Information-Disturbance theorem and Uncertainty Relation

Takayuki Miyadera * and Hideki Imai *,†
* Research Center for Information Security (RCIS), National Institute of Advanced Industrial Science and Technology (AIST).
Daibiru building 1102, Sotokanda, Chiyoda-ku, Tokyo, 101-0021, Japan.
(e-mail: miyadera-takayuki@aist.go.jp)
† Graduate School of Science and Engineering, Chuo University.
1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan.

February 1, 2008

Abstract

It has been shown that Information-Disturbance theorem can play an important role in security proof of quantum cryptography. The theorem is by itself interesting since it can be regarded as an information theoretic version of uncertainty principle. It, however, has been able to treat restricted situations. In this paper, the restriction on the source is abandoned, and a general information-disturbance theorem is obtained. The theorem relates information gain by Eve with information gain by Bob.

1 Introduction

In 1984, Bennett and Brassard[1] proposed a protocol to realize key distribution that uses quantum theory in its essential part. In spite of simplicity of the protocol, its unconditional security proof[2, 3, 4, 5] appeared more than a decade later after its proposal. Among the various existing proofs, a proof by Biham et al.[5] employs a so-called information-disturbance theorem[6, 7, 8, 9] that can be regarded as an information theoretical version of the uncertainty relation. We, in [10], succeeded in deriving an improved variation of the theorem. Our theorem expressed a relation between information gain by Eve and randomness of error contained in Bob’s data. Although it has a natural form, its applicability is still restricted. In fact the state prepared by Alice has to be ensembles consisting of pure states with even probability. In this paper, we get
rid of this strong condition and show fairly generalized form of the information-disturbance theorem. Alice prepares an arbitrary state by one of two different ensembles. That is, Alice chooses one of two random variables to be encoded. Each ensemble does not need to consist of distinguishable states. Our new information-disturbance theorem represents a relation between Eve’s information gain and Bob’s information gain. According to the theorem, if Eve employs an attack that gives her large information on an encoded random variable, Bob could obtain small information on another random variable. This trade-off is determined by noncommutativity between the ensembles. The theorem is derived by using remote ensemble preparation technique and entropic uncertainty relation. These technique also allows us to obtain a simple derivation of the result in [10]. In section 2 we give a brief review on positive operator valued measure and entropic uncertainty relation that play central roles in our proof. In section 3 we introduce a method to prepare remotely an ensemble of quantum states by making a proper measurement on predistributed quantum state. In section 4 our main theorems are presented.

2 Preliminaries

We begin with a brief introduction of relevant notions in quantum theory: positive operator valued measure and entropic uncertainty relation.

2.1 Positive Operator Valued Measure (POVM)

A quantum system is described by a Hilbert space and operators acting on it. The most general observable is represented by a positive operator valued measure (POVM) (see, e.g. [11]). A positive operator valued measure $A(\cdot)$ is a map from measurable space $(\Omega, F)$ to a set of positive operators satisfying:

(i) For all $S_1, S_2 \in F$ satisfying $S_1 \cap S_2 = \phi$, $A(S_1 \cup S_2) = A(S_1) + A(S_2)$ holds.

(ii) $A(\Omega) = 1$ holds.

Hereafter we treat only the case that the measurable set is a finite set. Therefore the conditions above can be rephrased as follows. A POVM is a family of positive operators $\{A_a\}_{a \in \Omega}$ satisfying $\sum_{a \in \Omega} A_a = 1$. Each $a \in \Omega$ corresponds to a measurement outcome. A POVM is called as a projection valued measure (PVM) if $A_a$ is a projection operator for all $a \in \Omega$. A state is described by a so-called density operator. A density operator $\rho$ is defined by an operator satisfying $\rho \geq 0$ and $\text{tr} \rho = 1$. If one measures an observable $A = \{A_a\}$ in a state $\rho$, one obtains an outcome $a$ with probability $\text{tr}(\rho A_a)$. From a POVM $A = \{A_a\}$ one can construct a self adjoint operator $\hat{A} := \sum_{a \in \Omega} a A_a$. This operator is useful since it gives the expectation value for the measurements so that $\langle A \rangle_\rho = \text{tr}(\rho \hat{A})$. For PVM, the standard deviation can be calculated as $\Delta A_\rho = (\langle \hat{A}^2 \rangle_\rho - \langle \hat{A} \rangle_\rho^2)^{1/2}$. 

