

where $W_{\mathcal{A}} := \left(\sum_{x,y,a,b\in\{0,1\}}(-1)^{d'_{xyab}}\alpha^{1\oplus y}B_b^y \otimes A_a^x\right)$ which is the operator corresponding to $\beta_{\mathcal{A}}$ of OBStestAlice. Now, proceeding like the similar way as mentioned in the derivation of the simplified form for operator $W_{\mathcal{B}}^2$ in the proof of theorem 3.3, here we can get the following expression of $W_{\mathcal{A}}$,

$$W_{\mathcal{A}} = \alpha_{\mathcal{A}}B_0 \otimes (A_0 + A_1) + B_1 \otimes (A_0 - A_1). \tag{59}$$

Note that, the right-hand side of the OBStestAlice operator $W_{\mathcal{A}}$ is almost of the same form as the tilted CHSH operator, described in [20].

So the expression of $W_{\mathcal{A}}^2$ can be written as,

$$\begin{aligned}
W_{\mathcal{A}}^2 &= \alpha_{\mathcal{A}}^2 B_0^2 \otimes (A_0^2 + A_1^2 + \{A_0, A_1\}) \\
&\quad + B_1^2 \otimes (A_0^2 + A_1^2 - \{A_0, A_1\}) \\
&= (\alpha_{\mathcal{A}}^2 B_0^2 + B_1^2 + \alpha_{\mathcal{A}}\{B_0, B_1\}) \otimes A_0^2 \\
&\quad + (\alpha_{\mathcal{A}}^2 B_0^2 + B_1^2 - \alpha_{\mathcal{A}}\{B_0, B_1\}) \otimes A_1^2 \\
&\quad + (\alpha_{\mathcal{A}}^2 B_0^2 - B_1^2) \otimes \{A_0, A_1\} - \alpha_{\mathcal{A}}[B_0, B_1] \otimes [A_0, A_1].
\end{aligned}$$

Using the property $B_j^2 \leq \mathbb{I}$, we can rewrite this expression as,

$$\begin{aligned}
W_{\mathcal{A}}^2 &\leq [(\alpha_{\mathcal{A}}^2 + 1).\mathbb{I} + \alpha_{\mathcal{A}}\{B_0, B_1\}] \otimes A_0^2 \\
&\quad + [(\alpha_{\mathcal{A}}^2 + 1).\mathbb{I} - \alpha_{\mathcal{A}}\{B_0, B_1\}] \otimes A_1^2 \\
&\quad + \mathbb{I} \otimes (\alpha_{\mathcal{A}}^2 - 1)\{A_0, A_1\} - \alpha[B_0, B_1] \otimes [A_0, A_1].
\end{aligned}$$

Since $-2.\mathbb{I} \leq \{B_0, B_1\} \leq 2.\mathbb{I}$, we have,

$$[(\alpha_{\mathcal{A}}^2 + 1).\mathbb{I} \pm \alpha_{\mathcal{A}}\{B_0, B_1\}] \geq 0$$

We can use the property $A_k^2 \leq \mathbb{I}$ and get the following simplified expression

$$\begin{aligned}
W_{\mathcal{A}}^2 &\leq 2(\alpha_{\mathcal{A}}^2 + 1).\mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes (\alpha_{\mathcal{A}}^2 - 1)\{A_0, A_1\} \\
&\quad - \alpha[B_0, B_1] \otimes [A_0, A_1]
\end{aligned}$$

We can further upper bound the commutators by their matrix moduli and use the relation $||[B_0, B_1]|| \leq 2.\mathbb{I}$ to get the following expression

$$W_{\mathcal{A}}^2 \leq 2(\alpha_{\mathcal{A}}^2 + 1).\mathbb{I} \otimes \mathbb{I} + T_{\alpha_{\mathcal{A}}} \otimes \mathbb{I} \tag{60}$$

where $T_{\alpha_{\mathcal{A}}} = (\alpha_{\mathcal{A}}^2 - 1)\{A_0, A_1\} + 2\alpha_{\mathcal{A}}||[A_0, A_1]||$

Now the expression of $T_{\alpha_{\mathcal{A}}}$ can also be upper bounded by upper bounding the anticommutators by its matrix modulus. So, the value of $T_{\alpha_{\mathcal{A}}}$ will be upper bounded by,

$$T_{\alpha_{\mathcal{A}}} \leq (\alpha_{\mathcal{A}}^2 - 1)|\{A_0, A_1\}| + 2\alpha_{\mathcal{A}}||[A_0, A_1]||$$

