ISSN 2186-6570

# Coupling Lemma and Its Application to The Security Analysis of Quantum Key Distribution

Kentaro Kato

Quantum Communication Research Center
Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610, Japan

Tamagawa University Quantum ICT Research Institute Bulletin, Vol.4, No.1, 23-30, 2014

Tamagawa University Quantum ICT Research Institute Bulletin
Vol.4 No.1 : 23-30 (2014)


# Coupling Lemma and Its Application to The Security Analysis of Quantum Key Distribution

Kentaro Kato
Quantum Communication Research Center
Quantum ICT Research Institute, Tamagawa University
6-1-1 Tamagawa-gakuen, Machida, Tokyo 194-8610 Japan
E-mail: kkatop@lab.tamagawa.ac.jp

***Abstract*—It is known that the coupling lemma [1] provides a useful tool in the study of probability theory and its related areas; it describes the relation between the variational distance of two probability distributions and the probability that outcomes from the two random experiments associated with each distribution are not identical. In this paper, the failure probability interpretation problem that has been presented by Yuen and Hirota is discussed from the viewpoint of the application of the coupling lemma. First, we introduce the coupling lemma, and investigate properties of it. Next, it is shown that the claims for this problem in the literatures [10], [11] are justified by using the coupling lemma. Consequently, we see that the failure probability interpretation is not adequate in the security analysis of quantum key distribution.**

## I. Introduction

In the theoretical studies on quantum key distribution (QKD) done in the last decade, the trace distance criterion is widely used for evaluating the QKD system for one-time pad (*e.g.* [2], [3], [4], [5], [6], [7], [8], [9]). In the trace distance criterion, the security notion "$\varepsilon$-secure" is defined by

$$d = \frac{1}{2}\|\hat{\boldsymbol{\rho}}_{KE} - \hat{\rho}_U \otimes \hat{\rho}_E\| \leq \varepsilon, \tag{1}$$

where $\hat{\boldsymbol{\rho}}_{KE}$ corresponds to the *real* system of the key and environment, and $\hat{\rho}_U \otimes \hat{\rho}_E$ the *ideal* one. Further, it is claimed that the parameter $\varepsilon$ is interpreted as the so-called (maximal) failure probability of the QKD protocol. For this failure probability interpretation of the parameter $\varepsilon$, Yuen and Hirota have respectively presented their objection, together with various other theoretical lack in the security analysis of QKD (*e.g.* [10], [11], [12], [13], [14], [15], [16], [17]). In this paper, we treat this failure probability interpretation problem.

Tracking back through the development history of the trace distance criterion to seek the origin of the failure probability interpretation, one can arrive at Lemma 1 of the literature [2]: "*Let $P$ and $Q$ be two probability distributions. Then there exists a joint probability distribution $P_{XX'}$ such that $P_X = P$, $P_{X'} = Q$, and* $\Pr_{(x,x')\leftarrow P_{XX'}}[x \neq x'] = \delta(P,Q)$", where $\delta(P,Q)$ is the variational distance between $P$ and $Q$. Based on this lemma, they give the failure probability interpretation to $\varepsilon$. However, Yuen has repeatedly claimed that a "for every" statement would be needed rather than the "there exists" statement if it justifies the interpretation, and basically, there is no cause one must choose such favorable distribution other than the independent distribution $P_{XX'} = P \cdot Q$ ([10], [11] and its subsequent papers). But, despite the series of criticisms, the failure probability interpretation is still kept [9]. This situation motivates us to discuss on the failure probability interpretation problem.

To discuss the failure probability interpretation problem from a new point of view, we employ a lemma called the coupling lemma, which was given by Aldous [1]. This lemma describes the relation between the variational distance of two probability distributions and the probability that outcomes from the two random experiments associated with each distribution are not identical. As we will show later, the use of the coupling lemma would provide a clear perspective on this problem.

## II. Coupling Lemma

We first introduce the coupling lemma, together with the definition of the variational distance of two probability distributions. The original statement of the coupling lemma is found in the literature [1] (Lemma 3.6, p.249).

*Definition 1* (variational distance): Let $X$ and $Y$ be random variables that take on values from a finite alphabet $\mathcal{A} = \{a_1, a_2, \ldots, a_N\}$. Let $P_X$ and $P_Y$ denote probability distributions of $X$ and $Y$, respectively. The variational distance between $P_X$ and $P_Y$ is defined by

$$v(P_X, P_Y) = \max_{\mathcal{S}\subseteq\mathcal{A}} \Big[P_X(\mathcal{S}) - P_Y(\mathcal{S})\Big], \tag{2}$$

or equivalently, by

$$v(P_X, P_Y) = \frac{1}{2}\sum_{a\in\mathcal{A}} \Big|P_X(a) - P_Y(a)\Big|. \tag{3}$$

■

*Definition 2* (coupling): There are two probability distributions $P_X$ and $P_Y$ that are defined on a finite set $\mathcal{A}$. Consider a joint probability distribution $P_{XY}$ on $\mathcal{A}^2$ whose marginal distributions are $P_X$ and $P_Y$. We call this $P_{XY}$ a coupling of $P_X$ and $P_Y$. ■

*Theorem 3* (coupling lemma [1]): Suppose that $P_X$ and $P_Y$ are given.

