2017 IEEE 30th Computer Security Foundations Symposium

Differential Privacy in Quantum Computation

Li Zhou* and Mingsheng Ying†
Dept. of Computer Science and Technology, Tsinghua University, China
CQSI, FEIT, University of Technology Sydney, Australia
SKLCS, Institute of Software, Chinese Academy of Sciences, China
*Email: zhou-l14@mails.tsinghua.edu.cn
†Email: Mingsheng.Ying@uts.edu.au

***Abstract*—More and more quantum algorithms have been designed for solving problems in machine learning, database search and data analytics. An important problem then arises: how privacy can be protected when these algorithms are used on private data? For classical computing, the notion of differential privacy provides a very useful conceptual framework in which a great number of mechanisms that protect privacy by introducing certain noises into algorithms have been successfully developed. This paper defines a notion of differential privacy for quantum information processing. We carefully examine how the mechanisms using three important types of quantum noise, the amplitude/phase damping and depolarizing, can protect differential privacy. A composition theorem is proved that enables us to combine multiple privacy-preserving operations in quantum information processing.**

I. Introduction

One of the first quantum algorithms, the Grover algorithm [1], was designed for database search. In the last decade, more and more sophisticated and elaborate quantum algorithms for data analytics and related problems have been discovered, e.g. element distinctness [2], principal component analysis [3], [4], data classification [5], machine learning [6]. Experiments implementing these algorithms have also been reported, e.g. [7].

A huge amount of data contains private information, e.g. medical, insurance and banking data. Whenever quantum computers become commercially available and these quantum algorithms are used in real-world applications, an important problem will need to be addressed [8]: how privacy can be protected in quantum computation and quantum information processing?

In the realm of classical computing, Dwork et al. [9], [10] introduced the fundamental notion of differential privacy by observing that it is impossible to completely avoid the statistical disclosure defined by Dalenius [11]. Intuitively, differential privacy guarantees that, the answer to a query based on a statistical database is almost the same whether or not any individual participates in the data set, so the privacy of any participants will not be disclosed significantly. Within this conceptual framework, a great number of mechanisms and algorithms that protect privacy have been developed [12], [13], [14], and applied in different fields, such as data analysis [15], data mining [16] and machine learning [17].

Contributions of the paper: In this paper, we define a notion of differential privacy for quantum operations; that is, completely positive and trace-preserving (CPTP) maps, which are the most general (discrete-time) mathematical formalism of physically realisable operations on quantum systems [18]. It can be used to measure privacy leak in quantum computation and information processing. Furthermore, we introduce three privacy mechanisms employing three widely used models of quantum noises: generalized amplitude damping mechanism, phase-amplitude damping mechanism and depolarizing mechanism. The differential privacy parameters for each of these cases are settled. A composition theorem is established that enables us to accomplish a quantum computational or information processing task by combining multiple privacy-preserving operations.

Organisation of the paper: In Section II, for the convenience of the reader, we first review some basic notions in quantum theory and the definition of differential privacy for classical computing. Then we formalize the notion of quantum differential privacy. In Section III, we introduce three privacy mechanisms using important and widely used quantum noises: generalized amplitude damping mechanism, phase-amplitude damping mechanism and depolarizing mechanism, respectively. To further illustrate the notion of quantum differential privacy, the differential privacy parameters for each of these quantum noise models are calculated. We carefully compare these three mechanisms with their privacy parameters. A composition theorem for combining different privacy mechanisms is established in Section IV. In Section V, we present an algorithm for computing privacy parameters through sampling of inputs. Its accuracy is analysed at the end of this section. For readability, the very technical proofs of several theorems are postponed to Section VII-C. We draw conclusions and point out some issues for future research in Section VII.

II. Basic Definitions

In this section, we are going to formally define the notion of differential privacy for quantum computation.

A. Quantum States

For the convenience of the reader, let us first briefly review some basic notions in quantum theory; for details, the reader can consult the standard textbook [18]. According to a basic postulate of quantum mechanics, the state space of a quantum system is a Hilbert space $\mathcal{H}$, i.e. a complex vector space with an inner product that is complete in the sense that

DOI 10.1109/CSF.2017.23

every Cauchy sequence has a limit. We use Dirac's notation $|\varphi\rangle, |\psi\rangle, ...$ to denote vectors. The inner product of $|\varphi\rangle$ and $|\psi\rangle$ is denoted $\langle\varphi|\psi\rangle$. A pure quantum state is represented by a unit vector, i.e. a vector $|\psi\rangle$ with length

$$\|\psi\| = \sqrt{\langle\psi|\psi\rangle} = 1.$$

For example, for finite n, an n-dimensional Hilbert space is essentially the space $\mathbb{C}^n$ of complex vectors. In particular, the state space of a qubit (quantum bit) is the 2-dimensional Hilbert space. A qubit can be in the basis states $|0\rangle$, $|1\rangle$, and it can also be in their superpositions like

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

A mixed state of a quantum system is represented by an ensemble

$$\mathbb{E} = \{(p_1, |\psi_1\rangle), ..., (p_k, |\psi_k\rangle)\}$$

meaning that the system is in state $|\psi_i\rangle$ with probability p_i, where $0 \leq p_i$ and $\sum_i p_i = 1$. A convenient mathematical tool for the description of a mixed state is density operator. Let A be a linear operator in Hilbert space A. The trace of A is the complex number

$$tr(A) = \sum_i \langle i|A|i\rangle$$

where $\{|i\rangle\}$ is an orthonormal basis of the space, and $\langle i|A|i\rangle$ stands for the inner product of $|i\rangle$ and $A|i\rangle$. The external product $A = |\varphi\rangle\langle\psi|$ of two vectors $|\varphi\rangle, |\psi\rangle$ is an operator defined as follows: $A|\eta\rangle = \langle\psi|\eta\rangle|\varphi\rangle$ for each vector $|\eta\rangle$. For example, in the n-dimensional space $\mathbb{C}^n$, an operator is represented by an $n \times n$ complex matrix A and $tr(A) = \sum_i A_{ii}$ (the sum of the entries on the main diagonal); if $|\varphi\rangle, |\psi\rangle \in \mathbb{C}^n$, then its external product is the multiplication $|\varphi\rangle\langle\psi|$ of column vector $|\varphi\rangle$ and the row vector $\langle\psi|$ (the adjoint, i.e. conjugate and transpose of $|\psi\rangle$). An operator A is positive if $\langle\psi|A|\psi\rangle \geq 0$ for every vector $|\psi\rangle$. A positive operator ρ in $\mathcal{H}$ is called a density operator if $tr(\rho) = 1$. Now a mixed state represented by ensemble $\mathbb{E}$ can be described by the density operator

$$\rho_{\mathbb{E}} = \sum_i p_i |\psi_i\rangle\langle\psi_i|;$$

in particular, a pure state $|\psi\rangle$ can be identified with the density operator $\rho = |\psi\rangle\langle\psi|$. For example, if a qubit is in state $|0\rangle$ with probability $\frac{2}{3}$ and in state $|+\rangle$ with probability $\frac{1}{3}$, then it can be modelled by the density matrix:

$$\rho = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|+\rangle\langle +| = \frac{1}{6}\begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}. \tag{1}$$

B. Quantum Operations

A basic operation on a quantum system is a quantum gate. Mathematically, a quantum gate is modelled by a unitary operator. An operator U in Hilbert space $\mathcal{H}$ is unitary if:

$$U^\dagger U = UU^\dagger = I,$$

where $U^\dagger$ is the adjoint of U, and I is the identity operator in $\mathcal{H}$. For example, the Hadamard gate

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{2}$$

is a unitary operator in the 2-dimensional Hilbert space. After it applies to a qubit in state $|0\rangle$ (respectively, $|1\rangle$), the qubit will be in state $|+\rangle$ (respectively, $|-\rangle$).

Another basic operation on a quantum system is a quantum measurement, which is essentially the way to extract information about the quantum system. In quantum computation, measurement is usually used to read out a computational result. Mathematically, a measurement is modelled as a set of operators $M = \{K_m\}$ with the normalisation condition:

$$\sum_m K_m^\dagger K_m = I.$$

If we perform a measurement M on a system in state ρ, then an outcome m is observed with probability

$$p_m = tr(K_m \rho K_m^\dagger),$$

and after that, the system will be in state:

$$\frac{K_m \rho K_m^\dagger}{p_m}.$$

For example, the measurement on a qubit (in the computational basis $|0\rangle, |1\rangle$) is $M = \{K_0, K_1\}$, where:

$$K_0 = |0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad K_1 = |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

If we perform M on a qubit in (mixed) state ρ given in equation (1) then the probability that we get outcome 0 is

$$p(0) = tr\left(K_0 \rho K_0^\dagger\right) = \frac{5}{6}$$

and then the qubit is in state $|0\rangle$. Similarly, outcome 1 is obtained with probability $p(1) = \frac{1}{6}$ and after that the qubit is in $|1\rangle$. It is worth noting that a major difference between classical and quantum systems occurs here: measuring a classical system does not change its state, whereas after measuring it, the state of the qubit is changed from ρ to either $|0\rangle$ or $|1\rangle$.

A general operation on a quantum system with state Hilbert space $\mathcal{H}$ can be modelled as a mapping $\mathcal{E}$ from density operators in $\mathcal{H}$ to themselves, which is completely positive and satisfies the condition:

$$(\text{Trace Preserving}): \quad tr(\mathcal{E}(\rho)) = tr(\rho)(= 1)$$

for all density operators ρ. Quantum operations are the discrete-time formalism of open quantum systems, i.e. systems interacting with their environments. Kraus representation theorem provides a more convenient way to deal with quantum operations: for each quantum operation $\mathcal{E}$, there is a family $\{E_i\}$ of linear operators such that

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$$

for all density operators ρ, where the Kraus operators E_i are required to satisfy the normalisation condition:

$$\sum_i E_i^\dagger E_i = I \text{ (the identity operator)}.$$

For example, a unitary operator U can be seen as the quantum operation $\mathcal{E}$ defined by

$$\mathcal{E}(\rho) = U\rho U^\dagger$$

for all density operators ρ, and after discarding the measurement outcomes, a quantum measurement $M = \{K_m\}$ can be seen as a quantum operation $\mathcal{E}$ defined by

$$\mathcal{E}(\rho) = \sum_m K_m \rho K_m^\dagger$$

for all density operators ρ.

C. Classical Differential Privacy

To motivate the notion of quantum differential privacy, we recall the definition of differential privacy for classical computing from [9].