Again one can easily check that,

$$\begin{aligned}
&|\{A_0, A_1\}|^2 + |[A_0, A_1]|^2 \\
&= |A_0A_1 + A_1A_0|^2 + |A_0A_1 - A_1A_0|^2 \\
&= (A_0A_1 + A_1A_0)^\dagger(A_0A_1 + A_1A_0) \\
&\quad + (A_0A_1 + A_1A_0)^\dagger(A_0A_1 + A_1A_0) \\
&= 2(A_0A_1)^\dagger(A_0A_1) + 2(A_1A_0)^\dagger(A_1A_0)
\end{aligned} \tag{61}$$

Let us consider that the measurement operators are projective i.e., $(A_c^s)^2 = A_c^s$ and $(B_b^r)^2 = B_b^r$. Now for the projectors A_0^0 and A_1^0, $(A_0^0 + A_1^0) = \mathbb{I}$. From this relation we can write,

$$\begin{aligned}(A_0^0 + A_1^0)(A_0^0 + A_1^0)^\dagger &= \mathbb{I} \\ A_0^0.A_0^{0\dagger} + A_0^0.A_1^{0\dagger} + A_1^0.A_0^{0\dagger} + A_1^0.A_1^{0\dagger} &= \mathbb{I} \\ (A_0^0 + A_1^0) + (A_0^0.A_1^{0\dagger} + A_1^0.A_0^{0\dagger}) &= \mathbb{I}\end{aligned}$$

This implies,

$$(A_0^0.A_1^{0\dagger} + A_1^0.A_0^{0\dagger}) = 0$$

Now $A_0 = (A_0^0 - A_1^0)$. From this we can get,

$$\begin{aligned}A_0 A_0^\dagger &= (A_0^0 - A_1^0)(A_0^0 - A_1^0)^\dagger \\ &= A_0^0.A_0^{0\dagger} - A_0^0.A_1^{0\dagger} - A_1^0.A_0^{0\dagger} + A_1^0.A_1^{0\dagger} \\ &= (A_0^0 + A_1^0) - (A_0^0.A_1^{0\dagger} + A_1^0.A_0^{0\dagger}) \\ &= \mathbb{I} + 0 = \mathbb{I}\end{aligned}$$

Similarly, it can be shown that, $A_1 A_1^\dagger = A_1^\dagger A_1 = \mathbb{I}$.

So, from equation 61, we can write that for unitary observables A_0 and A_1,

$$\begin{aligned}|\{A_0, A_1\}|^2 + |[A_0, A_1]|^2 &= 2(A_0 A_1)^\dagger (A_0 A_1) \\ &\quad + 2(A_1 A_0)^\dagger (A_1 A_0) \\ &= 2\mathbb{I} + 2\mathbb{I} = 4\mathbb{I}\end{aligned}$$

This implies,

$$|\{A_0, A_1\}| = \sqrt{4.\mathbb{I} - |[A_0, A_1]|^2}$$

So, the simplified expression of $T_{\alpha_\mathcal{A}}$ will be of the form

$$T_{\alpha_\mathcal{A}} = (\alpha_\mathcal{A}^2 - 1)\sqrt{4.\mathbb{I} - |[A_0, A_1]|^2} + 2\alpha_\mathcal{A}|[A_0, A_1]|$$

Now one can easily check that the value of $|[A_0, A_1]|$ for which the value of $T_{\alpha_\mathcal{A}}$ becomes maximum is $|[A_0, A_1]| = \frac{4\alpha_\mathcal{A}}{(\alpha_\mathcal{A}^2+1)}.\mathbb{I}$ and the corresponding value of $T_{\alpha_\mathcal{A}}$ is $2(\alpha_\mathcal{A}^2 + 1).\mathbb{I}$. This implies that,

$$T_{\alpha_\mathcal{A}} = 2(\alpha_\mathcal{A}^2 + 1).\mathbb{I}$$

From this value of $T_{\alpha_\mathcal{A}}$ and from the expression of $W_\mathcal{A}^2$ mentioned in equation 60, we can easily write that the value of $W_\mathcal{A}$ is upper bounded by the following quantity-

$$W_\mathcal{A} \leq \sqrt{2(\alpha_\mathcal{A}^2 + 1)\mathbb{I} \otimes \mathbb{I} + T_{\alpha_\mathcal{A}} \otimes \mathbb{I}} \tag{62}$$

where $T_{\alpha_\mathcal{A}} = 2(\alpha_\mathcal{A}^2 + 1).\mathbb{I}$.