2
2.2 Entropic Uncertainty Relation

As is widely known, the uncertainty relation is the most fundamental result of quantum theory. It, in general, is expressed by an inequality. The uncertainty relation treats two (or more) observables. Incompatibility of probability distributions of their measurement outcomes is bounded by noncommutativity between them. The most famous one is the Robertson-type uncertainty relation for PVMs:

$$\Delta A_\rho \Delta B_\rho \geq \frac{1}{2} |\text{tr}(\rho [\hat{A}, \hat{B}])|$$

where $\Delta A_\rho$ ($\Delta B_\rho$) represents standard deviation of the outcome of the corresponding observables. However, to characterize randomness of measurement outcomes, the standard deviation is often insufficient. The standard deviation depends on how to assign a value of measurement outcome to each event. For instance, let us imagine an observable which takes 0, 1 and 2 as its value of measurement outcome. If a state gives an outcome 0 or 1 with probability $\frac{1}{2}$, its standard deviation is $\frac{1}{2}$. On the other hand, if we shuffle the values of outcome so that the new observable takes an outcome 0 or 2 with probability $\frac{1}{2}$, its standard deviation becomes 1. In addition, the above Robertson-type formulation cannot deal with the most general type of measurement, positive operator value measure (POVM) measurement. The entropic uncertainty relation can cover this type of measurement and is of advantage to its application. It has the following form:

$$H(A|\rho) + H(B|\rho) \geq -2 \log \max_{a,b} \| A_a^{1/2} B_b^{1/2} \|,$$

where $A := \{A_a\}$ and $B := \{B_b\}$ are POVMs and $H(A|\rho)$ ($H(B|\rho)$) represents Shannon entropy of the probability distribution of the measurement outcome of $A$ ($B$) in a state $\rho$, i.e., $H(A|\rho) = - \sum_{a \in \Omega} \text{tr}(\rho A_a) \log \text{tr}(\rho A_a)$. This type of uncertainty relation was first proposed by Deutsch[12] and was improved by Maassen and Uffink[13]. The above general form for POVMs was obtained by Krishna and Parthasarathy[14].

3 Remote Ensemble Preparation

In this section we explain a way to prepare an ensemble of quantum states on a remotely located quantum system by using predistributed entangled state. It plays an essential role to prove impossibility of the bit commitment. It has been used to translate the BB84 quantum key distribution into E91 quantum key distribution. The theorem was first proved by Hughston, Jozsa and Wootters[15], and generalized by Halvorson[16] for the most general quantum system including infinite systems. We, in this paper, treat only finite quantum systems that are described by finite dimensional Hilbert spaces. Suppose there exist two characters: Alice and Bob. Each of them has a quantum system. The system
possessed by Alice (Bob) is called as system $A$ (system $B$). Alice (Bob) can manipulate only the system $A$ (system $B$). The system $A$ (system $B$) is described by a Hilbert space $\mathcal{H}_A$ ($\mathcal{H}_B$). We assume that they have an identical finite dimension, $\mathcal{H}_A \simeq \mathcal{H}_B \simeq \mathbb{C}^N$. We consider a method to prepare an ensemble of states on the system $B$ by Alice’s operation on a predistributed entangled state $|\Phi\rangle$. A normalized vector of the composite system, $|\Phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, can be written as, thanks to Schmidt decomposition theorem,

$$|\Phi\rangle = \sum_k \sqrt{\lambda_k} |e^A_k\rangle \otimes |e^B_k\rangle,$$

where $\{|e^A_k\rangle\}$ ($\{|e^B_k\rangle\}$) is an orthonormal basis of $\mathcal{H}_A$ ($\mathcal{H}_B$). We hereafter fix a normalized vector $|\Phi\rangle$ and its corresponding basis. $\mathcal{H}_A$ and $\mathcal{H}_B$ are identified with respect to these basis. We write its reduced state on each system as,