Definition 1 (Classical Differential Privacy)**.** *A randomized function $\mathcal{K}$ gives (ε, δ)-differential privacy if for all data sets D and D' differing on one single participant, and for any set of possible outcomes $S \subseteq Range(\mathcal{K})$,*

$$\Pr[\mathcal{K}(D) \in S] \leq \exp(\varepsilon) \cdot \Pr[\mathcal{K}(D') \in S] + \delta. \qquad (3)$$

In the above definition, a classical computation process is modelled by a randomised function $\mathcal{K}$. A data set D is an input to the computation $\mathcal{K}$, and the computational result (i.e. output) is $\mathcal{K}(D)$, which belongs to $Range(\mathcal{K})$, the range of $\mathcal{K}$. Since $\mathcal{K}$ is a randomised function, it is reasonable to consider the probability $\Pr[\mathcal{K}(D) \in S]$ that the output is within a subset S of $Range(\mathcal{K})$ rather than an exact value of $\mathcal{K}(D)$. Now suppose that one of two data sets D and D' has a record of some participant, but the other has not. Then inequality (3) means that the difference between probabilities $\Pr[\mathcal{K}(D) \in S]$ and $\Pr[\mathcal{K}(D') \in S]$ is very small, and from them one cannot observe the difference between D and D'. Thus, the participant's privacy is preserved.

D. Quantum Differential Privacy

Now we are ready to formally define quantum differential privacy. To this end, we have several design decisions to make:

1) We will use a general density operator rather than a pure state to represent data in quantum computation. This generality is very helpful in applications. For example, in a quantum random access memory (QRAM) [19], [20], we can query a superposition of addresses to the address register $\sum_j \phi_j |j\rangle_a$, and the QRAM allows the data register d to obtain the data on each queried memory cell simultaneously and return a state formed by the address and data registers:

$$\sum_j \psi_j |j\rangle_a \xrightarrow{\text{QRAM}} \sum_j \psi_j |j\rangle_a |D_j\rangle_d$$

where $|D_j\rangle_d$ represents the state stored in the jth memory cell. If we discard the address register and only leave the data as an input to the quantum algorithm, it would be an ensemble of quantum states and should be represented by a density matrix.

2) In Definition 1, the difference between two classical data sets D and D' is be measured by l_1-norm. We will use trace distance to measure the difference between quantum data because it is a natural quantum generalisation of l_1-norm (see [18], Section 9.2.1). The trace distance between two density operators ρ and σ is defined to be

$$\tau(\rho, \sigma) = \frac{1}{2}\text{Tr}|\rho - \sigma|$$

where for an operator A, we define:

$$|A| = \sqrt{A^\dagger A}.$$

3) Roughly speaking, a quantum computation process or algorithm consists of a series of quantum gates (i.e. unitary transformations) and measurements with an input and an output of either classical information or quantum states. Obviously, it can be appropriately treated as an open quantum system. To simplify our model and capture the key point, we choose to describe a quantum computation process by a quantum operation $\mathcal{E}$.

4) The output $\mathcal{E}(\rho)$ of quantum computation $\mathcal{E}$ upon input ρ is a quantum state. So, we have to perform a measurement, say $M = \{K_m\}$, in order to acquire classical information from $\mathcal{E}(\rho)$. Since M happens at the end of the computational process, we are not concerned with its post-measurement states. Then M can be simplified as a POVM (Positive Operator-Valued Measure) $M = \{M_m\}$ with $M_m = K_m^\dagger K_m$ for every m because

$$p_m = tr(K_m \rho K_m^\dagger) = tr(M_m \rho)$$

(for details, see [18], Section 2.2.6). We write $Out(M) = \{m\}$ for the set of all possible outcomes of M. If we perform M on a quantum system in state ρ, then the probability of obtaining outcome m is $\text{Tr}[\rho M_m]$. For any subset $S \subseteq Out(M)$, we write $\Pr[\rho \in_M S]$ for the probability that the outcome falls into S when performing M on ρ; that is,

$$\Pr[\rho \in_M S] = \sum_{m \in S} \text{Tr}[\rho M_m].$$

The above design decisions straightforwardly lead to the following:

Definition 2 (Quantum Differential Privacy)**.** *Let $d \in (0, 1]$ and $\varepsilon, \delta > 0$ be three constants. A quantum operation $\mathcal{E}$ is (ε, δ)-differentially private if for every POVM $M = \{M_m\}$, for all $S \subseteq Out(M)$, and for all inputs ρ, σ such that $\tau(\rho, \sigma) \leq d$, it holds that*

$$\Pr[\mathcal{E}(\rho) \in_M S] \leq \exp(\varepsilon) \cdot \Pr[\mathcal{E}(\sigma) \in_M S] + \delta. \qquad (4)$$

In particular, if $\delta = 0$, we say that $\mathcal{E}$ is ε-differentially private.

In the above definition, the trace distance τ is chosen to characterize the difference between two adjacent databases ρ and σ. Indeed, it can be replaced by other distances between mixed quantum states. The reason for using it is that it is commonly used in quantum information theory and is easy to compute in some cases. Moreover, the parameter d appeared in the above definition, but its value is not essential because our theory of quantum differential privacy does not depends on a concrete value of d.

Similar to the case of classical computing, quantum differential privacy is immune to post-processing: without additional knowledge about the private databases, any quantum operation performed on the output of a quantum differentially private mechanism does not increase privacy loss.

Proposition 1 (Post-Processing)**.** *Let $\mathcal{E}$ be a quantum operation that is (ε, δ)-differentially private. Let $\mathcal{F}$ be an arbitrary quantum operation. Then the composition of $\mathcal{E}$ and $\mathcal{F}$:*

$$\mathcal{F} \circ \mathcal{E} : \rho \mapsto \mathcal{F}(\mathcal{E}(\rho))$$

is (ε, δ)-differentially private too.

The proof of this proposition is straightforward from Prop. 2 and Prop. 4 in Sec. V.

III. Quantum Differential Private Mechanism

To prevent that a user acquires private information contained in the input from the precise output, various privacy mechanisms have been proposed in the studies of differential privacy and its applications [21]. The basic idea is to deliberately introduce a certain noise to the output in order to protect the privacy. This idea can be easily generalized to the quantum case. Let ρ be the input. Then $\mathcal{E}(\rho)$ is the precise output of quantum computation modelled by quantum operation $\mathcal{E}$. The computation agent might not directly give the output $\mathcal{E}(\rho)$ to the user. Instead, it introduces a quantum noise, which is described by another quantum operation $\mathcal{E}_N$, at the end of the computation. Thus, the user will receive $\mathcal{E}_N(\mathcal{E}(\rho))$. Then the quantum noise mechanism for privacy-preserving computation can be modelled by quantum operation:

$$\mathcal{E}_N \circ \mathcal{E} : \rho \mapsto \mathcal{E}_N(\mathcal{E}(\rho)).$$

In this section, we consider the quantum privacy mechanisms using three different models of quantum noises: amplitude damping, phase damping and depolarizing, and compute their differential privacy parameters. Each of the three kind noises has a corresponding physical realization and can be actually applied. For example, generalized amplitude damping can be regarded as energy dissipation - an open system interacts with its environment (bath) of nonzero temperature [18], [22]. Phase damping can be realized as the evolution of a photon which scatters randomly through a waveguide [18], and an experiment realization using nuclear spin systems was reported by Leung et al. in [23]. Depolarizing channel is essentially a description of depolarizing phenomenon of mixed quantum states.

A. Generalized amplitude damping

We first examine how can differential privacy be protected by a generalized amplitude damping channel [18] for a single qubit, which is defined as the quantum operation

$$\mathcal{E}_{\mathrm{GAD}}(\rho) = \sum_{k=0}^{3} E_k \rho E_k^\dagger$$

in the 2-dimensional Hilbert space $\mathcal{H}_2$, where

$$E_0 = \sqrt{p}\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad E_1 = \sqrt{p}\begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix},$$
$$E_2 = \sqrt{1-p}\begin{bmatrix} \sqrt{1-\gamma} & 0 \\ 0 & 1 \end{bmatrix}, \quad E_3 = \sqrt{1-p}\begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix}$$

and p and γ are two parameters. Suppose that a quantum computation is modelled by a quantum operation $\mathcal{E} : \mathcal{H} \to \mathcal{H}_2$. Then the privacy-preserving computation with generalized amplitude damping as noise is depicted as the mapping:

$$\rho \mapsto \mathcal{M}_{\mathrm{GAD}}(\rho) = \mathcal{E}_{\mathrm{GAD}}[\mathcal{E}(\rho)].$$

To simplify the calculation, we only consider the case of $p = 0.5$ and regard γ as a parameter.

Theorem 1. *For all inputs ρ and σ with $\tau(\rho, \sigma) \leq d$, the generalized amplitude damping noise mechanism $\mathcal{M}_{\mathrm{GAD}}$ provides ε-differential private where*

$$\varepsilon = \ln\left[1 + \frac{2d\sqrt{1-\gamma}}{1-\sqrt{1-\gamma}}\right].$$

In particular, if $d/\gamma \ll 1$, then

$$\varepsilon \approx \frac{2d\sqrt{1-\gamma}}{1-\sqrt{1-\gamma}},$$

which is linear in d.

The privacy parameter ε given here is optimal.

The proof of this theorem is very technical and daunting. For readability, we postpone it to Subsection VI-A.

B. Composition of phase and amplitude damping

The phase damping channel [18] is defined by the operator-sum representation

$$\mathcal{E}_{\mathrm{PD}}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger$$

where

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix}, \quad E_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{bmatrix}$$

Let see how the phase damping can be combined with the amplitude damping to protect differential privacy. For a quantum computation modelled by quantum operation $\mathcal{E} : \mathcal{H} \to \mathcal{H}_2$, the phase-amplitude damping mechanism can be depicted by the mapping:

$$\rho \mapsto \mathcal{M}_{\mathrm{PAD}}(\rho) = \mathcal{E}_{\mathrm{GAD}}[\mathcal{E}_{\mathrm{PD}}(\mathcal{E}(\rho))].$$

Theorem 2. *Let $p = 0.5$ in $\mathcal{E}_{\mathrm{GAD}}$. If the parameter γ in $\mathcal{E}_{\mathrm{GAD}}$ and λ in $\mathcal{E}_{\mathrm{PD}}$ satisfy $\lambda \leq \gamma$, then for all inputs ρ and σ such*

that $\tau(\rho,\sigma) \leq d$, the phase-amplitude damping mechanism $\mathcal{M}_{\rm PAD}$ provides ε-differential private where

$$\varepsilon = \ln\left[1 + \frac{2d\sqrt{1-\gamma}\sqrt{1-\lambda}}{1-\sqrt{1-\gamma}\sqrt{1-\lambda}}\right].$$

We omit the proof of the above theorem because it is almost the same as the proof of Theorem 1.