Now, the value $\beta_\mathcal{A}$ obtained in OBStestAlice of our algorithm can be written alternatively as $\beta_\mathcal{A} = \frac{\mathrm{Tr}(W_\mathcal{A}\rho_{\mathcal{B}\mathcal{A}})}{4}$ where $\rho_{\mathcal{B}\mathcal{A}}$ is the density matrix representation of the shared states $|\psi\rangle_{\mathcal{B}\mathcal{A}}$ i.e., $\rho_{\mathcal{B}\mathcal{A}} = |\psi\rangle_{\mathcal{B}\mathcal{A}} \langle\psi|$. From this expression of $\beta_\mathcal{A}$, one can easily derive that the value of $\beta_\mathcal{A}^2$ is upper bounded by the following quantity,

$$\beta_\mathcal{A}^2 \leq \frac{\mathrm{Tr}(W_\mathcal{A}^2 \rho_{\mathcal{B}\mathcal{A}})}{16} \tag{63}$$

Now if we assume $t_{\alpha_{\mathcal{A}}} := \frac{1}{4}\mathrm{Tr}(T_{\alpha_{\mathcal{A}}}\rho_{\mathcal{A}}) - \frac{1}{2}(\alpha_{\mathcal{A}}^2 - 1)$ (where $\rho_{\mathcal{A}}$ is the reduced state at Alice's side) then using this value of t_α along with the value of $W_{\mathcal{A}}$ obtained from expression 62 and the upper bound on the value of $\beta_{\mathcal{A}}^2$, we can write that the $\beta_{\mathcal{A}}$ value mentioned in OBStestAlice is upper bounded by the following quantity,

$$\beta_{\mathcal{A}} \leq \frac{\sqrt{\alpha_{\mathcal{A}}^2 + t_{\alpha_{\mathcal{A}}}}}{2}, \tag{64}$$

where, $t_{\alpha_{\mathcal{A}}} := \frac{1}{4}\mathrm{Tr}(T_{\alpha_{\mathcal{A}}}\rho_A) - \frac{1}{2}(\alpha_{\mathcal{A}}^2 - 1)$.

Now here, the observables are projective (i.e., $A_j^2 = \mathbb{I}$) and the anticommutator $\{A_0, A_1\}$ is a positive semi definite operator. Since we have already shown that the value of the anti-hermitian operator $|[A_0, A_1]|$ is $|[A_0, A_1]| = \frac{4\alpha_{\mathcal{A}}}{(\alpha_{\mathcal{A}}^2+1)}.\mathbb{I}$ for the maximum value of $T_{\alpha_{\mathcal{A}}}$, the spectral decomposition of $[A_0, A_1]$ can be written as,

$$[A_0, A_1] = \frac{4\alpha_{\mathcal{A}}.i}{(\alpha_{\mathcal{A}}^2 + 1)}(P_+^{\mathcal{A}} - P_-^{\mathcal{A}})$$

for some orthogonal projectors $P_+^{\mathcal{A}}$ and $P_-^{\mathcal{A}}$ such that $(P_+^{\mathcal{A}} + P_-^{\mathcal{A}}) = \mathbb{I}$. As it is well-known that for projective observables, the commutator holds the property $A_0[A_0, A_1]A_0 = -[A_0, A_1]$, we can easily conclude that $A_0 P_{\pm}^{\mathcal{A}} A_0 = P_{\mp}^{\mathcal{A}}$. Let us consider that $\{|e_j^0\rangle\}_j$ is an orthonormal basis for the support of $P_+^{\mathcal{A}}$ and $\{|e_j^1\rangle\}_j$ is an orthonormal basis for the support of $P_-^{\mathcal{A}}$ where $|e_j^1\rangle = A_0 |e_j^0\rangle$. We define the unitary operator U_0 as

$$U_0 |e_j^d\rangle = \frac{1}{\sqrt{2}}[|0\rangle + (-1)^d i |1\rangle] |j\rangle$$

for $d \in \{0, 1\}$. Then we can easily verify that,

$$U_0[A_0, A_1]U_0^\dagger = \frac{4\alpha_{\mathcal{A}}.i}{(\alpha_{\mathcal{A}}^2 + 1)}\sigma_Y \otimes \mathbb{I}$$

Since $\{\mathbb{I}, \sigma_X, \sigma_Y, \sigma_Z\}$ constitute an operator basis for linear operators acting on $\mathbb{C}^2$, without loss of generality we can write

$$U_0 A_0 U_0^\dagger = \mathbb{I} \otimes K_0 + \sigma_X \otimes K_x + \sigma_Y \otimes K_y + \sigma_Z \otimes K_z$$

for some hermitian operator K_0, K_x, K_y, K_z. For projective observable A_0, one can easily check that $\{A_0, [A_0, A_1]\} = 0$. This relation satisfies only when $K_0 = K_y = 0$. As $A_0^2 = \mathbb{I}$, K_x and K_z must satisfy the relation

$$K_x^2 + K_z^2 = \mathbb{I} \text{ and } [K_x, K_z] = 0$$

So, we can easily write K_x and K_z in the following form.