$$\rho^A = \sum_k \lambda_k |e^A_k\rangle \langle e^A_k|,$$

$$\rho^B = \sum_k \lambda_k |e^B_k\rangle \langle e^B_k|.$$

When we identify these two Hilbert space, we simply write them as $\rho(\equiv \rho^A \equiv \rho^B)$. Suppose that the state $\rho^A$ can be decomposed into a mixture of the states as $\rho^A \equiv \sum_i p_i \rho^A_i$, where $\rho^A_i$ is a state of the system $A$ for each $i$ and $\{p_i\}$ satisfies $\sum_i p_i = 1$ and $p_i \geq 0$. Hereafter, for simplicity, we assume rank $\rho = N$. In the following, we consider a measurement by Alice that prepares the state $\rho_i$ with the probability $p_i$ on the system $B$ attached to Bob. We define transpose operation with respect to the basis $\{|e^A_k\rangle\}$. Since the transpose operation $A \mapsto A^\dagger$ preserves the positivity of the operator, a family of operators,

$$F[\{p_j, \rho_j\}] := \{F[\{p_j, \rho_j\}]_i : \{p_i \rho^{-1/2} A^\dagger \rho^{-1/2}\}\}$$

(1)

forms a POVM. Let us take the state $|\Phi\rangle$ and consider an a-posteriori state with respect to the POVM $F[\{p_j, \rho_j\}]$. A probability to obtain an outcome $i$ is calculated as

$$\langle \Phi | F[\{p_j, \rho_j\}] | \Phi \rangle = \sum_k \lambda_k \langle e^A_k | F[\{p_j, \rho_j\}]_i | e^A_k \rangle = \sum_k \lambda_k p_i \langle e^A_k | \rho^{-1/2} A^\dagger \rho^{-1/2} | e^A_k \rangle = p_i \text{tr}(\rho^{-1/2} A^\dagger \rho^{-1/2}) = p_i.$$

Since Alice does not make any operation on the system $B$, the a-posteriori state of the system $B$ for the outcome $i$ is calculated as $\langle \Phi | F[\{p_j, \rho_j\}]_i | A | \Phi \rangle / p_i$. Since for each operator $A$ on $\mathcal{H}_B$,

$$\langle \Phi | F[\{p_j, \rho_j\}]_i | A | \Phi \rangle = \sum_{k,l} \sqrt{\lambda_k \lambda_l} \langle e^A_k | \rho^{-1/2} A^\dagger \rho^{-1/2} | e^A_l \rangle \langle e^A_l | A | e^A_l \rangle$$

4
\[
\sum_{k,l} \sqrt{\lambda_k \lambda_l} \langle e_l | \rho^{-1/2} \rho_i \rho^{-1/2} | e_k \rangle \langle e_k | A | e_l \rangle
= p_i \text{tr}(\rho_i A)
\]

holds, where we used a relation \( ^t \rho = \rho \), the a-posteriori state of the system \( B \) is \( \rho_i \). We thus proved the following theorem.

**Theorem 1** Suppose that there exist Hilbert spaces \( \mathcal{H}_A, \mathcal{H}_B \) and a normalized vector \( |\Phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \). Assume that the reduced density operator \( \rho := \text{tr}_{\mathcal{H}_B}(|\Phi\rangle\langle \Phi|) \) can be decomposed into a mixture of states as, \( \rho = \sum_i p_i \rho_i \). There exists a POVM \( F[\{p_j, \rho_j\}] = \{F[\{p_j, \rho_j\}]\} \) on \( \mathcal{H}_A \) that prepares the ensemble \( \{p_i, \rho_i\} \) on \( \mathcal{H}_B \). That is, the probability to obtain an outcome \( i \) is \( p_i \), and the a-posteriori state of \( \mathcal{H}_B \) then is \( \rho_i \).

### 4 Information-Disturbance theorem

In this section, we derive two types of Information-Disturbance theorem. Both treat a cryptographic setting. The first one relates information gain by Eve with information gain by Bob. The second one relates information gain by Eve with randomness of error contained in Bob’s outcome.