C. Depolarizing mechanism

Depolarizing operation is an important type of quantum noise which can be represented by the quantum operation:

$$\mathcal{E}_{\rm Dep}(\rho) = \frac{pI}{D} + (1-p)\rho$$

where D is the dimension of the state Hilbert space and p is the probability parameter. Let a quantum computation be described by quantum operation $\mathcal{E} : \mathcal{H} \to \mathcal{H}_D$, where $\mathcal{H}_D$ is the D-dimensional Hilbert space. Then the depolarizing mechanism is defined as the mapping:

$$\rho \mapsto \mathcal{M}_{\rm Dep}(\rho) = \mathcal{E}_{\rm Dep}(\mathcal{E}(\rho)).$$

Theorem 3. *For all inputs ρ and σ such that $\tau(\rho,\sigma) \leq d$, the depolarizing mechanism $\mathcal{M}_{\rm Dep}$ in the D-dimension Hilbert space provides ε-differential private where*

$$\varepsilon = \ln\left[1 + \frac{1-p}{p} dD\right].$$

If $dD/p \ll 1$, then

$$\varepsilon \approx (1-p)dD/p,$$

which is linear in d.

For readability, the proof of this theorem is deferred to Subsection VI-B.

D. Comparison between GAD, PAD and Dep mechanisms

The differential privacy parameters were computed for the three mechanisms $\mathcal{M}_{\rm GAD}$, $\mathcal{M}_{\rm PAD}$ and $\mathcal{M}_{\rm Dep}$. Now, let us briefly compare the differences between the disturbed output in these mechanisms and the exact output in the 2-dimensional case. For any input ρ, assume the exact output is

$$\mathcal{E}(\rho) = \begin{bmatrix} a & b \\ b^* & c \end{bmatrix}$$

then the trace distances between the disturbed output and the exact output are given as follows:

$$\tau(\mathcal{E}(\rho), \mathcal{M}_{\rm GAD}(\rho)) = \sqrt{b^2(1-\sqrt{1-\gamma_g})^2 + (a-c)^2\gamma_g^2/4};$$

$$\tau(\mathcal{E}(\rho), \mathcal{M}_{\rm PAD}(\rho)) = \sqrt{b^2(1-\sqrt{1-\gamma_p}\sqrt{1-\lambda})^2 + (a-c)^2\gamma_p^2/4};$$

$$\tau(\mathcal{E}(\rho), \mathcal{M}_{\rm Dep}(\rho)) = \sqrt{b^2p^2 + (a-c)^2p^2/4}.$$

Here, the parameter γ in the GAD mechanism $\mathcal{M}_{\rm GAD}$ and in the PAD mechanism $\mathcal{M}_{\rm PAD}$ are denoted by γ_g and γ_p, respectively. If we set:

$$1-\gamma_g = (1-\gamma_p)(1-\lambda),$$

the differential privacy parameter ε for GAD and PAD given in Theorems 1 and 2 are the same. It is easy to see that

$$\tau(\mathcal{E}(\rho), \mathcal{M}_{\rm GAD}(\rho)) \geq \tau(\mathcal{E}(\rho), \mathcal{M}_{\rm PAD}(\rho)).$$

So, we conclude that whenever PAD and GAD preserve the same level of privacy, the disturbance of PAD is smaller than that of GAD; that is, PAD gives a higher authenticity than GAD. We further see that whenever $\gamma_p = \lambda$, PAD reaches the highest authenticity. On the other hand, we note that the PAD mechanism and the depolarizing mechanism are actually the same when $\gamma_p = \lambda = p$. Therefore, for the most cases, the disturbance of output in the three mechanisms has the following ordering:

$$\tau(\mathcal{E}(\rho), \mathcal{M}_{\rm Dep}(\rho)) \leq \tau(\mathcal{E}(\rho), \mathcal{M}_{\rm PAD}(\rho)) \leq \mathcal{M}_{\rm GAD}(\rho)).$$

E. An Illustrative Example

To conclude this section, we give a simple example that shows how the quantum privacy mechanisms presented above work. Suppose the curator holds a database which stores one qubit information; that is, a quantum state $|\psi\rangle \in {\rm Span}\{|0\rangle, |1\rangle\}$, for each participant. The curator only accepts the query of averaging and she/he responds with a quantum state of which the density operator is:

$$\mathcal{E}(\rho) = \frac{1}{n}\sum_i |\psi_i\rangle\langle\psi_i|$$

where ρ represents the database, n is the number of participants, $|\psi_i\rangle$ is the information of ith participants, and the sum is over all participants.

To convince a potential participant A that joining this database would not leak his personal information, we need to ensure that it is almost impossible to decide whether an individual A participates or not from the curator's response for any given query. So, we consider the worst case that there is a very powerful adversary who knows the information of all the participants except A. We further assume that A's information is represented by state $|1\rangle$ and all of the others' are represented by $|0\rangle$. Thus, without A's information, we have $\mathcal{E}(\rho_1) = |0\rangle\langle 0|$. If A joins the database, then

$$\mathcal{E}(\rho_2) = \frac{n-1}{n}|0\rangle\langle 0| + \frac{1}{n}|1\rangle\langle 1|.$$

The adversary has a certain probability to fully distinguish ρ_1 and ρ_2; for example, she/he can use POVM $\{M_{yes}, M_{no}|\}$ to measure the respond state, where $M_{yes} = |0\rangle\langle 0|$, $M_{no} = |1\rangle\langle 1|$. If the measurement result is "no", then the adversary knows that A participates in the database.

To avoid this, one can use, for example, the depolarizing mechanism and choose the parameter $p = \frac{2}{n+2}$. Then:

$$\mathcal{M}_{\mathrm{Dep}}(\rho_1) = \begin{pmatrix} \frac{n+1}{n+2} & 0 \\ 0 & \frac{1}{n+2} \end{pmatrix},$$
$$\mathcal{M}_{\mathrm{Dep}}(\rho_2) = \begin{pmatrix} \frac{n}{n+2} & 0 \\ 0 & \frac{2}{n+2} \end{pmatrix}.$$

Now, it is impossible to fully distinguish $\mathcal{M}_{\mathrm{Dep}}(\rho_1)$ and $\mathcal{M}_{\mathrm{Dep}}(\rho_2)$ with a small number of copies of the response. Using Theorem 3 we obtain the differential privacy parameter

$$\varepsilon = \ln\left[1 + \frac{1-p}{p} \cdot \frac{1}{n} \cdot 2\right] = \ln 2,$$

which means that, for any POVM $M = \{M_m\}$ and any possible outcome m, we have:

$$\frac{1}{2} \leq \frac{p_m[\mathcal{M}_{\mathrm{Dep}}(\rho_1)]}{p_m[\mathcal{M}_{\mathrm{Dep}}(\rho_2)]} \leq 2.$$

The above inequality actually characterizes the difference between computational outcomes for ρ_1 and ρ_2 in the worst case.

IV. Composition Theorems

In the previous sections, we only considered the differential privacy for a single quantum operation. In many practical applications, however, we often need to deal with much more complicated situations where a data user may combine the responses of several queries based on different databases, and then perform the linkage attacks or re-identification [21] to discover personal information. Thus, it is desirable to establish some laws for the differential privacy of the combination of several quantum operations. Such laws will be very useful in the design of sophisticated quantum algorithms for privacy preserving data analytics.

The aim of this section is to prove several combination theorems for quantum differential privacy. These theorems are quantum generalisations of the combination theorems for classical differential privacy [21]. Let us start from the simple case of ε-differential privacy (without parameter δ).

Theorem 4. *Let $\mathcal{M}_1 : \mathcal{H}_1 \mapsto \mathcal{H}_A$ be an ε_1-differentially private quantum algorithm, and let $\mathcal{M}_2 : \mathcal{H}_2 \mapsto \mathcal{H}_B$ be an ε_2-differentially private quantum algorithm, and assume that they are independent. Their combination*

$$\mathcal{M}_{1,2} : \mathcal{H}_1 \otimes \mathcal{H}_2 \mapsto \mathcal{H}_A \otimes \mathcal{H}_B$$

is defined by

$$\mathcal{M}_{1,2}(\rho, \sigma) = \mathcal{M}_1(\rho) \otimes \mathcal{M}_2(\sigma)$$

for all density operators ρ in $\mathcal{H}_1$ and σ in $\mathcal{H}_2$. Then $\mathcal{M}_{1,2}$ is $(\varepsilon_1 + \varepsilon_2)$-differentially private.

Before we start to prove the theorem, let us make some clarifications. First, the input ρ and σ of $\mathcal{M}_{1,2}$ should not be entangled because they are assumed to be independent of each other. Second, two input states $\varphi = \rho \otimes \sigma$ and $\varphi' = \rho' \otimes \sigma'$ of $\mathcal{M}_{1,2}$ are considered as adjacent if and only if ρ and ρ' are adjacent databases, and σ and σ' are adjacent. Moreover, the parameter d in Def. 2 is not essential here, as we only care about when the databases inputted to $\mathcal{M}_{1,2}$ are adjacent.

Proof. Suppose $\rho_1 = \mathcal{M}_1(\rho)$ and $\rho_2 = \mathcal{M}_1(\rho')$, where quantum states ρ and ρ' are used to denote two adjacent databases. The same way, assume $\sigma_1 = \mathcal{M}_2(\sigma)$ and $\sigma_2 = \mathcal{M}_2(\sigma')$, where σ and σ' represent two adjacent databases. Now, we prove that for any pure state $|\phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, we have:

$$\mathrm{Tr}[\rho_1 \otimes \sigma_1 |\phi\rangle\langle\phi|] \leq e^{\varepsilon_1 + \varepsilon_2} \mathrm{Tr}[\rho_2 \otimes \sigma_2 |\phi\rangle\langle\phi|],$$

which directly leads to the theorem. Using the Schmidt decomposition theorem, we can write:

$$|\phi\rangle = \sum_i p_i |\alpha_i\rangle |\beta_i\rangle$$

where $|\alpha_i\rangle$ and $|\beta_i\rangle$ are orthonormal states for system A and B, respectively. We also decompose σ_1 in these bases as follows:

$$\begin{aligned}\sigma_1 &= \sum_k r_k \left(\sum_l b_{kl} |\beta_l\rangle \sum_{l'} b^\dagger_{kl'} \langle\beta_{l'}| \right) \\ &= \sum_k r_k \sum_{ll'} b_{kl} b^\dagger_{kl'} |\beta_l\rangle\langle\beta_{l'}|.\end{aligned}$$

Then we have:

$$\begin{aligned}&\mathrm{Tr}[\rho_1 \otimes \sigma_1 |\phi\rangle\langle\phi|] = \sum_{ij} p_i p_j^\dagger \langle\alpha_j|\rho_1|\alpha_i\rangle\langle\beta_j|\sigma_1|\beta_i\rangle \\ &= \sum_{ij} p_i p_j^\dagger \langle\alpha_j|\rho_1|\alpha_i\rangle \sum_k r_k \sum_{ll'} b_{kl} b^\dagger_{kl'} \langle\beta_j|\beta_l\rangle\langle\beta_{l'}|\beta_i\rangle \\ &= \sum_k r_k \sum_{ij} (p_i b^\dagger_{ki})(p_j b^\dagger_{kj})^\dagger \langle\alpha_j|\rho_1|\alpha_i\rangle \\ &= \sum_k r_k \mathrm{Tr}[\rho_1 |\psi_k\rangle\langle\psi_k|] \\ &\leq \sum_k r_k e^{\varepsilon_1} \mathrm{Tr}[\rho_2 |\psi_k\rangle\langle\psi_k|] \\ &= e^{\varepsilon_1} \mathrm{Tr}[\rho_2 \otimes \sigma_1 |\phi\rangle\langle\phi|] \\ &\leq e^{\varepsilon_1 + \varepsilon_2} \mathrm{Tr}[\rho_2 \otimes \sigma_2 |\phi\rangle\langle\phi|].\end{aligned}$$

Here, the states

$$|\psi_k\rangle = \sum_i p_i b^\dagger_{ki} |\alpha_i\rangle$$

for every k. □

As a straightforward corollary of the above theorem, we can combine more than two quantum algorithms.