(a) *For every* coupling $P_{XY}$ of $P_X$ and $P_Y$,

$$v(P_X, P_Y) \leq \Pr\{x \neq y\}. \tag{4}$$

(b) *There exists* a coupling $P_{XY}$ such that

$$v(P_X, P_Y) = \Pr\{x \neq y\}. \tag{5}$$

□

*Proof*: One would be able to find the coupling lemma and its proof in the textbooks on probability theory and its applications (*e.g.* [18], [19], [20]). The following proof is obtained by modifying the proof of Theorem A.6 of the literature [9].

(a) Let $P_{XY}$ be an arbitrarily chosen coupling of $P_X$ and $P_Y$. It is clear that

$$P_{XY}(a,a) \leq P_X(a) \quad \text{and} \quad P_{XY}(a,a) \leq P_Y(a)$$

for every $a \in \mathcal{A}$, because of the relationship between a joint probability distribution and the associated marginal distributions. So, we have

$$P_{XY}(a,a) \leq \min\{P_X(a), P_Y(a)\} \quad \forall a \in \mathcal{A}. \tag{6}$$

Note that the converse of the inequality above, $P_{XY}(a,a) > \min\{P_X(a), P_Y(a)\}$, is never established. Summing up the both sides of Eq.(6) with respect to $a$, we have

$$\begin{aligned} \Pr\{x = y\} &= \sum_{a \in \mathcal{A}} P_{XY}(a,a) \\ &\leq \sum_{a \in \mathcal{A}} \min\{P_X(a), P_Y(a)\}. \end{aligned} \tag{7}$$

This inequality is arranged to the following form.

$$\begin{aligned} \Pr\{x \neq y\} &= 1 - \Pr\{x = y\} \\ &\geq 1 - \sum_{a \in \mathcal{A}} \min\{P_X(a), P_Y(a)\} \\ &= \sum_{a \in \mathcal{A}} \Big[ P_X(a) - \min\{P_X(a), P_Y(a)\} \Big]. \end{aligned} \tag{8}$$

Here let us define the following partition of $\mathcal{A}$:

$$\begin{aligned} \mathcal{B} &= \{b : P_X(b) \geq P_Y(b)\}, \\ \overline{\mathcal{B}} &= \{b : P_X(b) < P_Y(b)\}. \end{aligned}$$

Then

$$\begin{aligned} \text{Eq.}(8) &= \sum_{b \in \mathcal{B}} \Big[ P_X(b) - \min\{P_X(b), P_Y(b)\} \Big] \\ &\quad + \sum_{b \in \overline{\mathcal{B}}} \Big[ P_X(b) - \min\{P_X(b), P_Y(b)\} \Big] \\ &= \sum_{b \in \mathcal{B}} \Big[ P_X(b) - P_Y(b) \Big] \\ &= P_X(\mathcal{B}) - P_Y(\mathcal{B}) \\ &= \max_{\mathcal{S} \subseteq \mathcal{A}} \Big[ P_X(\mathcal{S}) - P_Y(\mathcal{S}) \Big] \\ &= v(P_X, P_Y). \end{aligned} \tag{9}$$

This completes the proof for (a).

(b) In the preceding part, we observed that the inequality (4) follows from Eq.(6). This implies that the equality in Eq.(4) is established when the equality in Eq.(6) holds. The question is whether there exists a coupling $P_{XY}$ such that

$$P_{XY}(a,a) = \min\{P_X(a), P_Y(a)\} \quad \forall a \in \mathcal{A}. \tag{10}$$

The existence of such a coupling can be shown by the following constructive method.

step 1. Define $P_{XY}(a,a)$ by Eq.(10). It is clear that $P_{XY}(a,a) \geq 0$ for every $a \in \mathcal{A}$.

step 2. If $v(P_X, P_Y) = 0$, it means that the two random variables, $X$ and $Y$, are equal: $X = Y$. In this case, we let $P_{XY}(a,b) = 0$ for $(a,b) \in \mathcal{A}^2$ such that $a \neq b$. Clearly, $\Pr\{x \neq y\} = 0 = v(P_X, P_Y)$, and

$$\sum_{y \in \mathcal{A}} P_{XY}(a,y) = P_X(a) = P_Y(a) = \sum_{x \in \mathcal{A}} P_{XY}(x,a)$$

for every $a \in \mathcal{A}$.