#### 4.1 Information v.s. Information

We deal with a quantum cryptographic setting. It is a simplified version of the BB84 protocol. Three characters: Alice, Bob, and Eve, play their roles. Alice has a quantum system described by an \( N \)-dimensional Hilbert space, \( \mathcal{H}_A \). She prepares a state \( \rho \) of this system in one of the two different methods: (a) she prepares a state \( \rho_i \) with probability \( p_i \) for each \( i \), or (b) she prepares a state \( \sigma_l \) with probability \( q_l \) for each \( l \). To assure that both procedures actually give the state \( \rho \), we impose a condition, \( \rho = \sum_i p_i \rho_i = \sum_l q_l \sigma_l \). We write \( X \) (\( Y \)) a random variable whose value takes \( i \) (\( l \)) with the probability \( p_i \) (\( q_l \)). The preparation can be regarded as encoding \( X \) or \( Y \) to the state \( \rho \). The full protocol runs as follows:

(i) Alice chooses one of the two methods, (a) or (b), to prepare the state \( \rho \).

(ii) Alice prepares the state \( \rho \) according to her choice on the method. That is, Alice encodes \( X \) or \( Y \) to the state.

(iii) Alice sends the system to Bob.

(iv) After confirming that Bob has actually received the system, Alice publishes the method ((a) or (b)) which she employed to prepare the state \( \rho \).

(v) Bob makes a measurement on his received system to extract the encoded information. We write Hilbert space of the received system as \( \mathcal{H}_B \) instead of \( \mathcal{H}_A \) for convenience.
Note that even if there is no eavesdropper between Alice and Bob, Bob does not obtain in general the full information Alice has encoded. That is, the encoding employed by Alice may be ambiguous one. $\rho_i$ and $\rho_j$ for $i \neq j$ may not be distinguishable perfectly. In the following we will see that Eve’s eavesdropping in addition makes Bob’s information gain less. Let us see what Eve can do. Eve who wants to obtain information encoded by Alice can make her own apparatus interact with $\mathcal{H}_A$ when it is sent to Bob. She may keep the apparatus and only after knowing Alice’s announcement, may make a measurement on it to obtain a classical output. Denote $\mathcal{H}_E$ the Hilbert space of Eve’s apparatus. Eve’s operation is described by a unitary operator $U : \mathcal{H}_B \otimes \mathcal{H}_E \to \mathcal{H}_B \otimes \mathcal{H}_E$. A state of the apparatus before the interaction is written as $|\Omega\rangle \in \mathcal{H}_E$. Without loss of generality, we can assume it as vector state. Eve’s attack is determined by the triplet, $(\mathcal{H}_E, |\Omega\rangle, U)$. After Alice’s announcement, Eve tries to make an optimal measurement $Z$, a POVM, on her apparatus to extract the encoded information. What we are interested in is the trade-off between the information gain by Bob and one by Eve. Let us suppose a fixed Eve’s attack $(\mathcal{H}_E, |\Omega\rangle, U)$. We define $I(X : B)$ as optimal information gain by Bob on random variable $X$. That is, if Alice has encoded $X$ to the quantum state and Eve employs an attack $(\mathcal{H}_E, |\Omega\rangle, U)$, Bob’s optimal information gain on $X$ is $I(X : B)$. In the same manner, $I(Y : B)$ is defined as information gain by Bob on random variable $Y$. $I(Y : E)$, on the other hand, is defined as optimal information gain by Eve on $Y$ if Alice has encoded $Y$ to the quantum state $\rho$ and Eve employs the attack $(\mathcal{H}_E, |\Omega\rangle, U)$. $I(X : E)$ is defined as optimal information gain by Eve on $X$.

**Theorem 2** For a fixed Eve’s attack $(\mathcal{H}_E, |\Omega\rangle, U)$, the following inequalities hold:

$$I(X : B) + I(Y : E) \leq H(X) + H(Y) + 2 \log \max_{i,k} \| F[\{p_j, \rho_j\}]_i^{1/2} F[\{q_i, \sigma_i\}]_k^{1/2} \|,$$

$$I(X : E) + I(Y : B) \leq H(X) + H(Y) + 2 \log \max_{i,k} \| F[\{p_j, \rho_j\}]_i^{1/2} F[\{q_i, \sigma_i\}]_k^{1/2} \|,$$

where POVMs, $F[\{p_j, \rho_j\}]$ and $F[\{q_i, \sigma_i\}]$ are defined by (1).