Corollary 1. *For each $i \in [k]$, let $\mathcal{M}_i : \mathcal{H}_i \mapsto \mathcal{H}'_i$ be an ε_i-differentially private quantum algorithm, and assume that they are independent. Their combination*

$$\mathcal{M}_{[k]} : \bigotimes_{i=1}^k \mathcal{H}_i \mapsto \bigotimes_{i=1}^k \mathcal{H}'_i$$

is defined by

$$\mathcal{M}_{[k]}(\rho_1, \rho_2, \cdots, \rho_k) = \bigotimes_{i=1}^{k} \mathcal{M}_i(\rho_i)$$

for all density operators ρ_i in $\mathcal{H}_i$. Then $\mathcal{M}_{[k]}$ is $(\sum_{i=1}^{k} \varepsilon_i)$-differentially private.

We now extend Theorem 4 to the general case of (ε, δ)-differential privacy.

Theorem 5. *Let $\mathcal{M}_1$ be an $(\varepsilon_1, \delta_1)$-differentially private quantum algorithm and $\mathcal{M}_2$ an $(\varepsilon_2, \delta_2)$-differentially private quantum algorithm, and assume that they are independent. Then their combination $\mathcal{M}_{1,2}(\rho, \sigma)$ is $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-differentially private.*

The proof of the above theorem requires a technical lemma, which is essentially a quantum generalisation of Lemma 3.17(1) in [21].

Lemma 1. *Consider two arbitrary density operators ρ and σ satisfying the condition:*

$$\mathrm{Tr}[\rho M] \leq e^{\varepsilon} \mathrm{Tr}[\sigma M] + \delta$$

for any positive semi-definite matrices $M \leq \mathbb{I}$. Then there exists a density operator ϱ such that for any positive semi-definite matrices $M \leq \mathbb{I}$, the following two conditions always holds:

$$\mathrm{Tr}[\rho M] \leq \mathrm{Tr}[\varrho M] + \delta \Leftrightarrow \tau(\rho, \varrho) \leq \delta$$
$$\mathrm{Tr}[\varrho M] \leq e^{\varepsilon} \mathrm{Tr}[\sigma M]$$

Proof. We first present Algorithm 1 to find such ϱ and then prove that it satisfies the above conditions. Using an orthonormal basis $|\psi_1\rangle, \cdots, |\psi_n\rangle$, operator $\rho - e^{\varepsilon}\sigma$ can be diagonalised, as so ρ and σ have the property: $\rho_{ij} = e^{\varepsilon}\sigma_{ij}$ for all $i \neq j$. Furthermore, we perform a quantum measurement

Algorithm 1: Separation of ε and δ

1. Given inputs ρ and σ and ε, δ.
2. `// ρ and σ are two density operators in Hilbert space H.`
3. Let n be the dimension of the Hilbert space $\mathcal{H}$.
4. Let $|\psi_i\rangle$ be the eigenvectors of $\rho - e^{\varepsilon}\sigma$.
5. Let $p_i = \mathrm{Tr}[\rho|\psi_i\rangle\langle\psi_i|]$, and $q_i = \mathrm{Tr}[\sigma|\psi_i\rangle\langle\psi_i|]$.
6. For two distributions $P = \{p_i\}$ and $Q = \{q_i\}$, find the distribution $R = \{r_i\}$ s.t.:
$$\Delta(P, R) \leq \delta \ , \quad D_{\infty}(R||Q) \leq \varepsilon.$$
7. Let $\varrho = \rho + \sum_{i=1}^{n}(r_i - p_i)|\psi_i\rangle\langle\psi_i|$.
8. **Return** ϱ;

$\{M_i = |\psi_i\rangle\langle\psi_i|\}$ on states ρ and σ in the same basis. Suppose that the distributions of outcomes are $P = \{p_i\}$ and $Q = \{q_i\}$, respectively. Then these two distributions satisfy:

$$\sum_{i \in S} p_i \leq e^{\varepsilon} \sum_{i \in S} q_i + \delta$$

for all $S \subseteq [n]$. Adopting the definitions of max divergence D_{∞}, δ-approximate max divergence D_{∞}^{δ} and δ-close Δ in [21], we have:

$$D_{\infty}^{\delta}(P||Q) \leq \varepsilon.$$

So, there exists $R = \{r_i\}$ such that $\Delta(R, P) \leq \delta$ and $D_{\infty}(R||Q) \leq \varepsilon$ according to Lemma 3.17 in [21]. We now see that

$$\begin{aligned}\tau(\rho, \varrho) &= \frac{1}{2}\mathrm{Tr}|\rho - \varrho| \\ &= \frac{1}{2}\mathrm{Tr}\left|\sum_{i=1}^{n}(r_i - p_i)|\psi_i\rangle\langle\psi_i|\right| \\ &= \Delta(R, P) \leq \delta.\end{aligned}$$

Moreover, for all $|\phi\rangle = \sum_{i=1}^{n} \lambda_i |\psi_i\rangle$, we have:

$$\begin{aligned}\mathrm{Tr}[\varrho|\phi\rangle\langle\phi|] =& \sum_{i=1}^{n} \lambda_i \bar{\lambda_i} \varrho_{ii} + \sum_{i \neq j} \lambda_i \bar{\lambda_j} \varrho_{ij} \\ =& \sum_{i=1}^{n} \lambda_i \bar{\lambda_i} r_i + \sum_{i \neq j} \lambda_i \bar{\lambda_j} \rho_{ij} \\ \leq& \sum_{i=1}^{n} \lambda_i \bar{\lambda_i} e^{\varepsilon} q_i + \sum_{i \neq j} \lambda_i \bar{\lambda_j} e^{\varepsilon} \sigma_{ij} \\ =& \sum_{i=1}^{n} \lambda_i \bar{\lambda_i} e^{\varepsilon} \sigma_{ii} + \sum_{i \neq j} \lambda_i \bar{\lambda_j} e^{\varepsilon} \sigma_{ij} \\ =& e^{\varepsilon}\mathrm{Tr}[\sigma|\phi\rangle\langle\phi|]\end{aligned}$$

Finally, combining the above two inequalities immediately yields the lemma. □

Now, we are ready to prove Theorem 5.

Proof of Theorem 5. Let ρ_1 and σ_1 be the outputs of $\mathcal{M}_1$ of two adjacent databases, and let ρ_2 and σ_2 be the outputs of $\mathcal{M}_2$ of other two adjacent databases. Using Algorithm 1, we can first find ϱ_1 and ϱ_2 such that:

$$\begin{aligned}&\tau(\rho_1, \varrho_1) \leq \delta_1, \ \tau(\rho_2, \varrho_2) \leq \delta_2, \\ &\forall\ 0 \leq M \leq \mathbb{I}_A, \ \mathrm{Tr}[\varrho_1 M] \leq e^{\varepsilon_1}\mathrm{Tr}[\sigma_1 M], \\ &\forall\ 0 \leq M \leq \mathbb{I}_B, \ \mathrm{Tr}[\varrho_2 M] \leq e^{\varepsilon_2}\mathrm{Tr}[\sigma_2 M].\end{aligned}$$

Then for any measurement M in $\mathcal{H}_A \otimes \mathcal{H}_B$ $(0 \leq M \leq \mathbb{I}_{AB})$, using Theorem 4 we obtain:

$$\begin{aligned}&\mathrm{Tr}[\rho_1 \otimes \rho_2 M] \\ &= \mathrm{Tr}[\varrho_1 \otimes \varrho_2 M] + \{\mathrm{Tr}[\rho_1 \otimes \rho_2 M] - \mathrm{Tr}[\varrho_1 \otimes \varrho_2 M]\} \\ &\leq e^{\varepsilon_1 + \varepsilon_2}\mathrm{Tr}[\sigma_1 \otimes \sigma_2 M] + \tau(\rho_1 \otimes \rho_2, \varrho_1 \otimes \varrho_2) \\ &\leq e^{\varepsilon_1 + \varepsilon_2}\mathrm{Tr}[\sigma_1 \otimes \sigma_2 M] + \tau(\rho_1, \varrho_1) + \tau(\rho_2, \varrho_2) \\ &= e^{\varepsilon_1 + \varepsilon_2}\mathrm{Tr}[\sigma_1 \otimes \sigma_2 M] + (\delta_1 + \delta_2).\end{aligned}$$

Symmetrically, we can prove:

$$\mathrm{Tr}[\sigma_1 \otimes \sigma_2 M] \leq e^{\varepsilon_1 + \varepsilon_2}\mathrm{Tr}[\rho_1 \otimes \rho_2 M] + (\delta_1 + \delta_2)$$

for all measurements M in $\mathcal{H}_A \otimes \mathcal{H}_B$ $(0 \leq M \leq \mathbb{I}_{AB})$. Thus, we complete the proof. □

Theorem 5 can also be easily generalised to the case of combining more than two quantum algorithms.