$$K_x = \sum_j \sin 2\gamma_j |j\rangle \langle j|$$

$$K_z = \sum_j \cos 2\gamma_j |j\rangle \langle j|$$

for some angle γ_j and some orthonormal basis $\{|j\rangle\}$. This implies that,

$$U_0 A_0 U_0^\dagger = \sum_j (\sin\gamma_j \sigma_X + \cos\gamma_j \sigma_Z) \otimes |j\rangle \langle j|$$

We now consider the following controlled unitary to align the qubit observables.

$$U_1 = \sum_j \exp(-i0.\sigma_Y) \otimes |j\rangle \langle j|$$

Now for this defined unitary operator, one can easily check that,

$$U_1U_0A_0U_0^\dagger U_1^\dagger = (\sin\gamma_j\sigma_X + \cos\gamma_j\sigma_Z)\otimes\mathbb{I}$$
$$U_1U_0[A_0,A_1]U_0^\dagger U_1^\dagger = \frac{4\alpha_{\mathcal{A}}.i}{(\alpha_{\mathcal{A}}^2+1)}\sigma_Y\otimes\mathbb{I}$$

Like observable A_0, an analogous reasoning can also be applied for observable A_1 and from that, without loss of generality we can write

$$U_1U_0A_1U_0^\dagger U_1^\dagger = \sigma_X\otimes K_x' + \sigma_Z\otimes K_z'$$

Since the commutators are positive semi definite and the observables are projective, we can easily check that

$$\{A_0,A_1\} = |\{A_0,A_1\}| = \sqrt{4.\mathbb{I} - |[A_0,A_1]|^2}$$
$$= \frac{2(\alpha_{\mathcal{A}}^2-1)}{(\alpha_{\mathcal{A}}^2+1)}.\mathbb{I}$$

Now we define $2\gamma_j := \arccos\left(\frac{\alpha_{\mathcal{A}}^2-1}{\alpha_{\mathcal{A}}^2+1}\right) = 0$. From this relation, imposing consistency on the anticommutator, we get,

$$K_x'\sin\gamma_j + K_z'\cos\gamma_j = \cos 2\gamma_j \tag{65}$$

On the other hand, imposing consistency on the commutator, we get,

$$K_x'\cos\gamma_j - K_z'\sin\gamma_j = -\sin 2\gamma_j \tag{66}$$

Now, solving equation 65 and 66, we get,

$$K_x' = \sin\gamma_j \text{ and } K_z' = \cos\gamma_j$$

From the relation $2\gamma_j := \arccos\left(\frac{\alpha_{\mathcal{A}}^2-1}{\alpha_{\mathcal{A}}^2+1}\right) = 0$, we can get the value of $\alpha_{\mathcal{A}}$ which is

$$\alpha_{\mathcal{A}} = \cot\gamma_j$$

For this value of $\alpha_{\mathcal{A}}$, we can easily derive that $t_{\alpha_{\mathcal{A}}} = 1$. This implies that the value of $\beta_{\mathcal{A}}$ corresponding to these observables A_0 and A_1 will be,

$$\beta_{\mathcal{A}} = \frac{1}{2\sin\gamma_j}. \tag{67}$$

If we consider $U_{\mathcal{A}} = U_0^\dagger U_1^\dagger$ then the observables A_0 and A_1 will be of the form

$$A_0 = U_{\mathcal{A}}(\cos\gamma_j\sigma_Z + \sin\gamma_j\sigma_X\otimes\mathbb{I})U_{\mathcal{A}}^\dagger$$
$$A_1 = U_{\mathcal{A}}(\cos\gamma_j\sigma_Z - \sin\gamma_j\sigma_X\otimes\mathbb{I})U_{\mathcal{A}}^\dagger$$

Setting $\gamma_j = \theta$ shows that in OBStestAlice, if the value of the parameter $\beta_{\mathcal{A}}$ is equal to $\frac{1}{2\sin\theta}$, then the measurement operators at Alice's side are same as the one described in the OBStestAlice. In our DI proposal, whenever the devices involved in OBStestAlice do not achieve the value $\beta_{\mathcal{A}} = \frac{1}{2\sin\theta}$, the protocol aborts. This concludes the proof. □

Appendix D: Statement and proof of Theorem 4.4.