**Proof:**

To calculate Bob’s and Eve’s information gain, we construct an appropriate probability distribution. We apply the remote ensemble preparation technique. Suppose that $\rho$ can be diagonalized as $\rho = \sum_k \lambda_k |e_k^A\rangle \langle e_k^A|$ and thus $\{ |e_k^A\rangle \}$ forms a basis of $\mathcal{H}_A$. We introduce $\{ |e_k^B\rangle \}$, a basis of $\mathcal{H}_B$ and use $\{ |e_k^A\rangle \}$ and $\{ |e_k^B\rangle \}$ to identify both Hilbert spaces. Let us introduce a virtual entangled state on $\mathcal{H}_A \otimes \mathcal{H}_B$. A normalized vector $|\Phi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is defined as,

$$|\Phi\rangle := \sum_k \sqrt{\lambda_k} |e_k^A\rangle \otimes |e_k^B\rangle.$$  

As we have explained, this state can be used for the remote ensemble preparation in case of existence of Eve. In fact, if Alice operates a POVM $F[\{p_i, \rho_i\}]$ ($F[\{q_i, \sigma_i\}]$) on this state, Bob obtains a state $\rho_i$ ($\sigma_i$) with probability $p_i$ ($q_i$).
The effect Eve gives on it can be included by defining a new state,

\[ |\Psi\rangle := (1 \otimes U)|\Phi\rangle \otimes |\Omega\rangle. \]

If Alice applies a POVM \( F([q_1, \sigma_1]) \) on this state, she obtains an outcome \( i \) (\( l \)) with probability \( p_i(q_1) \) and the state of Bob and Eve then is \( U(\rho_2 \otimes |\Omega\rangle\langle\Omega|)U^* \) \( (U(\sigma_l \otimes |\Omega\rangle\langle\Omega|)U^*) \). Let us consider arbitrary POVMs \( \tilde{B} := \{B_b\} \) of Bob’s and \( Z := \{Z_z\} \) of Eve’s. We write the random variable representing the outcome of Bob’s (Eve’s) measurement also as \( \tilde{B} (Z) \). A-posteriori state \( \rho_{B=b, Z=z} \) of \( H_A \) with respect to these POVMs is written as

\[
\text{tr}(\rho_{\tilde{B}=b, Z=z}A) := \frac{\text{tr}(\rho A\tilde{B}_bZ_z)}{\text{tr}(\rho B_bZ_z)}
\]

for an arbitrary operator \( A \) on \( H_A \). We apply the entropic uncertainty relation to this state. The observables to be concerned are POVMs: \( F([p_j, \rho_j]) \) and \( F([q_l, \sigma_l]) \). We obtain,

\[
H(X|\tilde{B} = b, Z = z) + H(Y|\tilde{B} = b, Z = z) \geq -2 \log \max_{i,k} \|F([p_j, \rho_j])_{i|}^{1/2} F([q_l, \sigma_l])_{k|}^{1/2}\|
\]

Subtracting \( H(X) + H(Y) \) from both sides and summing them up with \( \langle \Phi|\tilde{B}_bZ_z|\Phi\rangle \), we obtain,

\[
I(X : \tilde{B}, Z) + I(Y : \tilde{B}, Z) \leq H(X) + H(Y) + 2 \log \max_{i,k} \|F([p_j, \rho_j])_{i|}^{1/2} F([q_l, \sigma_l])_{k|}^{1/2}\|
\]

Using \( I(X : \tilde{B}) \leq I(X : \tilde{B}, Z) \) and \( I(Y : \tilde{B}) \leq I(Y : \tilde{B}, Z) \), or \( I(X : Z) \leq I(X : \tilde{B}, Z) \) and \( I(Y : Z) \leq I(Y : \tilde{B}, Z) \) we obtain,

\[
I(X : \tilde{B}) + I(Y : Z) \leq H(X) + H(Y) + 2 \log \max_{i,k} \|F([p_j, \rho_j])_{i|}^{1/2} F([q_l, \sigma_l])_{k|}^{1/2}\|
\]

\[
I(X : Z) + I(Y : \tilde{B}) \leq H(X) + H(Y) + 2 \log \max_{i,k} \|F([p_j, \rho_j])_{i|}^{1/2} F([q_l, \sigma_l])_{k|}^{1/2}\|
\]

Since the POVMs \( \tilde{B} \) and \( Z \) are arbitrary, we can take the optimal one for both. Q.E.D.