Next let us consider the case for $v(P_X, P_Y) \neq 0$. In this case, the two random variables, $X$ and $Y$, are not equal: $X \neq Y$. This implies $\Pr\{x \neq y\} > 0$. Then we define

$$P_{XY}(a,b) = \frac{R_X(a) R_Y(b)}{\Pr\{x \neq y\}} \quad \forall (a,b) \in \mathcal{A}^2 \text{ s.t. } a \neq b, \tag{11}$$

where the functions in the numerator are given by

$$R_X(a) = P_X(a) - P_{XY}(a,a) \quad \forall a \in \mathcal{A}, \tag{12}$$

$$R_Y(b) = P_Y(b) - P_{XY}(b,b) \quad \forall b \in \mathcal{A}. \tag{13}$$

Since $R_X(a) \geq 0$ and $R_Y(b) \geq 0$, we see that $P_{XY}(a,b) \geq 0$ for every $(a,b) \in \mathcal{A}^2$. In addition, we observe that

$$\sum_{a \in \mathcal{A}} R_X(a) = \sum_{b \in \mathcal{A}} R_Y(b) = \Pr\{x \neq y\} \tag{14}$$

and

$$R_X(a) R_Y(a) = R_X(b) R_Y(b) = 0 \quad \forall a, b \in \mathcal{A}. \tag{15}$$

By using Eqs.(14) and (15), we can verify that the constructed joint probability distribution has the given distributions $P_X$ and $P_Y$ as the associated marginal

distributions. For any fixed $a \in \mathcal{A}$,

$$\begin{aligned}
&\sum_{b\in\mathcal{A}} P_{XY}(a,b) \\
=\ & \sum_{b:b\neq a} P_{XY}(a,b) + \sum_{b:b=a} P_{XY}(a,b) \\
=\ & \sum_{b:b\neq a} \frac{R_X(a)R_Y(b)}{\Pr\{x\neq y\}} + P_{XY}(a,a) \\
=\ & \frac{R_X(a)}{\Pr\{x\neq y\}}\left(\sum_{b:b\neq a} R_Y(b)\right) + P_{XY}(a,a) \\
=\ & \frac{R_X(a)}{\Pr\{x\neq y\}}\left(\sum_{b\in\mathcal{A}} R_Y(b) - R_Y(a)\right) + P_{XY}(a,a) \\
=\ & \frac{R_X(a)}{\Pr\{x\neq y\}}\left(\Pr\{x\neq y\} - R_Y(a)\right) + P_{XY}(a,a) \\
=\ & R_X(a) - \frac{R_X(a)R_Y(a)}{\Pr\{x\neq y\}} + P_{XY}(a,a) \\
=\ & R_X(a) + P_{XY}(a,a) \\
=\ & P_X(a).
\end{aligned}$$

With the same manner, we also have

$$\sum_{a\in\mathcal{A}} P_{XY}(a,b) = P_Y(b) \quad \forall b \in \mathcal{A}.$$

Thus, it was shown that the joint probability distribution $P_{XY}$ constructed from Eqs.(10) and (11) has the marginal distributions $P_X$ and $P_Y$. Since Eq.(10) holds, we have $v(P_X, P_Y) = \Pr\{x \neq y\}$ with this constructed distribution $P_{XY}$. ■

A concrete example of couplings of $P_X$ and $P_Y$ is shown in the appendix *A*. This illustrates the coupling lemma (*Theorem 3*) in a case of $X \neq Y$.

It should be emphasized that the coupling lemma consists of two parts: (a) "for every" part and (b) "there exists" part. Lemma 1 of the literature [2], Proposition 2.1.1 of the literature [4], and Theorem A.6 of the literature [9], these are essentially identical to the "there exists" part of the coupling lemma. On the other hand, an example of independent joint probability distribution, which was treated in the literatures [10], [11], can be explained by the "for every" part of the coupling lemma. For the later discussion, we treat the independent joint distribution case here again.

Suppose that $X$ and $Y$ are independent. Then the joint probability distribution $P_{XY}$ is given by

$$P_{XY}(a,b) = P_X(a)P_Y(b) \quad \forall (a,b) \in \mathcal{A}^2.$$

Since $0 \leq P_X(a) \leq 1$ and $0 \leq P_Y(b) \leq 1$, we have

$$P_X(a)P_Y(a) \leq \min\{P_X(a), P_Y(a)\} \quad \forall a \in \mathcal{A}.$$

If some $a' \in \mathcal{A}$ satisfies the conditions $0 < P_X(a') < 1$ and $0 < P_Y(a') < 1$, then it is reduced to

$$P_X(a')P_Y(a') < \min\{P_X(a'), P_Y(a')\}. \tag{16}$$

Taking this fact into account, we have

$$1 - \sum_{a\in\mathcal{A}} \min\{P_X(a), P_Y(a)\} < 1 - \sum_{a\in\mathcal{A}} P_X(a)P_Y(a),$$

when at least one $a' \in \mathcal{A}$ satisfies the conditions $0 < P_X(a') < 1$ and $0 < P_Y(a') < 1$. With the help of the calculation of Eq.(9), the above inequality is summarized to the following statement:

*Corollary to Theorem 3*: Suppose that $X$ and $Y$ are independent, and at least one $a' \in \mathcal{A}$ satisfies the conditions $0 < P_X(a') < 1$ and $0 < P_Y(a') < 1$. Then,

$$v(P_X, P_Y) < \Pr\{x \neq y\}. \tag{17}$$

■

The next example was used in the literatures [10], [11] to explain the incorrectness of the failure probability interpretation.