Statement of Theorem 4.4: POVMtestAlice either results in a high probability of termination of this modified proposal (as the limit approaches infinity), or it guarantees that Alice's measurement devices attain $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$, meaning they are of this specified form (up to a local unitary),

$$D_0 = \frac{1}{(1+\cos\theta)}(|\phi_1^{\perp}\rangle\langle\phi_1^{\perp}|)$$
$$D_1 = \frac{1}{(1+\cos\theta)}(|\phi_0^{\perp}\rangle\langle\phi_0^{\perp}|)$$
$$D_2 = \mathbb{I} - D_0 - D_1,$$

where $|\phi_1^{\perp}\rangle = \left(\sin\frac{\theta}{2}|0\rangle + \cos\frac{\theta}{2}|1\rangle\right)$ and $|\phi_0^{\perp}\rangle = \left(\sin\frac{\theta}{2}|0\rangle - \cos\frac{\theta}{2}|1\rangle\right)$.

Proof. In algorithm KeyGenAlice of this modified protocol, Alice applies the POVM D on a single qubit state ρ_{R_i} (where R_i is the raw key bit indexed by i at Bob's side). So, without any loss of generality we can assume that $D_i \in D$ has the following form,

$$D_i = \lambda_i(\mathbb{I} + \vec{d_i}.\vec{\sigma}), \tag{68}$$

where $\vec{d_i} = [d_{i0}, d_{i1}, d_{i2}]$ and it is the Bloch vector with length at most one, $\vec{\sigma} = [\sigma_X, \sigma_Y, \sigma_Z]$ are the Pauli matrices and $\lambda_i \geq 0$.

In this case, one may wonder how we can fix the dimension of D_i here in the proof in DI scenario? The answer to this question is that here we are able to fix the dimension of D_i and choose this particular general form because of the tests mentioned earlier in the *source device and Bob's measurement device verification phase* (corresponding result mentioned in Theorem 4.3) which certifies that the states shared between Alice and Bob are of the specified form (upto a unitary) as mentioned in [35] and after Bob's projective measurements, the reduced states at Alice's side are one qubit states.

Now, the condition $\sum_{i=0}^{2} D_i = \mathbb{I}$ leads us to the following relations,

$$\sum_{i=0}^{2} \lambda_i = 1 \tag{69}$$

$$\sum_{i=0}^{2} \lambda_i \vec{d_i} = 0. \tag{70}$$

In terms of Bloch vector we can rewrite ρ_0, ρ_1 in following way,

$$\rho_0 = \frac{1}{2}(\mathbb{I} + \cos\theta\sigma_Z + \sin\theta\sigma_X) \tag{71}$$

$$\rho_1 = \frac{1}{2}(\mathbb{I} + \cos\theta\sigma_Z - \sin\theta\sigma_X). \tag{72}$$

In the algorithm POVMtestAlice, if Alice would like to maximize her winning probability then she needs to maximize the following expression,

$$\Omega = \sum_{R_i, R_{\mathcal{A}_i} \in 0,1} (-1)^{R_i \oplus R_{\mathcal{A}_i}} \mathrm{Tr}[D_{R_{\mathcal{A}_i}} \rho_{R_i}]. \tag{73}$$

In terms of $\lambda_i, \vec{d_i}, \vec{\sigma}$ we have,

$$\mathrm{Tr}[D_0\rho_0] = \lambda_0(1 + d_{00}\sin\theta + d_{02}\cos\theta)$$

$$\begin{aligned}\mathrm{Tr}[D_0\rho_1] &= \lambda_0(1 - d_{00}\sin\theta + d_{02}\cos\theta)\\ \mathrm{Tr}[D_1\rho_0] &= \lambda_1(1 + d_{10}\sin\theta + d_{12}\cos\theta)\\ \mathrm{Tr}[D_1\rho_1] &= \lambda_1(1 - d_{10}\sin\theta + d_{12}\cos\theta)\end{aligned}$$

In terms of $\lambda_i, \vec{d_i}, \vec{\sigma}$ can rewrite Ω as,

$$\begin{aligned}\Omega = &\lambda_0(1 + d_{00}\sin\theta + d_{02}\cos\theta)\\ &+ \lambda_1(1 - d_{10}\sin\theta + d_{12}\cos\theta)\\ &- \lambda_0(1 - d_{00}\sin\theta + d_{02}\cos\theta)\\ &- \lambda_1(1 + d_{10}\sin\theta + d_{12}\cos\theta)\end{aligned} \tag{74}$$