This theorem gives nontrivial bounds if POVMs \( F([p_j, \rho_j]) \) and \( F([q_l, \sigma_l]) \) do not commute with each other. That is, when Eve employs an operation that should yield herself to obtain large information if the encoded random variable was \( Y (X) \), Bob cannot obtain large information on \( X (Y) \) that was actually employed by Alice.

Let us consider the simplest example. The system consists of \( N \)-qubits. \( \rho \) is the maximally mixed state, \( \rho = \frac{1}{2^n} \). Each bit has two natural basis corresponding to the eigenvectors of \( \sigma_z \) and \( \sigma_x \). Let \( b \) be an element of \( \{z, x\}^N \). \( b \) naturally determines a basis of \( N \)-qubit and an observable \( X(b) \) that is diagonalized by this basis. We write \( \tilde{b} \), the conjugate basis of \( b \). It is defined by exchange all \( z \) (\( x \)) to \( x \) (\( z \)). We write its corresponding observable as \( X(\tilde{b}) \). In this situation, we obtain,

\[
I(X(b) : B) + I(X(\tilde{b}) : E) \leq N.
\]
4.2 Information v.s. Randomness of Error

In [10] we derived a theorem [17] that relates information gain by Eve and randomness of error contained in Bob’s data. Its derivation, however, relied upon symmetrization technique and Holevo bound, and was complicated. We here give another simple proof of the theorem by the remote ensemble preparation technique and the entropic uncertainty relation.

Let us first begin with the setting. It is a special case of the above general one. Let us consider two pairs of orthogonal states, \(\{|0\rangle, |1\rangle\}\) and its conjugate \(\{|\overline{0}\rangle, |\overline{1}\rangle\}\) in \(\mathbb{C}^2\). They are assumed mutually unbiased and thus

\[ |\langle i | j \rangle|^2 = \frac{1}{2} \]

holds for each \(i, j = 0, 1\). Alice has \(N\)-qubits described by a Hilbert space \(\mathcal{H}_A = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2\) (\(N\) times). For each \(i = i_1i_2\cdots i_N \in \{0, 1\}^N\), we write \(|i\rangle = |i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_N\rangle\) and \(|\overline{i}\rangle := |\overline{i_1}\rangle \otimes |\overline{i_2}\rangle \otimes \cdots \otimes |\overline{i_N}\rangle\). She prepares a maximally mixed state \(\rho = \frac{1}{2^N}\) of this system in one of the two different methods:

(a) she prepares a state \(|i\rangle\langle i|\) with probability \(\frac{1}{2^N}\) for each \(i \in \{0, 1\}^N\), or
(b) she prepares a state \(|\overline{j}\rangle\langle \overline{j}|\) with probability \(\frac{1}{2^N}\) for each \(j \in \{0, 1\}^N\). We write a random variable \(A\) which takes value \(i \in \{0, 1\}^N\) with probability \(\frac{1}{2^N}\). Alice encodes this random variable to quantum state \(\rho\) by one of the methods (a) or (b).

(i) Alice first selects (a) or (b) which is used to encode a random number.

(ii) Alice encodes the random variable \(A\) to the state \(\rho = \frac{1}{2^N}\) according to her choice on the method. That is, if she has chosen (a), Alice prepares \(|i\rangle\langle i|\) with probability \(\frac{1}{2^N}\) for each \(i \in \{0, 1\}^N\). On the other hand, if her choice was (b), she prepares \(|\overline{j}\rangle\langle \overline{j}|\) with probability \(\frac{1}{2^N}\).

(iii) Alice sends the system to Bob.

(iv) Alice, after confirming that Bob actually has received \(N\)-qubits, informs him of the method ((a) or (b)) she used.

(v) Bob makes a measurement with respect to the basis and obtains an outcome. Let us write \(B\) the random variable representing this outcome. If there is no eavesdropper, \(A = B\) naturally follows.