Corollary 2. *For each $i \in [k]$, let $\mathcal{M}_i : \mathcal{H}_i \mapsto \mathcal{H}'_i$ be an $(\varepsilon_i, \delta_i)$-differentially private quantum algorithm, and assume that they are independent. Then their combination $\mathcal{M}_{[k]}$ is $(\sum_{i=1}^k \varepsilon_i, \sum_{i=1}^k \delta_i)$-differentially private.*

A. Advanced Composition Theorem

Theorems 4 and 5 show that once an individual shares her/his information in several databases, the risk of leaking his information will increase; more precisely, the differential privacy parameter ε and δ will grow linearly in the number of databases. Of course, if these parameters can degrade more slowly, then we can design quantum algorithms with less noise and preserve more useful information. Dwork and Roth [21] introduced an adaptive combination of classical algorithms and proved an advanced composition theorem that allows the privacy parameters to degrade much slower. In this subsection, we generalise this advanced composition theorem into a quantum setting where a certain measurement is performed at each step.

Let us first formally define the computational model. Assume that $\mathcal{F}$ is a set of quantum algorithms. We introduce two experiments that can be seen as adaptive combination of quantum algorithms in $\mathcal{F}$.

Definition 3 (Experiment $b \in \{0,1\}$)**.** *The experiment b consists of k rounds of data access experiment. For each $i = 1, \cdots, k$:*

1) *The input at step i is two adjacent databases represented by quantum states (density operators) σ_i^0 and σ_i^1.*
2) *The adversary A chooses:*
   a) *a quantum algorithm $\mathcal{M}_i \in \mathcal{F}$;*
   b) *a POVM P_i used to measure the output quantum state; and*
   c) *parameters w_i^0 and w_i^1. The parameter w_i^b is determined by the previous results; that is,*

$$w_i^b = f(y_1^b, y_2^b, \cdots, y_{i-1}^b).$$

3) *The adversary A receives a quantum state*

$$\rho_i^b = \mathcal{M}_i(w_i^b, \sigma_i^b),$$

performs the POVM P_i, and gains the result y_i^b.

It is clear that in the experiments defined above, the adversary A is given the most powerful ability to influence the databases and queries. So, these experiments can be regarded as the worst case where no more information can be leaked. The situation considered in the above definition is referred as the *measure-each-step scenario* because at each round the adversary A performs a POVM.

For each $b \in \{0,1\}$, the view V^b of adversary A in experiment b is defined to be the vector $(y_1^b, y_2^b, \cdots, y_k^b)$ of the results at all steps.

Definition 4. *We say the set $\mathcal{F}$ of quantum algorithms is (ε, δ)-differential private under k-fold adaptive composition if for any adversary A, we have:*

$$D_\infty^\delta(V^0||V^1) \leq \varepsilon$$

where $D_\infty^\delta(Y||Z)$ stands for the δ-max divergence of random variables Y and Z; that is,

$$D_\infty^\delta(Y||Z) = \max_{\substack{S \subseteq \mathrm{Supp}(Y); \\ \Pr[Y \in S] \geq \delta}} \ln \frac{\Pr[Y \in S] - \delta}{\Pr[Z \in S]}.$$

Theorem 6 (Advanced Composition Theorem for Measure-Each-Step Scenario)**.** *Let $\varepsilon, \delta, \delta' \geq 0$, and let $\mathcal{F}$ be a set of quantum algorithms. If each algorithm in $\mathcal{F}$ is (ε, δ)-differentially private, then $\mathcal{F}$ is $(\varepsilon', k\delta + \delta')$-differential private under k-fold adaptive composition, where:*

$$\varepsilon' = \sqrt{2k \ln(1/\delta')}\varepsilon + k\varepsilon(e^\varepsilon - 1).$$

Proof. If we combine the quantum algorithm and POVM together, with the same parameter $w_i^0 = w_i^1$, then we know that

$$D_\infty^\delta(y_i^0||y_i^1) \leq \varepsilon, \qquad D_\infty^\delta(y_i^1||y_i^0) \leq \varepsilon.$$

The rest part of the proof is exactly the same as that for the classical case given in [21]. □

V. Algorithm for Computing Quantum Privacy Parameters

In Section III, we gave the privacy parameters λ and δ for three simple quantum noise models. In general, however, it is very hard to compute the privacy parameters ε and δ directly using Definition 2. In this section, we develop an algorithm for the calculation of privacy parameters for a general quantum operation.

A. Proportional Distance

Let us first prepare a very useful mathematical tool.

Definition 5 (Proportional Distance)**.** *For two density operators ρ and σ, their proportional distance is defined as:*

$$PD(\rho, \sigma) = \sup_{\{M_m\}, m} \max \left(\ln \frac{q_m}{p_m}, \ln \frac{p_m}{q_m} \right) \tag{5}$$

where the supremum is over all POVMs $\{M_m\}$ and all possible outcomes m,

$$p_m \equiv \mathrm{Tr}[\rho M_m], \ q_m \equiv \mathrm{Tr}[\sigma M_m]$$

are the probabilities of obtaining outcome m when the measurement is performed on ρ, σ, respectively, and we make the conventions $\frac{0}{0} = 1$, $\frac{1}{0} = +\infty$, $\ln 0 = -\infty$, $\ln +\infty = +\infty$.

From the defining equation (4) of quantum differential privacy, we see that a key step to compute privacy parameters ε and δ is to evaluate the quantity

$$\ln \frac{\Pr[\mathcal{E}(\rho) \in_M S]}{\Pr[\mathcal{E}(\sigma) \in_M S]}$$

which can often be done by calculating the proportional distance between $\mathcal{E}(\rho)$ and $\mathcal{E}(\sigma)$, as clearly indicated by equation (5).

Some basic properties of proportional distance are collected in the following:

Proposition 2. 1) *PD is a distance; that is,*

a) *$PD(\rho,\sigma) \geq 0$, and $PD(\rho,\sigma) = 0$ if and only if $\rho = \sigma$.*

b) *$PD(\rho,\sigma) = PD(\sigma,\rho)$.*

c) *Triangle inequality:*

$$PD(\rho,\sigma) \leq PD(\rho,\delta) + PD(\delta,\sigma).$$

2) *PD is preserved by unitary transformations, and every quantum operation $\mathcal{E}$ is contractive with respect to PD; that is,*

$$PD(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leq PD(\rho,\sigma).$$

The second part of this proposition shows that no physical process can increase the proportional distance between two quantum states. A similar result for a different distance between quantum states, namely trace distance, was presented in [18], Theorem 9.2.

The proof of the first part of the above proposition is trivial, and the second part is a special case of Proposition 4 below.

As we pointed out before, proportional distance PD can be used to compute the differential privacy parameters ε and δ for quantum operations. However, PD itself is also not easy to compute if we simply use Definition 5. The following proposition gives a way for computing the proportional distance much simpler than its definition.

Proposition 3.

$$PD(\rho,\sigma) = \sup_{|\psi\rangle} \left| \ln \frac{\langle\psi|\rho|\psi\rangle}{\langle\psi|\sigma|\psi\rangle} \right| \tag{6}$$

where the supremum is over all pure states.

Proof. For any POVM $M = \{M_m\}$ and any outcome m of M, M_m is a positive operator. So, M_m can be diagonalized as

$$M_m = \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

where $0 < p_i \leq 1$ (and $\sum_i p_i \leq 1$). Then we have:

$$\begin{aligned} \frac{\mathrm{Tr}[\rho M_m]}{\mathrm{Tr}[\sigma M_m]} &= \frac{\sum_i p_i \mathrm{Tr}[\rho P_i]}{\sum_i p_i \mathrm{Tr}[\sigma P_i]} \\ &\leq \max_i \frac{\mathrm{Tr}[\rho P_i]}{\mathrm{Tr}[\sigma P_i]} \end{aligned}$$

where $P_i = |\psi_i\rangle\langle\psi_i|$. This directly leads to equation (6). □

B. Computing Proportional Distance

As discussed in the above subsection, computation of privacy parameters ε, δ depends heavily on the calculation of the proportional distance of quantum states. Furthermore, Proposition 3 provides a way to compute proportional distance PD. But actually, it is mainly useful for the theoretical calculation.

In this subsection, we present an efficient algorithm that can compute approximate values of PD in practical applications. To simplify the presentation, let us first introduce a notation:

$$dPD(\rho,\sigma) = \sup_{0<M\leq\mathbb{I}} \ln \frac{\mathrm{Tr}(M\rho)}{\mathrm{Tr}(M\sigma)}.$$

Obviously, we have:

$$PD(\rho,\sigma) = \max(dPD(\rho,\sigma), dPD(\sigma,\rho)). \tag{7}$$

Now we can present Algorithm 2, which efficiently calculate dPD with an error smaller than θ. Then the proportional distance PD between ρ and σ can be computed using equation (7).

Remark 1. *Essentially, calculating PD is a Linear Programming Problem and can be solved in polynomial time (w.r.t. the dimension of density operator). It can be done using cvx toolkit based on convex geometry [24]. For a fixed accuracy θ, we guess that the complexity of Algorithm 2 is $O(N^3 \log N)$, where N is the dimension of the density operators ρ, σ and the factor N^3 comes from that of diagonalizing the matrices, but we failed to prove it.*

Algorithm 2: Computing dPD

Function DPD (ρ, σ, θ)
`// ρ and σ are two density operators in Hilbert space H and θ is the desired accuracy.`
Let n be the dimension of the Hilbert space $\mathcal{H}$.
Set real number $max = 0, \lambda = \infty$.
while $|max - \lambda| > \theta$ **do**
  $\lambda = max$;
  $max = \max_i(\rho_{ii}/\sigma_{ii})$;
  $\eta = \rho - max * \sigma$;
  Diagonalize η: $\eta = PDP^\dagger$;
  `// P is unitary and D is diagonal matrix.`
  $\rho = P^\dagger \rho P$;
  $\sigma = P^\dagger \sigma P$;
end
Return $\ln max$;

C. Accuracy of Sampling Inputs

For a given pair ρ, σ, we can compute $PD(\rho,\sigma)$ using Algorithm 2. However, it is clear from Definition 2 that computing the privacy parameters ε, δ requires us to consider all adjacent inputs ρ, σ (i.e. input pairs ρ, σ with $\tau(\rho,\sigma) \leq d$). To address this issue, our strategy is to find a sample set $\mathcal{P}'$ to represent $\mathcal{P}$, where $\mathcal{P}$ denotes the set of all possible inputs. Now the question is: how accurate can be the computed values of ε, δ by a sampling of inputs?

To answer the above question, we need a notion of δ-approximation of proportional distance.

Definition 6 (δ-Proportional Distance)**.** *Using the same notations as in Def. 5, δ-proportional distance is defined to be:*

$$PD^{\delta}(\rho,\sigma) = \sup_{\substack{\{M_m\},\\ \mathcal{S}\subseteq\{m\}}} \ln\max\left(\frac{\sum_{m\in\mathcal{S}} q_m-\delta}{\sum_{m\in\mathcal{S}} p_m}, \frac{\sum_{m\in\mathcal{S}} p_m-\delta}{\sum_{m\in\mathcal{S}} q_m}, 0\right).$$

Obviously, if $\delta = 0$, then

$$PD^{\delta}(\rho,\sigma) = PD(\rho,\sigma).$$

The following proposition is a generalisation of clause 2) of Proposition 2 , and it shows that any physical process cannot increase the δ-proportional distance between two quantum states.