As both $\mathrm{Tr}[D_0\rho_1]$ and $\mathrm{Tr}[D_1\rho_0]$ are positive quantity, hence

$$\Omega \leq \lambda_0(1 + d_{00}\sin\theta + d_{02}\cos\theta) + \lambda_1(1 - d_{10}\sin\theta + d_{12}\cos\theta) \tag{75}$$

and this implies that for maximum value of Ω,

$$\lambda_0(1 - d_{00}\sin\theta + d_{02}\cos\theta) = 0 \tag{76}$$

$$\lambda_1(1 + d_{10}\sin\theta + d_{12}\cos\theta) = 0. \tag{77}$$

As both of ρ_0, ρ_1 lie on the XZ plane and due to the freedom of global unitary, without loss of generality we can assume $d_{01} = d_{11} = d_{21} = 0$. Due to the positivity constraint ($D_i \geq 0$) we have,

$${d_{00}}^2 + {d_{02}}^2 \leq 1 \tag{78}$$

$${d_{10}}^2 + {d_{12}}^2 \leq 1 \tag{79}$$

$${d_{20}}^2 + {d_{22}}^2 \leq 1. \tag{80}$$

Without any loss of generality we can assume that for the maximum value of Ω, ${d_{00}}^2 + {d_{02}}^2 = 1$. So, we can parameterize d_{00}, d_{02} as $\cos\alpha, \sin\alpha$ ($-2\pi \leq \alpha \leq 2\pi$). By substituting $d_{00} = \cos\alpha, d_{02} = \sin\alpha$ in equation 76 we get,

$$1 - \cos\alpha\sin\theta + \sin\alpha\cos\theta = 0$$

This implies,

$$\sin(\theta - \alpha) = 1 = \sin\frac{\pi}{2}.$$

As $-2\pi \leq \alpha \leq 2\pi$, so $\sin(\theta - \alpha) = 1$ implies,

$$\theta - \alpha = \frac{\pi}{2} \quad \text{and,}$$

$$\alpha = \left(\theta - \frac{\pi}{2}\right). \tag{81}$$

From the equation 81 we get,

$$\vec{d_0} = [\sin\theta, 0, -\cos\theta]. \tag{82}$$

Similarly, for maximum value of Ω, ${d_{10}}^2 + {d_{12}}^2 = 1$. So, we can parameterize d_{10}, d_{12} as $\cos\alpha, \sin\alpha$ ($-2\pi \leq \alpha \leq 2\pi$). By substituting $d_{10} = \cos\alpha, d_{12} = \sin\alpha$ in equation 77 we get,

$$1 + \cos\alpha\sin\theta + \sin\alpha\cos\theta = 0$$

This implies,

$$\sin(\theta + \alpha) = -1 = \sin\frac{3\pi}{2}.$$

As $-2\pi \leq \alpha \leq 2\pi$, so $\sin(\theta+\alpha) = -1$ implies,

$$\theta + \alpha = \frac{3\pi}{2} \quad \text{and,}$$
$$\alpha = \left(\frac{3\pi}{2} - \theta\right). \tag{83}$$

From the equation 83 we get,

$$\vec{d_1} = [-\sin\theta, 0, -\cos\theta]. \tag{84}$$

By substituting the expression of $\vec{d_0}, \vec{d_1}$ in equation 75 we get,

$$\Omega \leq (\lambda_0 + \lambda_1)(1 - \cos 2\theta). \tag{85}$$

Now again substituting the values of $\vec{d_0}, \vec{d_1}$ in equation 70 we get,

$$\lambda_0 \sin\theta - \lambda_1 \sin\theta + \lambda_2 d_{20} = 0 \tag{86}$$
$$-\lambda_0 \cos\theta - \lambda_1 \cos\theta + \lambda_2 d_{22} = 0. \tag{87}$$

Due to the constraint equation 80, similar to $\vec{d_0}$ and $\vec{d_1}$, here also we parameterize the expression of d_{20}, d_{22} as $\sin\beta, \cos\beta$ respectively. By substituting $d_{20} = \sin\beta$ and $d_{22} = \cos\beta$ in the equations 86 and 87 we get,

$$\lambda_0 \sin\theta - \lambda_1 \sin\theta + \lambda_2 \sin\beta = 0 \tag{88}$$
$$-\lambda_0 \cos\theta - \lambda_1 \cos\theta + \lambda_2 \cos\beta = 0. \tag{89}$$

By solving equation 88 and equation 89 together with equation 69 we get,

$$\lambda_0 = \frac{\sin(\theta-\beta)}{[\sin(\theta+\beta) + \sin(\theta-\beta) + \sin 2\theta]} \tag{90}$$
$$\lambda_1 = \frac{\sin(\theta+\beta)}{[\sin(\theta+\beta) + \sin(\theta-\beta) + \sin 2\theta]} \tag{91}$$

Hence,

$$\lambda_0 + \lambda_1 = \frac{\sin(\theta+\beta) + \sin(\theta-\beta)}{[\sin(\theta+\beta) + \sin(\theta-\beta) + \sin 2\theta]} \tag{92}$$
$$= \frac{\cos\beta}{(\cos\beta + \cos\theta)}. \tag{93}$$