Eve wants to obtain the information of the random variable \(A\). For the purpose, Eve prepares an apparatus and makes it interact with the \(N\)-qubits sent to Bob by Alice. Denote \(\mathcal{H}_E\) the Hilbert space of Eve’s apparatus. Eve’s operation is described by a unitary operator \(U : \mathcal{H}_B \otimes \mathcal{H}_E \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E\). A state of the apparatus before the interaction is written \(|\Omega\rangle \in \mathcal{H}_E\). Thus Eve’s attack is determined by the triplet, \((\mathcal{H}_E, |\Omega\rangle, U)\). After the publication of the basis, Eve tries to make an optimal measurement \(Z = \{Z_z\}\), a POVM (positive operator valued measure), on her apparatus to extract the information of \(A\).
What we are interested in is the trade-off between the information gain by Eve and the errors contained in Bob’s outcome. Let us suppose a fixed Eve’s attack. We define $I(A : E|a)$ as Eve’s optimal information gain on $A$ if Alice has chosen the method $(a)$ for encoding. $I(A : E|b)$ is defined as Eve’s optimal information gain on $A$ if Alice has chosen the method $(b)$. We can show the following theorem.

**Theorem 3** Information gain by Eve inevitably makes Bob’s data in another basis random. More precisely, for the fixed Eve’s attack $(H_E, |\Omega\rangle, U)$, the following inequality holds:

$$I(A : E|a) \leq H(A \oplus B|b),$$

where $H(A \oplus B|a)$ is the Shannon entropy of the error contained in Bob’s outcome when Alice has chosen the method $(a)$ for encoding.

**Proof:** For the proof, we employ the remote ensemble preparation technique. Since the state prepared by Alice is maximally mixed state, the relevant entangled state of $H_A \otimes H_B$ is maximally entangled state: $|\Phi\rangle := \frac{1}{\sqrt{N}} \sum_i |i\rangle \otimes |i\rangle$. The effect of Eve’s attack can be included by defining a new state,

$$|\Psi\rangle := (1 \otimes U)|\Phi\rangle \otimes |\Omega\rangle.$$

Now suppose that Eve employed a POVM $Z := \{Z_z\}$ and obtained a value $z$. We write $\rho_z$, a-posteriori state on $H_A \otimes H_B$. To this state, we apply the entropic uncertainty relation. To introduce relevant POVMs, we fix a basis to define transpose operation as $\{|i\rangle\}$. It is convenient to introduce a new basis $\{|\tilde{i}\rangle\}$ as $|\tilde{i}\rangle := \sum_j |j\rangle \langle \tilde{j}|$. In fact, the transposition of $|\tilde{i}\rangle \langle \tilde{j}| = \sum_{i,k} |j\rangle \langle i| \langle \tilde{i}| \langle \tilde{k}|$ with respect to the basis $\{|i\rangle\}$ can be simply written as, $|\tilde{i}\rangle \langle \tilde{j}|$. Let us define observables $F_A$ and $F_A$ on $H_A$ as $F_A := \sum_{i\in\{0,1\}^N} i |i\rangle \langle i|$ and $F_A := \sum_{k\in\{0,1\}^N} k |k\rangle \langle k|$. Let us define observable $G_{ZF}$ on $H_B$ as $G_{ZF} := \sum_j |j\rangle \langle j|$. Observables to be treated are $F_A \oplus G_{ZF} = \sum_j |E_j\rangle \langle E_j|$ that gives probability distribution for $A \oplus B$ in $(b)$ and $F_A \otimes 1 = \sum_j j |P_j\rangle \langle P_j|$ that gives probability distribution for $A$ in $(a)$. The following inequality holds:

$$H(F_A \oplus G_{ZF}|Z = z) + H(F_A \otimes 1|Z = z) \geq -2 \log \left( \max_{l,j} \|E_l P_j\| \right). \quad (2)$$

Thus we must estimate $\|E_l P_j\|$. From

$$F_A \oplus G_{ZF} = \sum_{l,i} |l\rangle \langle l| \otimes |\overline{i} \oplus \overline{l}\rangle \langle \overline{i} \oplus \overline{l}|,$$

we obtain $E_l = \sum_i |\tilde{i}\rangle \langle \tilde{i}| \otimes |\overline{i} \oplus \overline{l}\rangle \langle \overline{i} \oplus \overline{l}|$. Therefore

$$E_l P_j = \sum_i |\tilde{i}\rangle \langle j| \otimes |\overline{i} \oplus \overline{l}\rangle \langle \overline{i} \oplus \overline{l}|.$$
holds. To estimate the norm of this operator, we introduce a normalized vector $|\Phi\rangle := \sum_{k\mu} \alpha_{k\mu} |k\rangle \otimes |\mu\rangle$ and apply $E_i P_j$ on it.