*Example 4* ([10], [11]): Let $\mathcal{A} = \{1, 2, \ldots, N\}$, and $P_X(x) = 1/N\ \forall x \in \mathcal{A}$ and $P_Y(y) = 1/N\ \forall y \in \mathcal{A}$. Suppose that $X$ and $Y$ are independent. Then $P_{XY}(x,y) = P_X(x)P_Y(y) = 1/N^2$ for every $(x,y) \in \mathcal{A}^2$. In this case, we have

$$v(P_X, P_Y) = 0 < 1 - \frac{1}{N} = \Pr\{x \neq y\}$$

for $N \geq 2$. ■

Up to this point, we have considered the case of one-dim random variables. It is easy to extend the case of $v(P_X, P_Y)$ to that of $v(P_{X_1X_2}, P_{Y_1Y_2})$.

*Theorem 5*: Let $(X_1, X_2)$ and $(Y_1, Y_2)$ are two-dim random variables that take on values from the same finite alphabet $\mathcal{A}^2$, where $\mathcal{A} = \{a_1, a_2, \ldots, a_N\}$. Suppose that $P_{X_1X_2}$ and $P_{Y_1Y_2}$ are given.

(a) *For every* coupling $P_{X_1X_2Y_1Y_2}$ of $P_{X_1X_2}$ and $P_{Y_1Y_2}$,

$$v(P_{X_1X_2}, P_{Y_1Y_2}) \leq \Pr\{(x_1,x_2) \neq (y_1,y_2)\} \tag{18}$$

where

$$\begin{aligned}
&v(P_{X_1X_2}, P_{Y_1Y_2}) \\
&\quad = \frac{1}{2}\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{A}}\Big|P_{X_1X_2}(a,b) - P_{Y_1Y_2}(a,b)\Big|.
\end{aligned}$$

(b) *There exists* a coupling $P_{X_1X_2Y_1Y_2}$ such that

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \Pr\{(x_1,x_2) \neq (y_1,y_2)\}. \tag{19}$$

□

*Proof*: This is due to the inequality

$$\begin{aligned}
&P_{X_1X_2Y_1Y_2}(a,b,a,b) \\
&\quad \leq \min\{P_{X_1X_2}(a,b), P_{Y_1Y_2}(a,b)\}
\end{aligned} \tag{20}$$

for every $(a,b) \in \mathcal{A}^2$. ■

A concrete example of couplings of $P_{X_1X_2}$ and $P_{Y_1Y_2}$ is shown in the appendix *B*. This illustrates the coupling lemma for two-dim random variables (*Theorem 5*) in a case of $(X_1, X_2) \neq (Y_1, Y_2)$.

The reader might recall a description on the variational distance for two-dim random variables $(\tilde{X}, X)$ and $(\tilde{X}, Y)$ in the textbook [21] of Nielsen and Chuang (p.402). From the point of view of the coupling lemma, it can be understood as a special case that has the condition $X_1 = X_2 = Y_1$.

*Corollary to Theorem 5*: When the condition $X_1 = X_2 = Y_1$ is imposed, then

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \Pr\{x_2 \neq y_2\}. \tag{21}$$

□

*Proof*: From the condition $X_1 = X_2 = Y_1$, we have $P_{X_1X_2Y_1Y_2}(x_1, x_2, y_1, b) = 0$ if $x_2 \neq x_1$ or $y_1 \neq x_1$. Hence, for every $a, b \in \mathcal{A}$,

$$\begin{aligned} P_{X_1Y_2}(a,b) &= \sum_{x_2\in\mathcal{A}}\sum_{y_1\in\mathcal{A}} P_{X_1X_2Y_1Y_2}(a,x_2,y_1,b) \\ &= P_{X_1X_2Y_1Y_2}(a,a,a,b); \qquad (22) \\ P_{X_2Y_2}(a,b) &= \sum_{x_1\in\mathcal{A}}\sum_{y_1\in\mathcal{A}} P_{X_1X_2Y_1Y_2}(x_1,a,y_1,b) \\ &= P_{X_1X_2Y_1Y_2}(a,a,a,b); \qquad (23) \\ P_{Y_1Y_2}(a,b) &= \sum_{x_1\in\mathcal{A}}\sum_{x_2\in\mathcal{A}} P_{X_1X_2Y_1Y_2}(x_1,x_2,a,b) \\ &= P_{X_1X_2Y_1Y_2}(a,a,a,b). \qquad (24) \end{aligned}$$

If $a = b$, then

$$\begin{aligned} P_{X_1X_2}(a,b) &= P_{X_1X_2}(a,a) \\ &= \sum_{y_1\in\mathcal{A}}\sum_{y_2\in\mathcal{A}} P_{X_1X_2Y_1Y_2}(a,a,y_1,y_2) \\ &= \sum_{y_2\in\mathcal{A}} P_{X_1X_2Y_1Y_2}(a,a,a,y_2) \\ &\geq P_{X_1X_2Y_1Y_2}(a,a,a,y_2') \quad \forall y_2' \in \mathcal{A}. \end{aligned}$$