Proposition 4. *PD^{δ} is preserved by unitary transformations, and every quantum operation $\mathcal{E}$ is contractive with respect to PD^{δ}; that is,*

$$PD^{\delta}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leq PD^{\delta}(\rho,\sigma).$$

Proof. Let $\mathcal{E}(\rho) = \sum_k E_k\rho E_k^{\dagger}$ be an operator-sum representation of $\mathcal{E}$. For an arbitrary POVM $M = \{M_m\}$, we define:

$$Q_{mk} = E_k^{\dagger}M_m^{\dagger}M_mE_k$$

for every m,k. Then $\sum_{k,m} Q_{mk} = I$ and $Q = \{Q_{mk}\}$ is a POVM too. We assume that $\epsilon = PD^{\delta}(\rho,\sigma) < \infty$, and write:

$$p_{mk} = \mathrm{Tr}[\rho Q_{mk}], \qquad q_{mk} = \mathrm{Tr}[\rho Q_{mk}].$$

Then for any subset $\mathcal{S}\subseteq Out(M)$, note that

$$\{mk | m\in\mathcal{S}, k\in\{k\}\} \subseteq Out(Q)$$

and so we obtain:

$$\begin{aligned} \frac{\sum_{m\in\mathcal{S},k} p_{mk}-\delta}{\sum_{m\in\mathcal{S},k} q_{mk}} &\leq e^{\epsilon}, \\ \frac{\sum_{m\in\mathcal{S},k} q_{mk}-\delta}{\sum_{m\in\mathcal{S},k} p_{mk}} &\leq e^{\epsilon} \end{aligned} \tag{8}$$

Now we perform POVM P on $\mathcal{E}(\rho)$ and $\mathcal{E}(\sigma)$. For every m, we have:

$$\begin{aligned} p'_m &= \mathrm{Tr}[\mathcal{E}(\rho)M_m] \\ &= \mathrm{Tr}\left[\sum_k E_k\rho E_k^{\dagger}M_m\right] \\ &= \sum_k \mathrm{Tr}\left[E_k\rho E_k^{\dagger}M_m\right] \\ &= \sum_k \mathrm{Tr}\left[\rho Q_{mk}\right] = \sum_k p_{mk} \end{aligned}$$

and

$$q'_m = \mathrm{Tr}[\mathcal{E}(\rho)M_m] = \sum_k q_{mk}.$$

It follows from equation (8) that for any $\mathcal{S}\subseteq Out(M)$

$$\begin{aligned} \frac{\sum_{m\in\mathcal{S}} p'_m-\delta}{\sum_{m\in\mathcal{S}} q'_m} &= \frac{\sum_{m\in\mathcal{S},k} p_{mk}-\delta}{\sum_{m\in\mathcal{S},k} q_{mk}} \leq e^{\epsilon}, \\ \frac{\sum_{m\in\mathcal{S}} q'_m-\delta}{\sum_{m\in\mathcal{S}} p'_m} &= \frac{\sum_{m\in\mathcal{S},k} q_{mk}-\delta}{\sum_{m\in\mathcal{S},k} p_{mk}} \leq e^{\epsilon} \end{aligned}$$

and $PD^{\delta}(\mathcal{E}(\rho),\mathcal{E}(\sigma)) \leq \epsilon$. □

Similar to Algorithm 2, we can develop an algorithm to compute δ-proportional distance. We write:

$$dPD^{\delta}(\rho,\sigma) = \max_{\substack{0<M\leq\mathbb{I}\\ \mathrm{Tr}(M\rho)\geq\delta}} \ln\frac{\mathrm{Tr}(M\rho)-\delta}{\mathrm{Tr}(M\sigma)}$$

Then it is obvious that

$$PD^{\delta}(\rho,\sigma) = \max(dPD^{\delta}(\rho,\sigma), dPD^{\delta}(\sigma,\rho)).$$

Algorithm 3 can efficiently compute $dPD^{\delta}(\rho,\sigma)$ (and thus $PD^{\delta}(\rho,\sigma)$) with an error smaller than θ.

Algorithm 3: Computing dPD^{δ}

Function $DPDD$ $(\rho,\sigma,\delta,\theta)$
// ρ and σ are two density operators in Hilbert space $\mathcal{H}$ and θ is the desired accuracy.
Let n be the dimension of the Hilbert space $\mathcal{H}$.
Set real number $max = 0, \lambda = \infty$.
while $|max-\lambda| > \theta$ **do**
  $\lambda = max$;
  $max = \max_{S\subseteq[n]} \frac{(\sum_{i\in S}\rho_{ii})-\delta}{\sum_{i\in S}\sigma_{ii}}$;
  $\eta = \rho - max*\sigma$;
  Diagonalize η: $\eta = PDP^{\dagger}$;
  // P is unitary and D is diagonal matrix.
  $\rho = P^{\dagger}\rho P$;
  $\sigma = P^{\dagger}\sigma P$;
end
Return $\ln max$;

The following proposition is required to estimate the accuracy of our sampling strategy.

Proposition 5. *Suppose that ρ, σ, $\rho+\varphi$, $\rho+\theta$ and $\sigma+\varphi$ are all density operators, and parameter $\delta\geq 0$. Then:*

1) $PD^{\delta}(\rho,\rho+p\varphi) \leq PD^{\delta}(\rho,\rho+\varphi)$ *for* $p\in[0,1]$*;*
2) *For $p,q\in[0,1]$ and $p+q\leq 1$, we have:*

$$\begin{aligned} &PD^{\delta}(\rho,\rho+p\varphi+q\theta) \\ &\qquad \leq \max\{PD^{\delta}(\rho,\rho+\varphi), PD^{\delta}(\rho,\rho+\theta)\}; \end{aligned}$$

3) *For $p\in[0,1]$, we have:*

$$\begin{aligned} &PD^{\delta}(p\rho+(1-p)\sigma, p\rho+(1-p)\sigma+\varphi) \\ &\qquad \leq \max\{PD^{\delta}(\rho,\rho+\varphi), PD^{\delta}(\sigma,\sigma+\varphi)\}; \end{aligned}$$

4) *If* $PD^{\delta}(\rho,\rho+\theta)=\epsilon$, *then* $PD^{\delta'}(\sigma,\sigma+\varphi)\leq\epsilon$, *where:*

$$\delta'=\delta+(e^{\epsilon}-1)\tau(\rho,\sigma)+e^{\epsilon}\tau(\theta,\varphi).$$

In particular, all of the above inequalities hold for proportional distance PD.

Now we are ready to figure out the accuracy of sampling inputs. Let Δ denote the set of the differences of two adjacent inputs; that is,

$$\Delta=\{\rho-\sigma|\rho,\sigma\in\mathcal{P}\text{ with }\tau(\rho,\sigma)\leq d\},$$

where $\mathcal{P}$ is the set of all possible inputs, and d is the constant fixed in Definition 2. Then a quantum algorithm $\mathcal{M}$ is (ε,δ)-differentially private if and only if:

$$\max_{\substack{\rho\in\mathcal{P},\eta\in\Delta;\\ \rho+\eta\in P}}\{PD^{\delta}(\mathcal{M}(\rho),\mathcal{M}(\rho+\eta))\}\leq\varepsilon.$$

We define the following indexes of the sample sets:

$$\begin{aligned}
d_{\mathcal{P}}&=\max_{\substack{\rho_1\notin\mathrm{Conv}(\mathcal{P}'),\rho_1\in\mathcal{P};\\ \rho_2\in\mathrm{Conv}(\mathcal{P}')}}\tau(\rho_1,\rho_2);\\
d_{\Delta}&=\max_{\substack{\eta_1\notin\mathrm{Conv}(\Delta'),\eta_1\in\Delta;\\ \eta_2\in\mathrm{Conv}(\Delta')}}\tau(\eta_1,\eta_2);\\
p_{\mathcal{P}}&=\Pr[\text{input }\rho\in\mathrm{Conv}(\mathcal{P}')];\\
p_{\Delta}&=\Pr[\text{difference }\eta\in\mathrm{Conv}(\Delta')]
\end{aligned}\tag{9}$$

where $\mathrm{Conv}()$ denotes the convex hull, and $\mathcal{P}',\Delta'$ are the sample sets of $\mathcal{P},\Delta$, respectively. In order to understand why we use the convex sets in equation (9), we may regard a density operator as a point in the Bloch sphere. Note that in above properties only the convex hulls of the sample sets appears, and this implies we can just choose those points near the boundary of $\mathcal{P}$ (resp. Δ) to form $\mathcal{P}'$ (resp. Δ'). Now, with a pre-selected parameter δ_m, we calculate

$$\varepsilon_m=\max_{\rho\in\mathcal{P}',\eta\in\Delta'}\{PD^{\delta_m}(\mathcal{M}(\rho),\mathcal{M}(\rho+\eta))\}.$$

Then we conclude:

1) with at least probability $(1-p_{\mathcal{P}}-p_{\Delta})$ (the probability space is over the choice of databases), the algorithm $\mathcal{M}$ is (ε_m,δ_m)-differentially private;
2) the algorithm $\mathcal{M}$ is $(\varepsilon_m,\delta'_m)$-differentially private where: $\delta'_m=\delta_m+(e^{\varepsilon_m}-1)d_{\mathcal{P}}+e^{\varepsilon_m}d_{\Delta}$.

VI. Technical Proofs

The proofs of the theorems in Section III are very technical and involved. So, for readability we presented these theorems there without proofs. In this section, we complete the picture by providing their proofs.