$$E_i P_j |\Phi\rangle = \sum_i \alpha_{ji\oplus l} |\bar{\nu}_i\rangle \otimes |\bar{\nu}_i\oplus l\rangle \langle \bar{\nu}_j|$$

gives

$$\|E_i P_j |\Phi\rangle\|^2 = \sum_i |\alpha_{ji\oplus l}|^2 |\langle \bar{\nu}_j|\rangle|^2 \leq \max_i |\langle \bar{\nu}_j|\rangle|^2 \sum_i |\alpha_{ji\oplus l}|^2 \leq \max_i |\langle \bar{\nu}_j|\rangle|^2.$$ 

Thanks to $|\langle \bar{\nu}_j|\rangle| = |\langle \bar{\nu}_l|\rangle|$, we obtain

$$\max_{ij} \|E_i P_j\| \leq \left(\frac{1}{2}\right)^{N/2}.$$ 

Application of this inequality to (2) leads us

$$H(F_A \oplus G_{1B}|Z = z) + H(F_A \otimes 1 |Z = z) \geq N.$$ 

Taking an average with respect to $z$ and adding $N$ to both sides, we obtain

$$I(A : Z|a) \leq H(A \oplus B|b).$$ 

Since the POVM $Z$ is arbitrary, we obtain the theorem. Q.E.D.

5 Summary

In conclusion, we derived a generalization of the information-disturbance theorem. Our generalized theorem can treat a general source (a pair of ensembles that give the same state) and relate Eve’s information gain for an ensemble with Bob’s information gain for another ensemble. The result is a direct consequence of the entropic uncertainty relation.

Acknowledgment

The authors would like to thank Kentaro Imafuku for helpful discussions.

References

[1] C. H. Bennett and G. Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. In Proc. of IEEE Int. Conf. on Computers, Systems and Signal Processing, pages 175–179, 1984.
[2] D. Mayers. Quantum key distribution and string oblivious transfer in noisy channel. In Advances in cryptology - CRYPTO'96, LNCS 1109, pages 343–357, 1996.

[3] H-K. Lo and H-F. Chau. Science, 283, pages 2050–2056, 1999.

[4] P. W. Shor and J. Preskill. Phys.Rev.Lett., 85, pages 441–444, 2000.

[5] E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury. A proof of the security of quantum key distribution. in Proc. of the 32nd Annual ACM Symposium on Theory of Computing, pages 715–724, 2000. : E. Biham, M. Boyer, P. O. Boykin, T. Mor, and V. Roychowdhury.

[6] C. A. Fuchs and A. Peres. Phys.Rev.A, 53(4), pages 2038–2045, 1996.

[7] C. A. Fuchs. Fortschrritte der Physik, 46(4,5), pages 535–565. 1998.

[8] M. Christandl and A. Winter. IEEE Trans Inf Theory, 51(9), pages 3159–3165, 2005.

[9] P. O. Boykin and V. P. Roychowdhury. QIC: Quantum Information and Computation, 5(5), pages 396–412, 2005.

[10] T. Miyadera and H. Imai, Phys.Rev.A. 73, pages 042317 2006.

[11] M. A. Nielsen, and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge press. 2000.

[12] D. Deutsch, Phys.Rev.Lett. 50.631 (1983).

[13] H. Maassen, and J. Uffink, Phys. Rev. Lett., 60, pages 1103 (1998).

[14] M. Krishna and K. R. Parthasarathy, Sankhya, Series A, 64(3), 842 (2002).

[15] L. Hughston, R. Jozsa, and W. Wootters, Phys. Lett. A. 183 pages 14 (1993).

[16] H. Halvorson, J. Math. Phys. 45, pages 4920 (2004).

[17] It was derived independently by M. Hayashi: M. Hayashi, Phys. Rev. A. 74, pages 022307 2006.