According to equation 85, for getting a tight upper bound on Ω we need to maximize $(\lambda_0 + \lambda_1)$. By equating $\frac{d(\lambda_0+\lambda_1)}{d\beta} = 0$ in equation 93 we get,

$$\frac{-\sin\beta\cos\theta}{(\cos\beta + \cos\theta)^2} = 0. \tag{94}$$

This implies,

$$\beta = 0. \tag{95}$$

It is also easy to check that for $\beta = 0$, the expression $\frac{d^2(\lambda_0+\lambda_1)}{d\beta^2} < 0$. Hence, the expression $\lambda_0 + \lambda_1$ maximizes at the point $\beta = 0$. Substituting this relation in equations 90 and 91 we get,

$$\lambda_0 = \lambda_1 = \frac{1}{2(1+\cos\theta)}. \tag{96}$$

By substituting the values of $\lambda_0 + \lambda_1$ in equation 69 we get,

$$\lambda_2 = \frac{\cos\theta}{1+\cos\theta}. \tag{97}$$

Hence, we get,

$$\Omega \leq \frac{2\sin^2\theta}{(1+\cos\theta)}, \tag{98}$$

and

$$D_0 = \frac{1}{2(1+\cos\theta)}(\mathbb{I} + \sin\theta\sigma_X - \cos\theta\sigma_Z) \tag{99}$$

$$D_1 = \frac{1}{2(1+\cos\theta)}(\mathbb{I} - \sin\theta\sigma_X - \cos\theta\sigma_Z) \tag{100}$$

$$D_2 = \frac{\cos\theta}{1+\cos\theta}(\mathbb{I} + \sigma_Z). \tag{101}$$

We can rewrite the above expressions as follows,

$$D_0 = \frac{1}{(1+\cos\theta)}(\left|\phi_1^\perp\right\rangle\left\langle\phi_1^\perp\right|)$$
$$D_1 = \frac{1}{(1+\cos\theta)}(\left|\phi_0^\perp\right\rangle\left\langle\phi_0^\perp\right|)$$
$$D_2 = \mathbb{I} - D_0 - D_1,$$

where $\left|\phi_1^\perp\right\rangle = \left(\sin\frac{\theta}{2}\left|0\right\rangle + \cos\frac{\theta}{2}\left|1\right\rangle\right)$ and $\left|\phi_0^\perp\right\rangle = \left(\sin\frac{\theta}{2}\left|0\right\rangle - \cos\frac{\theta}{2}\left|1\right\rangle\right)$. This implies that whenever the measurement devices at Alice's side achieve $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$, then it certifies that the measurement operators at Alice's side are the intended POVM devices. In our modified DI proposal, whenever the devices involved in POVMtestAlice do not achieve the value $\Omega = \frac{2\sin^2\theta}{(1+\cos\theta)}$, the protocol aborts. This concludes the proof.

□

REFERENCES

[1] A. Acin, S. Massar and S. Pironio, Randomness versus nonlocality and entanglement, *Phys. Rev. Lett.*, **108** (2012), 100402.

[2] N. Aharon, A. Chailloux, I. Kerenidis, S. Massar, S. Pironio and J. Silman, Weak coin flipping in a device-independent setting, *Theory of Quantum Computation, Communication, and Cryptography, Lecture Notes in Comput. Sci.*, **6745** (2014), 1-12.

[3] N. Aharon, S. Massar, S. Pironio and J. Silman, Device-independent bit commitment based on the CHSH inequality, *New J. Phys.*, **18** (2016), 025014.

[4] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, *Phys. Rev. A*, **91** (2015), 052111, 2015.

[5] J. Basak, K. Chakraborty, A. Maitra and S. Maitra, Improved and formal proposal for device-independent quantum private query, *J. Phys. A: Math. Theor.*, **57** (2024), 085302, https://arxiv.org/abs/1901.03042.

[6] J. Basak and S. Maitra, Clauser-Horne-Shimony-Holt versus three-party pseudo-telepathy: On the optimal number of samples in device-independent quantum private query, *Quantum Inf. Process.*, **17** (2018), Paper No. 77, 14 pp.

[7] C. H. Bennett, Quantum cryptography using any two nonorthogonal states, *Phys. Rev. Lett.*, **68** (1992), 3121-3124.

[8] C. H. Bennett and G. Brassard, Quantum cryptography: Public key distribution and coin tossing, *Theoretical Computer Science*, **560** (2014), 7-11.

[9] A. Broadbent and P. Yuen, Device-independent oblivious transfer from the bounded-quantum-storage-model and computational assumptions, *New J. Phys.*, **25** (2023), 053019, 21 pp, arXiv:2111.08595.

[10] K. Chakraborty, A. Chailloux and A. Leverrier, Arbitrarily long relativistic bit commitment, *Phys. Rev. Lett.*, **115** (2015), 250501.