A. Proof of Theorem 1

We write $\rho'=\mathcal{E}(\rho)$, $\sigma'=\mathcal{E}(\sigma)$, $\rho''=\mathcal{E}_{\mathrm{GAD}}(\rho')$ and $\sigma''=\mathcal{E}_{\mathrm{GAD}}(\sigma')$. By Definition 2, it suffices to show that $PD(\rho'',\sigma'')\leq\varepsilon$ whenever $\tau(\rho,\sigma)\leq d$. Since ρ' and σ' are density operators, we can assume that

$$\rho'=\begin{bmatrix}a & b\\ b^* & c\end{bmatrix},\quad \delta'=\rho'-\sigma'=\begin{bmatrix}\delta_a & \delta_b\\ \delta_b^* & -\delta_a\end{bmatrix}$$

where $c=1-a$. It follows from Theorem 9.2 in [18] that

$$\tau(\mathcal{E}(\rho),\mathcal{E}(\sigma))\leq\tau(\rho,\sigma)=d.$$

So, we have:

$$\begin{aligned}\tau(\rho',\sigma')&=\frac{1}{2}\mathrm{Tr}|\rho'-\sigma'|\\ &=\frac{1}{2}\mathrm{Tr}\,|\delta'|=\sqrt{\delta_a^2+\delta_b^2}\leq d\end{aligned}$$

After the generalized amplitude damping channel, we have:

$$\begin{aligned}\rho''&=\mathcal{E}_{\mathrm{GAD}}(\rho')=\begin{bmatrix}a+\frac{1}{2}(c-a)\gamma & b\sqrt{1-\gamma}\\ b^*\sqrt{1-\gamma} & c+\frac{1}{2}(a-c)\gamma\end{bmatrix}\\ \delta''&=\rho''-\sigma''\\ &=\mathcal{E}_{\mathrm{GAD}}(\delta')=\begin{bmatrix}\delta_a(1-\gamma) & \delta_b\sqrt{1-\gamma}\\ \delta_b^*\sqrt{1-\gamma} & -\delta_a(1-\gamma)\end{bmatrix}\end{aligned}$$

Now we compute the proportional difference $PD(\rho'',\sigma'')$. According to Proposition 3 we only need to consider all the projectors of the form $P=|\varphi\rangle\langle\varphi|$ where $|\varphi\rangle=(x,y)^{\top}$ is a pure state. A routine calculation yields:

$$\begin{aligned}p(\rho)=\mathrm{Tr}[\rho''P]=ax^2+cy^2+\frac{1}{2}(c-a)\gamma(x^2-y^2)+\\ (bx^*y+b^*xy^*)\sqrt{1-\gamma}\end{aligned}$$

and $p(\sigma)=\mathrm{Tr}[\sigma''P]=p(\rho)+\Delta$ where

$$\Delta=\delta_a(1-\gamma)(x^2-y^2)+(\delta_b x^*y+\delta_b^*xy^*)\sqrt{1-\gamma}.$$

We further calculate the supremum of $PD(\rho'',\sigma'')$ over all possible ρ'' and σ''. By symmetry we only need to compute the maximum of

$$p(\sigma)/p(\rho)=1+\Delta/p(\rho).$$

In order to find the maximum of $\Delta/p(\rho)$, we minimize the $(bx^*y+b^*xy^*)\sqrt{1-\gamma}$ and maximize $(\delta_b x^*y+\delta_b^*xy^*)\sqrt{1-\gamma}$. To this end, we choose:

$$\begin{aligned}(bx^*y+b^*xy^*)&=-\sqrt{ac}|xy|,\\ (\delta_b x^*y+\delta_b^*xy^*)&=2|\delta_b||xy|.\end{aligned}$$

Then, we only need to consider real number x, and

$$\begin{aligned}p(\rho)&=ax^2+(1-a)(1-x^2)+\frac{1}{2}(1-2a)(2x^2-1)\gamma\\ &\quad-2\sqrt{a(1-a)}x\sqrt{1-x^2}\sqrt{1-\gamma}\\ \Delta&=\delta_a(1-\gamma)(2x^2-1)+2|\delta_b|x\sqrt{1-x^2}\sqrt{1-\gamma}\end{aligned}$$

Let us first compute the maximum of Δ. We write

$$A=(1-\gamma)(2x^2-1),\quad B=2x\sqrt{1-x^2}\sqrt{1-\gamma}.$$

Then $\Delta=A\delta_a+B\delta_b$ with constrain $\delta_a^2+\delta_b^2\leq d^2$. It is clear that

$$\Delta\leq\sqrt{A^2+B^2}d$$

with the equality when

$$\delta_a=\frac{|A|}{\sqrt{A^2+B^2}}d,\qquad \delta_b=\frac{|B|}{\sqrt{A^2+B^2}}d.$$

On the other hand, it is not difficult to see that the maximum $1-\gamma$ of A^2+B^2 is attained when $x=1/\sqrt{2}$. To compute the minimum of $p(\rho)$, we put $\lambda_a=\frac{1}{2}-a$ and $\lambda_b=x^2-\frac{1}{2}$. Then $-\frac{1}{2}\leq\lambda_a,\lambda_b\leq\frac{1}{2}$ and

$$\begin{aligned}
p(\rho)&=\left(\frac{1}{2}-\lambda_a\right)\left(\frac{1}{2}+\lambda_b\right)+\left(\frac{1}{2}+\lambda_a\right)\left(\frac{1}{2}-\lambda_b\right)\\
&\quad+2\lambda_a\lambda_b\gamma-2\sqrt{\left(\frac{1}{4}-\lambda_a^2\right)\left(\frac{1}{4}-\lambda_b^2\right)}\sqrt{1-\gamma}\\
&=\frac{1}{2}-2\lambda_a\lambda_b(1-\gamma)-2\sqrt{\left(\frac{1}{4}-\lambda_a^2\right)\left(\frac{1}{4}-\lambda_b^2\right)}\sqrt{1-\gamma}\\
&\geq\frac{1}{2}-2\lambda_a\lambda_b(1-\gamma)-2\frac{1}{2}\left(\frac{1}{4}-\lambda_a^2+\frac{1}{4}-\lambda_b^2\right)\sqrt{1-\gamma}\\
&=\frac{1}{2}(1-\sqrt{1-\gamma})+(\lambda_a^2+\lambda_b^2)\sqrt{1-\gamma}-2\lambda_a\lambda_b(1-\gamma)\\
&\geq\frac{1}{2}(1-\sqrt{1-\gamma})+2|\lambda_a\lambda_b|\sqrt{1-\gamma}-2\lambda_a\lambda_b(1-\gamma)\\
&\geq\frac{1}{2}(1-\sqrt{1-\gamma})\qquad\left(\because\sqrt{1-\gamma}\geq(1-\gamma)\geq 0\right)
\end{aligned}$$

All of the equalities hold when $\lambda_a=\lambda_b=0$, or $x=1/\sqrt{2}$ and $a=1/2$. Combining with the condition for Δ, we see that when $x=1/\sqrt{2}$, $a=1/2$, $\delta_a=0$ and $\delta_b=d$, Δ has the maximum $\sqrt{1-\gamma}d$ and $p(\rho)$ has the minimum $\frac{1}{2}(1-\sqrt{1-\gamma})$. So, it holds that

$$\frac{p(\sigma)}{p(\rho)}=1+\frac{\Delta}{p(\rho)}\leq 1+\frac{2d\sqrt{1-\gamma}}{1-\sqrt{1-\gamma}} \tag{10}$$

and $PD(\rho'',\sigma'')\leq\varepsilon$.

Finally, we prove the optimality of ϵ. Since $d\leq 1$, we can choose $b=-1/2$ and thus $\det(\sigma')\geq 0$. Then the upper bound in equation (10) can be achieved.

B. Proof of Theorem 3

We write $\rho'=\mathcal{E}(\rho)$ and $\sigma'=\mathcal{E}(\sigma)$. Then it follows from Theorem 9.2 in [18] that $\tau(\rho',\sigma')\leq d$. For an arbitrary POVM $M=\{M_m\}$, we first prove that

$$\mathrm{Tr}[(\rho'-\sigma')M_m]\leq d\mathrm{Tr}(M_m) \tag{11}$$

for all outcomes m. In fact, if it is not true for some outcome n, then we diagonalize: $M_n=U\Lambda U^\dagger$, and define λ to be the maximal element of Λ. We further define the positive operator $Q_n=\frac{1}{\lambda}M_n$. Obviously, $\mathrm{Tr}(Q_n)\geq 1$ and $Q_n\leq I$. The second inequality implies that we can extend Q_n to a new POVM $Q=\{Q_k\}$ such that Q_n is one of its elements. Let

$$q_k(\rho')\equiv\mathrm{Tr}[\rho' Q_k],\qquad q_k(\sigma')\equiv\mathrm{Tr}[\sigma' Q_k]$$

be the probabilities of obtaining the measurement outcome labeled by k. Then we have:

$$\begin{aligned}
q_n(\rho')-q_n(\sigma')&=\mathrm{Tr}[(\rho'-\sigma')Q_n]\\
&=\frac{1}{\lambda}\mathrm{Tr}[(\rho'-\sigma')M_n]\\
&>\frac{1}{\lambda}d\mathrm{Tr}(M_n)=d\mathrm{Tr}(Q_n)\geq d.
\end{aligned}$$

Using Theorem 9.1 in [18], we see that the above inequality contradicts to $\tau(\rho',\sigma')\leq d$, and thus (11) is proved.

Now we put

$$p_m(\rho)\equiv\mathrm{Tr}[\rho'' M_m],\qquad p_m(\sigma)\equiv\mathrm{Tr}[\sigma'' M_m],$$

where $\rho''=\mathcal{E}_{\mathrm{Dep}}(\rho')$ and $\sigma''=\mathcal{E}_{\mathrm{Dep}}(\sigma')$. Then it holds that

$$\begin{aligned}
\frac{p_m(\rho)}{p_m(\sigma)}-1&=\frac{\frac{p}{D}\mathrm{Tr}[M_m]+(1-p)\mathrm{Tr}[\rho' M_m]}{\frac{p}{D}\mathrm{Tr}[M_m]+(1-p)\mathrm{Tr}[\sigma' M_m]}-1\\
&=\frac{(1-p)\mathrm{Tr}[(\rho'-\sigma')M_m]}{\frac{p}{D}\mathrm{Tr}[M_m]+(1-p)\mathrm{Tr}[\sigma' M_m]}\\
&\leq\frac{(1-p)d\mathrm{Tr}[M_m]}{\frac{p}{D}\mathrm{Tr}[M_m]}=\frac{1-p}{p}dD.
\end{aligned}$$

So we conclude:

$$e^{-\varepsilon}\leq\frac{p_m(\rho)}{p_m(\sigma)}\leq e^{\varepsilon}$$

and

$$\varepsilon=\ln\left[1+\frac{1-p}{p}dD\right].$$

The same holds if we exchange ρ and σ. For the case that dD/p is close to zero, using the Maclaurin series we can show that

$$\varepsilon\approx\frac{1-p}{p}dD.$$

VII. Discussions and Conclusion

In this paper, we generalise the notion of differential privacy for classical computation defined in [9] to quantum computation. Three simple privacy mechanisms using quantum noises are proposed, and their abilities for protecting privacy are examined by estimating their differential privacy parameters. Furthermore, we establish several composition theorems that allow us to figure out the differential privacy parameters of complex and sophisticated mechanisms where several quantum algorithms are combined. With the discovery of more and more quantum algorithms for data mining and big data analytics, we believe that the framework of quantum differential privacy developed in this paper will be very useful for privacy protection in quantum computing, e.g. against the so-called joint-measurement attack, which may cause a noticeable disclosure of data privacy.