This implies

$$P_{X_1X_2}(a,b) \geq P_{X_1X_2Y_1Y_2}(a,a,a,b) \quad \text{if } a = b.$$

Substituting Eq.(24) into this, we have

$$P_{X_1X_2}(a,a) \geq P_{X_1X_2Y_1Y_2}(a,a,a,a) = P_{Y_1Y_2}(a,a).$$

Therefore,

$$\begin{aligned} &P_{X_1X_2Y_1Y_2}(a,a,a,a) \\ &\quad = P_{Y_1Y_2}(a,a) \\ &\quad = \min\{P_{X_1X_2}(a,a), P_{Y_1Y_2}(a,a)\}. \end{aligned} \tag{25}$$

Next we assume that $a \neq b$. It is obvious that $P_{X_1X_2Y_1Y_2}(a,b,a,b) = 0 = P_{X_1X_2}(a,b)$. By using Eq.(24),

$$\begin{aligned} P_{Y_1Y_2}(a,b) &= P_{X_1X_2Y_1Y_2}(a,a,a,b) \\ &\geq 0 \\ &= P_{X_1X_2}(a,b) \\ &= P_{X_1X_2Y_1Y_2}(a,b,a,b). \end{aligned} \tag{26}$$

This yields

$$\begin{aligned} &P_{X_1X_2Y_1Y_2}(a,b,a,b) \\ &\quad = P_{X_1X_2}(a,b) \\ &\quad = \min\{P_{X_1X_2}(a,b), P_{Y_1Y_2}(a,b)\}. \end{aligned} \tag{27}$$

Summarizing Eqs.(25) and (27),

$$\begin{aligned} &P_{X_1X_2Y_1Y_2}(a,b,a,b) \\ &\quad = \min\{P_{X_1X_2}(a,b), P_{Y_1Y_2}(a,b)\} \end{aligned} \tag{28}$$

for every $(a,b) \in \mathcal{A}^2$. From the comparison between this and Eq.(20), we find that

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \Pr\{(x_1,x_2) \neq (y_1,y_2)\}.$$

Moreover, since $X_1 = X_2 = Y_1$, the probability $\Pr\{(x_1,x_2) \neq (y_1,y_2)\}$ is reduced to $\Pr\{x_2 \neq y_2\}$. Thus, we proved that $X_1 = X_2 = Y_1$ yields the equality of $v(P_{X_1X_2}, P_{Y_1Y_2})$ and $\Pr\{x_2 \neq y_2\}$. ■

Finally, let us summarize the relationship between the variational distance and the probability of $x \neq y$.

1) The variational distance $v(P_X, P_Y)$ does not always mean the probability $\Pr\{x \neq y\}$. $v(P_{X_1X_2}, P_{Y_1Y_2})$ too. Basically, the variational distance $v(P_X, P_Y)$ is a *lower* bound of $\Pr\{x \neq y\}$.
2) Suppose that the two random variables $X$ and $Y$ are independent, and some $a' \in \mathcal{A}$ satisfies the conditions $0 < P_X(a') < 1$ and $0 < P_Y(a') < 1$. In this case, we have $v(P_X, P_Y) < \Pr\{x \neq y\}$.
3) Conversely, a maximal coupling of $P_X$ and $P_Y$ that yields $v(P_X, P_Y) = \Pr\{x \neq y\}$ demands correlation between the outcomes of $X$ and $Y$.

## III. On The Failure Probability Interpretation For $\varepsilon$

Let us return to the main course of this paper. In the security analysis of QKD under the trace distance criterion, as mentioned in the section I, "$\varepsilon$-secure" is defined by

$$d = \frac{1}{2}\|\hat{\boldsymbol{\rho}}_{KE} - \hat{\rho}_U \otimes \hat{\rho}_E\| \leq \varepsilon.$$

The problem that Yuen and Hirota have discussed and we discuss here is whether or not the parameter $\varepsilon$ has a meaning of "the failure probability".

In the literatures [2], [3], [4], [5], [6], [7], [8], [9], it is claimed that the operational meaning of $\varepsilon$ is the (maximal) failure probability of the QKD protocol. To give such an interpretation, they first make the inequality

$$v(P_K, P_U) \leq d \leq \varepsilon, \tag{29}$$

where $P_K$ is a probability distribution of the *real* key and $P_U$ is that of the *ideal* key (uniform key). After that, they apply the statement (b) of the coupling lemma to give the failure probability interpretation. If we dare to ignore the physical restrictions and focus only on the mathematical possibilities of the coupling lemma, a coupling $P_{KU}$ of $P_K$ and $P_U$ that yields $v(P_K, P_U) = \Pr\{k \neq u\}$ exists. However, we cannot drop physical restrictions.

The distribution $P_K$ describes the probabilistic behavior of measurement outcomes that are obtained through an appropriate POVM for the system of $\hat{\boldsymbol{\rho}}_{KE}$. In addition, the distribution $P_U$ is obtained by the same POVM. If one supposes that the measurement outcomes of the *real* system and the *ideal* system are characterized by a maximal coupling, it means that measurement outcomes from the two distinct systems are correlated. In other words, it demands that measurement outcomes in the *real* system depends on that in the *ideal* system. There is no clear reason why one must choose such a coupling and why correlation between the *real* keys and the *ideal* keys is allowed.