[11] I. B. Damgaard, S. Fehr, R. Renner, L. Salvail and C. Schaffner, A tight high-order entropic quantum uncertainty relation with applications, *Advances in cryptology-CRYPTO, Lecture Notes in Comput. Sci.*, **4622** (2007), 360–378.
[12] S. Fehr and M. Fillinger, On the composition of two-prover commitments, and applications to multi-round relativistic commitments, *Advances in Cryptology—EUROCRYPT*, (2016), 477-496.
[13] F. Gao, B. Liu, Q. Y. Wen and H. Chen, Flexible quantum private queries based on quantum key distribution, *Opt. Express*, **20** (2012), 17411-17420.
[14] Y. Gertner, Y. Ishai, E. Kushilevitz and T. Malkin, Protecting data privacy in private information retrieval schemes, *Journal of Computer and System Sciences*, **60** (2000), 592-629.
[15] V. Giovannetti, S. Lloyd and L. Maccone, Quantum private queries, *Phys. Rev. Lett.*, **100** (2008), 230502, 4 pp.
[16] V. Giovannetti, S. Lloyd and L. Maccone, Quantum private queries: Security analysis, *IEEE Trans. Info. Theory*, **56** (2010), 3465-3477.
[17] W. Hoeffding, Probability inequalities for sums of bounded random variables, *Journal of the American Statistical Association*, **58** (1963), 13-30.
[18] I. D. Ivanovic, How to differentiate between non-orthogonal states, *Physics Lett. A*, **123** (1987), 257-259.
[19] M. Jakobi, C. Simon, N. Gisin, J. D. Bancal, C. Branciard, N. Walenta and H. Zbinden, Practical private database queries based on a quantum-key-distribution protocol, *Phys. Rev. A*, **83** (2011), 022301.
[20] J. Kaniewski, Self-testing of binary observables based on commutation, *Phys. Rev. A*, **95** (2017), 062323.
[21] J. Kaniewski and S. Wehner, Device-independent two-party cryptography secure against sequential attacks, *New J. Phys.*, **18** (2016), 055004, 21 pp.
[22] R. Konig, R. Renner and C. Schaffner, The operational meaning of min- and max-entropy, *IEEE Trans. Info. Theory*, **55** (2009), 4337-4347.
[23] S. Kundu, J. Sikora and E. Y.-Z. Tan, A device-independent protocol for XOR oblivious transfer, *Quantum*, **6** (2022), 725.
[24] T. Lunghi, J. Kaniewski, F. Bussiéres, R. Houlmann, M. Tomamichel, S. Wehner and H. Zbinden, Practical relativistic bit commitment, *Phys. Rev. Lett.*, **115** (2015), 030502.
[25] A. Maitra, G. Paul and S. Roy, Device-independent quantum private query, *Phys. Rev. A*, **95** (2017), 042344.
[26] M. Naor, Bit commitment using pseudorandomness, *J. Cryptology*, **4** (1991), 151-158.
[27] T. G. Noh, Counterfactual quantum cryptography, *Phys. Rev. Lett.*, **103** (2009), 230501, 4 pp.
[28] L. Olejnik, Secure quantum private information retrieval using phase-encoded queries, *Phys. Rev. A*, **84** (2011), 022313.
[29] M. P. Rao and M. Jakobi, Towards communication-efficient quantum oblivious key distribution, *Phys. Rev. A*, **87** (2013), 012331.
[30] https://www.cs.princeton.edu/courses/archive/spr08/cos598D/scribe5.pdf.
[31] http://theory.stanford.edu/~trevisan/cs276/lecture27.pdf.
[32] V. Scarani, A. Acín, G. Ribordy and N. Gisin, Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations, *Phys. Rev. Lett.*, **92** (2004), 057901.
[33] J. Silman, A. Chailloux, N. Aharon, I. Kerenidis, S. Pironio and S. Massar. Fully distrustful quantum bit commitment and coin flipping, *Phys. Rev. Lett.*, **106** (2011), 220501.
[34] A. Tavakolix, M. Smania, T. Vértesi, N. Brunner and M. Bourennane, Self-testing non-projective quantum measurements in prepare-and-measure experiments, *Science Advances*, **6** (2020).
[35] Y. G. Yang, S. J. Sun, P. Xu and J. Tiang, Flexible protocol for quantum private query based on B92 protocol, *Quant. Info. Proc.*, **13** (2014), 805-813.
[36] J. L. Zhang, F. Z. Guo, F. Gao, B. Liu and Q. Y. Wen, Private database queries based on counterfactual quantum key distribution, *Phys. Rev. A*, **88** (2013), 022334.

Received August 2023; 1st revision August 2023; 2nd revision January 2024; early access March 2024.