A. Physical Implementation of Quantum Privacy Mechanism

As pointed out in Section III, there are certain physical technologies that can realise the three types of quantum noises considered in this paper. However, implementing them in practical applications is very hard. In the current stage, controlling quantum noises is one of the greatest difficulties in building a quantum computer. The decoherence time and/or relaxation time of the materials used to implement quantum computation (e.g., [25]) is crucial. To ensure that the quantum states during the calculation are reliable, the evolution time of the whole computation should not be longer than the decoherence time and/or relaxation time. So, a crude way to implement the quantum noises needed for our quantum privacy mechanism is

to control the evolution time of the calculation; for example, if we leave the used quantum states in the material for some extra time, then the decoherence and relaxation is naturally happened. But such a method can only introduce the typical noise of the material, which may not be the type of the noises that we want. The theory of decoherence-free subspace, noiseless subsystems, and dynamical decoupling provides us with some further techniques to overcome decoherence and to introduce the required noise (see [26] for a review). Assume that the interaction of the system and the environment (bath) is described by a Hamiltonian H_{SB}. Then we can find a noiseless subsystem such that under the evolution of H_{SB}, the state in this subsystem is preserved up to a global phase. If the information is encoded into this subsystem, then the state undergoes the calculation is not disturbed by the environment. This gives rise to a way to implement noises: H_{SB} can be decomposed into two part $H_{SB} = H'_{SB} + H_N$, where H_N is the system where the desired noise arises. Now, if we encode the information into the noiseless subsystem H'_{SB}, then the state only suffers the noise that we desire while other types of noises do not work.

B. Related Work

A very interesting related work is the delegated quantum computation [27], [28], [29]. The basic idea of delegated quantum computation is as follows: similar to today's cloud computing, a large quantum computer plays a central role. A client which is not capable of a full-blown quantum computation, sends the input which may be either quantum or classical data as well as the program (the descriptions of the computation she/he wants to perform), to the server who is able to perform universal quantum computation, and then the server executes the desired computation and sends the output back to the client. Two important features of delegated quantum computation are: (1) Blindness: the computation remains hidden from the server, and (2) Verifiability: the client is able to confirm that the final output of the computation is correct. Furthermore, the composable security of delegated quantum computation was carefully studied by Dunjko et al. [30]. It worth noting that the client's security in delegated quantum computation is well-addressed, but no enough attentions has been paid to the server's security, i.e. the security of the private resources held by the server; for example, when the computation of the server is based on some private (either classical or quantum) databases, the output sent to the client should not leak private information of the databases. It seems that quantum differential privacy mechanism studied in this paper can be introduced into the framework of delegated quantum computation in order to protect the server's privacy.

C. Topics for Future Research

Only the measure-each-step scenario is dealt with in Theorem 6. As a topic for future research, we can consider an extension of Theorem 6 for a more "quantum" scenario where at each round only a part of the quantum state is measured, and the other part of the state is used as the input of next query. It seems that the proof techniques employed in the case of classical computing presented in [21] do not work for this case, and some radically new ideas are required to established such a generalised version of advanced composition theorem for quantum computing.

A probabilistic relational Hoare logic and a machine-checked framework for reasoning about differential privacy for (classical) programs were established by Barthe, Köpf, Olmedo et al. [31], [32]. On the other hand, a Hoare logic for proving correctness of quantum programs was proposed in [33] (see also [34], Chapter 4). Then another topic for future research is to build a logic and tools for derivation of differential privacy guarantees in quantum computation by combining the techniques developed in [31], [32], [33].

ACKNOWLEDGMENT

We would like to thank Professor Zhengfeng Ji for valuable discussions.

REFERENCES

[1] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, ser. STOC'96. New York: ACM, 1996, pp. 212–219. [Online]. Available: http://doi.acm.org/10.1145/237814.237866

[2] A. Ambainis, "Quantum walk algorithm for element distinctness," *SIAM J. Comput.*, vol. 37, no. 1, pp. 210–239, 2007. [Online]. Available: http://dx.doi.org/10.1137/S0097539705447311

[3] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nat. Phys.*, vol. 10, no. 9, pp. 631–633, 2014. [Online]. Available: http://dx.doi.org/10.1038/nphys3029

[4] A. Daskin, "Obtaining a linear combination of the principal components of a matrix on quantum computers," *arXiv:1512.02109*, 2015. [Online]. Available: https://arxiv.org/abs/1512.02109

[5] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Phys. Rev. Lett.*, vol. 113, p. 130503, Sep 2014. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.113.130503

[6] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," *arXiv:1307.0411*, 2013. [Online]. Available: http://arxiv.org/abs/1307.0411

[7] X.-D. Cai, D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, "Entanglement-based machine learning on a quantum computer," *Phys. Rev. Lett.*, vol. 114, p. 110504, Mar 2015. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.114.110504

[8] M. Ying, Y. Feng, and N. Yu, "Quantum information-flow security: Noninterference and access control," in *Proceedings of the 2013 IEEE 26th Computer Security Foundations Symposium*, ser. CSF'13. Washington: IEEE Computer Society, 2013, pp. 130–144. [Online]. Available: http://dx.doi.org/10.1109/CSF.2013.16

[9] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proceedings of the Third Conference on Theory of Cryptography*, ser. TCC'06. Berlin, Heidelberg: Springer, 2006, pp. 265–284.

[10] C. Dwork, "Differential privacy," in *Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II*, ser. ICALP'06. Berlin, Heidelberg: Springer, 2006, pp. 1–12.

[11] T. Dalenius, "Towards a methodology for statistical disclosure control," *statistik Tidskrift*, vol. 15, pp. 429–444, 1977.

[12] F. McSherry and K. Talwar, "Mechanism design via differential privacy," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, ser. FOCS'07. Washington: IEEE Computer Society, 2007, pp. 94–103. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2007.66

[13] K. Nissim, R. Smorodinsky, and M. Tennenholtz, "Approximately optimal mechanism design via differential privacy," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ser. ITCS'12. New York: ACM, 2012, pp. 203–213. [Online]. Available: http://doi.acm.org/10.1145/2090236.2090254

[14] X. Xiao, G. Wang, and J. Gehrke, "Differential privacy via wavelet transforms," *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 8, pp. 1200–1214, Aug 2011.

[15] S. P. Kasiviswanathan, K. Nissim, S. Raskhodnikova, and A. Smith, "Analyzing graphs with node differential privacy," in *Proceedings of the 10th Theory of Cryptography Conference on Theory of Cryptography*, ser. TCC'13. Berlin, Heidelberg: Springer, 2013, pp. 457–476.

[16] A. Friedman and A. Schuster, "Data mining with differential privacy," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD'10. New York: ACM, 2010, pp. 493–502. [Online]. Available: http://doi.acm.org/10.1145/1835804.1835868

[17] C. Dwork, G. N. Rothblum, and S. Vadhan, "Boosting and differential privacy," in *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, ser. FOCS'10. Washington: IEEE Computer Society, 2010, pp. 51–60. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2010.12

[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[19] V. Giovannetti, S. Lloyd, and L. Maccone, "Quantum random access memory," *Phys. Rev. Lett.*, vol. 100, p. 160501, Apr 2008. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.100.160501

[20] S. Arunachalam, V. Gheorghiu, T. Jochym-O'Connor, M. Mosca, and P. V. Srinivasan, "On the robustness of bucket brigade quantum ram," *New J. Phys.*, vol. 17, no. 12, p. 123010, 2015. [Online]. Available: http://stacks.iop.org/1367-2630/17/i=12/a=123010

[21] C. Dwork and A. Roth, "The algorithmic foundations of differential privacy," *Found. Trends Theor. Comput. Sci.*, vol. 9, no. 3-4, pp. 211–407, 2014. [Online]. Available: http://dx.doi.org/10.1561/0400000042

[22] R. Srikanth and S. Banerjee, "Squeezed generalized amplitude damping channel," *Phys. Rev. A*, vol. 77, p. 012318, Jan 2008. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevA.77.012318

[23] D. Leung, L. Vandersypen, X. Zhou, M. Sherwood, C. Yannoni, M. Kubinec, and I. Chuang, "Experimental realization of a two-bit phase damping quantum code," *Phys. Rev. A*, vol. 60, pp. 1924–1943, Sep 1999. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevA.60.1924

[24] M. Grant and S. Boyd, "Cvx: Matlab software for disciplined convex programming, version 1.21 (2011)," 2010. [Online]. Available: http://cvxr.com/cvx

[25] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, "Silicon quantum electronics," *Rev. Mod. Phys.*, vol. 85, no. 3, p. 961, 2013. [Online]. Available: https://doi.org/10.1103/RevModPhys.85.961

[26] D. A. Lidar, "Review of decoherence free subspaces, noiseless subsystems, and dynamical decoupling," *Adv. Chem. Phys.*, vol. 154, pp. 295–354, 2014. [Online]. Available: https://doi.org/10.1002/9781118742631.ch11

[27] A. M. Childs, "Secure assisted quantum computation," *Quantum Info. Comput.*, vol. 5, no. 6, pp. 456–466, 2005. [Online]. Available: http://dl.acm.org/citation.cfm?id=2011670.2011674

[28] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," in *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, ser. FOCS'09. Washington: IEEE Computer Society, 2009, pp. 517–526. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2009.36

[29] T. Morimae, "Continuous-bariable blind quantum computation," *Phys. Rev. Lett.*, vol. 109, no. 23, p. 230502, 2012.

[30] V. Dunjko, J. F. Fitzsimons, C. Portmann, and R. Renner, "Composable security of delegated quantum computation," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2014, pp. 406–425.

[31] G. Barthe, B. Köpf, F. Olmedo, and Z. S. Béguelin, "Probabilistic relational reasoning for differential privacy," in *Proceedings of the 39th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, ser. POPL'12, 2012, pp. 97–110. [Online]. Available: https://doi.org/10.1145/2103656.2103670

[32] ——, "Probabilistic relational reasoning for differential privacy," *ACM Transactions on Programming Languages and Systems*, vol. 35, pp. 9:1–9:49, 2013. [Online]. Available: https://doi.org/10.1145/2492061

[33] M. Ying, "Floyd-hoare logic for quantum programs," *ACM Transactions on Programming Languages and Systems*, vol. 33, pp. 19:1–19:49, 2011. [Online]. Available: https://doi.org/10.1145/2049706.2049708

[34] ——, *Foundations of Quantum Programming*. Morgan Kaufmann, 2016.