If the equality of the variational distance $v(P_K, P_U)$ and the probability $\Pr\{k \neq u\}$ were established for *every* coupling of $P_K$ and $P_U$, the failure probability interpretation might be justified. But, we have already seen that the equality does not always hold.

Further, we assume here that the measurement outcomes from the *real* and *ideal* systems are independent. It would be a natural situation, and is physically acceptable. As shown by Shannon [22] (p.681), all keys have to be equally likely to make one-time pad secure. This is reflected to the settings of the *ideal* probability distribution $P_U$. Indeed, $P_U$ is given to be a uniform distribution. So, $P_U$ satisfies the condition $0 < P_U(u) < 1$. On the other hand, if some sequence $k'$ generated by the QKD protocol possesses probability $P_K(k') = 0$ or $P_K(k') = 1$, it is clear that the generated sequence does not work as the secret key for the embedded one-time pad. Therefore, we can assume that every sequence $k$ satisfies the condition $0 < P_K(k) < 1$ without loss of generality. But, as shown in *Corollary to Theorem 3*, the strict inequality

$$v(P_K, P_U) < \Pr\{k \neq u\} \tag{30}$$

is established in this case. The juxtaposition of Eqs.(29) and (30) tells nothing about the relationship between $\varepsilon$ and $\Pr\{k \neq u\}$.

Consequently, we have the following facts:

1) If a maximal coupling of $P_K$ and $P_U$ is employed, correlation between the *real* and *ideal* keys is needed. But, there is no cause one must choose a maximal coupling.
2) $v(P_K, P_U) \leq \Pr\{k \neq u\}$ holds for every coupling of $P_K$ and $P_U$. That is, $v(P_K, P_U)$ is a lower bound of all the possible probabilities $\Pr\{k \neq u\}$.
3) If the *real* and *ideal* keys are statistically independent, $v(P_K, P_U) < \Pr\{k \neq u\}$. Clearly, $v(P_K, P_U)$ does not mean $\Pr\{k \neq u\}$.

Let us recall *Example 4*. In the viewpoint of the application of the coupling lemma, we can understand that this example illustrates all the facts listed above compactly; it immediately shows the fact 3), and by considering the "not independent" cases, one would arrive at the facts 1) and 2). Thus, the use of the coupling lemma justifies the Yuen's claim for the failure probability interpretation problem that are repeatedly stated in the literatures [10], [11] and its subsequent papers. In summary, we can say that $\varepsilon$ does not mean probability at all, in particular, it does not mean "the failure probability".

## IV. Conclusions

This paper is concerned with the failure probability interpretation problem that has been presented by Yuen and Hirota. To discuss the problem, we used the coupling lemma. First, we observed that the variational distance does not always mean probability due to the coupling lemma. In particular, the analysis of the independent distribution case showed that the variational distance of such a case is not a probability. Applying the coupling lemma to the discussions on the failure probability interpretation problem, we could justify the Yuen's claim stated in the literatures [10], [11] and its subsequent papers. As a result, we can conclude that the failure probability interpretation is not adequate in the security analysis of quantum key distribution.

## V. Acknowledgment

The author would like to thank Osamu Hirota for fruitful discussions and suggestions.

## References

[1] D. Aldous, "Random walks on finite groups and rapidly mixing Markov chains," Séminaire de Probabilités XVII 1981/82, Lecture Notes in Mathematics, vol.986, pp.243-297, 1983.
[2] R. Renner and R. König, "Universally composable privacy amplification against quantum adversaries," TCC 2005, LNCS 3378, pp.407-425, 2005.
[3] R. König, R. Renner, A. Briska, and U. Maurer, "Small accessible quantum information does not imply security," Phys. Rev. Lett., vol.98, no.14, 140502, 2007.
[4] R. Renner, "Security of quantum key distribution," Int. J. Quant. Inform., vol.6, no.1, pp.1-127, 2008.
[5] V. Scarani, and R. Renner, "Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way postprocessing," Phys. Rev. Lett., vol.100, no.20, 200501, 2008.
[6] J. Müller-Quade, and R. Renner, "Composability in quantum cryptography," New J. Phys., vol.11, 085006, 2009.
[7] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," Rev. Mod. Phys., vol.81, no.3, pp.1301-1350, 2009.
[8] M. Tomamichel, C. C. W. Lim, N. Gisin, R. Renner, "Tight finite-key analysis for quantum cryptography," Nature Commun, vol.3, 634, 2012.
[9] C. Portmann, and R. Renner, "Cryptographic security of quantum key distribution," quant-ph arXiv:1409.3525

[10] H. P. Yuen, "Key generation: foundations and a new quantum approach," IEEE J. Sel. Top. Quantum Electron., vol.15, no.6, pp.1630-1645, 2009.
[11] H. P. Yuen, "Fundamental quantitative security in quantum key generation," Phys. Rev. A, vol.82, no.6, 062304, 2010.
[12] H. P. Yuen, "Essential elements lacking in security proofs for quantum key distribution." Proc. SPIE, vol.8899, 88990J, 2013.
[13] H. P. Yuen, "On the foundations of quantum key distribution — Reply to Renner and beyond," Tamagawa University Quantum ICT Research Institute Bulletin, vol.3, no.1, pp.1-8, 2013; available at http://www.tamagawa.jp/research/quantum/bulletin/pdf/Tamagawa.Vol.3-1.pdf
[14] H. P. Yuen, "On the nature and claims of quantum key distribution (QKD)," Lecture at Tamagawa University, 5 Dec 2013; available at http://www.tamagawa.jp/research/quantum/openlecture/
[15] O. Hirota, "Incompleteness and limit of quantum key distribution theory - Yuen theory vs Renner theory -," Tamagawa University Quantum ICT Research Institute Bulletin, vol.2, no.1, pp.25-34, 2012; available at http://www.tamagawa.jp/research/quantum/bulletin/pdf/Tamagawa.Vol.2-6.pdf
[16] O. Hirota, "Misconception in theory of quantum key distribution - Reply to Renner -," quant-ph arXiv:1306.1277v1
[17] O. Hirota, "A correct security evaluation of quantum key distribution," quant-ph arXiv:1409.5991v1
[18] T. Lindvall, Lectures On The Coupling Method, Wiley&Sons 1992.
[19] H. Thorisson, Coupling, Stationarity, and Regeneration, Springer 2000.
[20] D. A. Levin, Y. Peres, and E. L. Wilmer, Markov Chains and Mixing Times, AMS 2009.
[21] M. A. Nielsen, and I. L. Chuang, Quantum Computation and Quantum Information, 10th anniversary edition, Cambridge University Press 2010.
[22] C. E. Shannon, "Communication theory of secrecy systems," Bell System Technical Journal, vol.28, no.4, pp.656-715, 1949.

## APPENDIX

### A. Coupling of $P_X$ and $P_Y$

In this section we treat a case of $X \neq Y$ to illustrate the mathematical notion of the coupling of two probability distributions.

Suppose that the distributions of $X$ and $Y$ are respectively given as

$$\begin{aligned} P_X &= (0.10000, 0.20000, 0.30000, 0.40000), \\ P_Y &= (0.25000, 0.25000, 0.25000, 0.25000). \end{aligned}$$

Then we have $v(P_X, P_Y) = 0.20000$.

*1) Case a ($X$ and $Y$ are independent):* If the random variables $X$ and $Y$ are independent, then the joint probability distribution is given as follows.

$$P_{XY}^{(a)} = \begin{pmatrix} 0.02500 & 0.02500 & 0.02500 & 0.02500 \\ 0.05000 & 0.05000 & 0.05000 & 0.05000 \\ 0.07500 & 0.07500 & 0.07500 & 0.07500 \\ 0.10000 & 0.10000 & 0.10000 & 0.10000 \end{pmatrix}.$$

With this joint probability distribution, we have

$$v(P_X, P_Y) = 0.20000 < 0.75000 = \Pr\{x \neq y\}.$$

This is also an example for *Corollary to Theorem 3*. Thus the variational distance $v(P_X, P_Y)$ behaves as a lower bound of the probability $\Pr\{x \neq y\}$ when $X$ and $Y$ are independent.

*2) Case b (maximal coupling):* Using the construction procedure shown in Section II, a maximal coupling of $P_X$ and $P_Y$ is given by

$$P_{XY}^{(b)} = \begin{pmatrix} 0.10000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.20000 & 0.00000 & 0.00000 \\ 0.03750 & 0.01250 & 0.25000 & 0.00000 \\ 0.11250 & 0.03750 & 0.00000 & 0.25000 \end{pmatrix}.$$

In this case, we have

$$v(P_X, P_Y) = 0.20000 = \Pr\{x \neq y\},$$

which is an example of *Theorem 3*(b). Observe that this coupling demands a strong correlation between $X$ and $Y$. Indeed the following events never happen if $X$ and $Y$ obey this coupling:

$$\begin{aligned} &\text{event} \quad (X = 1) \wedge (Y = 2), \\ &\text{event} \quad (X = 1) \wedge (Y = 3), \\ &\text{event} \quad (X = 1) \wedge (Y = 4), \\ &\text{event} \quad (X = 2) \wedge (Y = 1), \\ &\text{event} \quad (X = 2) \wedge (Y = 3), \\ &\text{event} \quad (X = 2) \wedge (Y = 4), \\ &\text{event} \quad (X = 3) \wedge (Y = 4), \text{ and} \\ &\text{event} \quad (X = 4) \wedge (Y = 4). \end{aligned}$$

Thus the half of the possible events never occur.

*3) Case c (not independent, not maximal):* When two probability distributions are given, there are infinitely many couplings in general. *Theorem 3*(a) covers all of the possible couplings, so that no coupling breaks the inequality (4). To see this, let us consider the following joint probability distribution of $P_X$ and $P_Y$.

$$P_{XY}^{(c)} = \begin{pmatrix} 0.06250 & 0.01250 & 0.01250 & 0.01250 \\ 0.02500 & 0.12500 & 0.02500 & 0.02500 \\ 0.05625 & 0.04375 & 0.16250 & 0.03750 \\ 0.10625 & 0.06875 & 0.05000 & 0.17500 \end{pmatrix}.$$

This is neither the case of independent random variables nor the case of maximal couplings. Even in this case, of course, the inequality holds.

$$v(P_X, P_Y) = 0.20000 < 0.47500 = \Pr\{x \neq y\}.$$

This is a typical behavior of the variational distance.

*B. Coupling of $P_{X_1X_2}$ and $P_{Y_1Y_2}$*

This section gives a concrete example of couplings for two two-dim probability distributions.

Suppose that the distributions of $(X_1, X_2)$ and $(Y_1, Y_2)$ are respectively given as

$$P_{X_1X_2} = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}, \tag{31}$$

and

$$P_{Y_1Y_2} = \begin{pmatrix} 1/9 & 2/9 & 0 \\ 1/9 & 1/9 & 1/9 \\ 0 & 1/9 & 2/9 \end{pmatrix}. \tag{32}$$

For these distributions, $v(P_{X_1X_2}, P_{Y_1Y_2}) = 5/9$.

*1) Case a (maximal coupling):* The joint probability $P^{(a)}_{X_1X_2Y_1Y_2}$ in TABLE I (of the next page) is constructed according to the procedure described in Section II. This provides a maximal coupling of $P_{X_1X_2}$ and $P_{Y_1Y_2}$. In this case we have

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \frac{5}{9} = \Pr\{(x_1, x_2) \neq (y_1, y_2)\}.$$

*2) Case b ($X_1 = X_2 = Y_1$ is imposed):* The joint distribution $P^{(b)}_{X_1X_2Y_1Y_2}$ in TABLE II (of the next page) is designed for satisfying the condition $X_1 = X_2 = Y_1$. For this joint distribution, we have

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \frac{5}{9} = \Pr\{(x_1, x_2) \neq (y_1, y_2)\}.$$

Observe that

$$\Pr\{(x_1, x_2) \neq (y_1, y_2)\} = \Pr\{x_2 \neq y_2\} = \frac{5}{9}.$$

Therefore we have

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \frac{5}{9} = \Pr\{x_2 \neq y_2\}.$$

This can also be understood as an example for the case of $X = \tilde{X}$ in the textbook [21] of Nielsen and Chuang.

*3) Case c ($(X_1, X_2)$ and $(Y_1, Y_2)$ are independent):* Let us consider the case that $(X_1, X_2)$ and $(Y_1, Y_2)$ are independent. The joint distribution $P^{(c)}_{X_1X_2Y_1Y_2}$ for this case is shown in TABLE III (of the next page). As expected from *Corollary to Theorem 3*, we have the following strict inequality.

$$v(P_{X_1X_2}, P_{Y_1Y_2}) = \frac{5}{9} < \frac{23}{27} = \Pr\{(x_1, x_2) \neq (y_1, y_2)\}.$$

Note that this example also implies that if the condition $X_1 = X_2$ is only imposed, that is, the condition $X_1 = Y_1$ is removed from the condition $X_1 = X_2 = Y_1$, then the equality of $v(P_{X_1X_2}, P_{Y_1Y_2})$ and $\Pr\{x_2 \neq y_2\}$ is not guaranteed.

TABLE I
MAXIMAL COUPLING $P^{(a)}_{X_1X_2Y_1Y_2}(x_1,x_2,y_1,y_2)$ CONSTRUCTED BY THE PROCEDURE DESCRIBED IN SECTION II.

| | | $Y_1=1$ | | | $Y_1=2$ | | | $Y_1=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ |
| $X_1=1$ | $X_2=1$ | 1/9 | 4/45 | 0 | 2/45 | 0 | 2/45 | 0 | 2/45 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 4/45 | 0 | 2/45 | 1/9 | 2/45 | 0 | 2/45 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 2/45 | 0 | 1/45 | 0 | 1/45 | 0 | 1/45 | 2/9 |

TABLE II
COUPLING $P^{(b)}_{X_1X_2Y_1Y_2}(x_1,x_2,y_1,y_2)$ UNDER THE CONDITION $X_1=X_2=Y_1$

| | | $Y_1=1$ | | | $Y_1=2$ | | | $Y_1=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ |
| $X_1=1$ | $X_2=1$ | 1/9 | 2/9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 1/9 | 1/9 | 1/9 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1/9 | 2/9 |

TABLE III
COUPLING $P^{(c)}_{X_1X_2Y_1Y_2}(x_1,x_2,y_1,y_2)$ UNDER THE CONDITIONS $(X_1,X_2)$ AND $(Y_1,Y_2)$ ARE INDEPENDENT.

| | | $Y_1=1$ | | | $Y_1=2$ | | | $Y_1=3$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ | $Y_2=1$ | $Y_2=2$ | $Y_2=3$ |
| $X_1=1$ | $X_2=1$ | 1/27 | 2/27 | 0 | 1/27 | 1/27 | 1/27 | 0 | 1/27 | 2/27 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 1/27 | 2/27 | 0 | 1/27 | 1/27 | 1/27 | 0 | 1/27 | 2/27 |
| | $X_2=3$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $X_1=2$ | $X_2=1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=2$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $X_2=3$ | 1/27 | 2/27 | 0 | 1/27 | 1/27 | 1/27 | 0 | 1/27 | 2/27 |