Channel estimation for information theoretically secure key agreement with finite number of pilot signals*

Ryutaroh Matsumoto\textsuperscript{1a)} and Masahiro Takeoka\textsuperscript{2}

\textsuperscript{1} Department of Information and Communications Engineering, Tokyo Institute of Technology, 152-8550, Japan
\textsuperscript{2} National Institute of Information and Communications Technology (NICT), Koganei, Tokyo 184-8795, Japan

\textit{a)} ryutaroh@rmatsumoto.org

\textbf{Abstract:} In this letter, we consider the information theoretically secure key agreement where the legitimate users have a virtual eavesdropper to estimate the statistics of wiretapping. We propose a computational procedure which can estimate the virtual eavesdropper's conditional entropy even with the finite number of pilot pulses and without restricting any probability family for the wiretap channel.

\textbf{Keywords:} channel estimation, information theoretic security, key agreement

\textbf{Classification:} Fundamental Theories for Communications

\textbf{References}

[1] U. M. Maurer, “Secret key agreement by public discussion from common information,” \textit{IEEE Trans. Inf. Theory,} vol. 39, no. 3, pp. 733–742, May 1993. \textit{DOI:10.1109/18.256484}

[2] R. Ahlswede and I. Csiszár, “Common randomness in information theory and cryptography—part I: Secret sharing,” \textit{IEEE Trans. Inf. Theory,} vol. 39, no. 4, pp. 1121–1132, July 1993. \textit{DOI:10.1109/18.243431}

[3] H. Endo, M Fujiwara, M. Kitamura, T Ito, M. Toyoshima, Y. Takayama, H. Takenaka, R. Shimizu, N. Laurenti, G. Vallone, P. Villoresi, T. Aoki, and M. Sasaki, “Free-space optical channel estimation for physical layer security,” \textit{Opt. Express,} vol. 24, no. 8, pp. 8940–8955, Aug. 2016. \textit{DOI:10.1364/OE.24.008940}

[4] M. Fujiwara, T. Ito, M. Kitamura, H. Endo, M. Toyoshima, H. Takenaka, Y. Takayama, R. Shimizu, M. Takeoka, and M. Sasaki, “Secret key agreement demonstration over 7.8 km free-space optical channel,” International Conference on Quantum Communication, Measurement and Computing 2016, Singapore, July 2016.

[5] V. Scarani and R. Renner, “Quantum cryptography with finite resources: Unconditional security bound for discrete-variable protocols with one-way

*The result in this letter appeared in Proc. the 39th Symposium on Information Theory and its Applications, Takayama, Japan, Dec. 13–16, 2016.
1 Introduction

Information theoretically secure key agreement was formulated by Maurer [1] and Ahlswede-Csiszar [2]. It has been actively studied, especially for applications in wireless communication (often referred as the physical layer security). In this scenario, a legitimate sender Alice sends signal through a wiretap channel, which is a broadcast channel to a legitimate receiver Bob and an eavesdropper Eve, to share \( n \)-symbol random variables \( X^n, Y^n, Z^n \), which are possessed by Alice, Bob and Eve, respectively. Throughout the letter, we assume that the channel is memoryless and thus \( X^n, Y^n, Z^n \) are generated from an i.i.d. information source. Alice and Bob also use a public communication channel where Eve can detect all the messages but cannot modify or forge them. The goal for Alice and Bob is to share a common random bit sequence (called a secret key) that is almost statistically independent of all information possessed by Eve. Recently, its experimental study is seriously pursued, for example, in a free-space optical (FSO) link [3].

In theory, it is usually assumed that Alice and Bob know the (full or partial) statistics of the wiretap channel or, in other words, the joint distribution \( P_{XZ} \) or \( P_{YZ} \) is available in advance. In practice, however, this is not always the case since Eve usually does not reveal her information to Alice and Bob. Therefore, Alice and Bob have to estimate \( P_{XZ} \) (for simplicity, hereafter we consider only \( P_{XZ} \). However, our procedure can also work for \( P_{YZ} \). To do this in practical FSO setting, Fujiwara et al. [4] proposed a scheme to place a third terminal Charlie near to Bob to experimentally estimate the upper bound on Eve’s information. Precisely, Charlie is under control of Alice and Bob, records his received signal \( Z^n \) and then reconstruct \( P_{XZ} \) to estimate \( P_{XZ} \). See Fig. 1 and its caption for details of the setup. Moreover, ref. [4] constructed this wiretap channel in the field FSO link with an on-off intensity modulated source and experimentally studied the feasibility of the secret key agreement (throughout the letter, we assume that \( Z \) and \( Z' \) are always discrete and finite alphabet since physically the detected signals are always quantized).

In this setting, the security and the key length to be generated heavily rely on how precisely the legitimate users can estimate \( P_{XZ} \) from a finite set of experimental data at Charlie. This is relatively straightforward when one can assume a family of probability distribution for \( P_{XZ} \). However, experimental data in [3, 4] suggest that specifying the statistical family of the free-space turbulence is sometimes not an easy task. In principle, even without restricting the statistical family, one can numerically perform the interval estimation. In practice, however, the required computational resource can easily get too large to perform an efficient numerical optimization. Therefore, the analytical approach is highly desirable.
In this letter, we describe an analytical procedure to compute the secure key length from the finite channel estimation data at Charlie without assuming any restriction on the statistical family of $P_{XZ}$. This is done by applying the technique developed in quantum key distribution (QKD) [5] to our (classical) physical layer security and use of the latest continuity results on entropy and variational distance [6]. The procedure described here would be useful for realistic implementation of the physical layer security in field. Due to the page limitation, we cannot review all the related literatures. The detailed background is for example given in [3] and references therein.

2 Result: the key length derivation via interval estimation

2.1 Application of the QKD security analysis technique

Scarani and Renner [5] proposed a formula for computing secure key length in the BB84 QKD protocol using a finite number of samples. Since the probability theory can be regarded as a special case of the quantum theory by restricting all matrices to diagonal ones, the security analysis in [5] is also capable of determining secure key length for (non-quantum) information theoretically secure key agreement [1, 2]. Since [5] uses trace distance as a security measure and trace distance is a quantum generalization of variational distance in probability theory, our proposed method is also based on the variational distance between the ideal secret key and actual key which is ensured below an arbitrarily specified value $\epsilon$.

Alice, Bob (and Charlie) initially have $(m+n)$ triples of correlated symbols. They randomly choose $m$ triples among $(m+n)$ ones and publicly disclose them for estimation of their joint probability distribution. $m$ triples serve as “pilot symbols” for channel estimation. We denote by $Q_{XZ}$ the empirical joint distribution of $m$ pairs possessed by Alice and Charlie. Remaining $n$ pairs of Alice and Bob’s symbols are used for generating a secret key. Let $\Gamma_{XZ}$ be a confidence set of $P_{XZ}$

Fig. 1. Schematic of the FSO secret key agreement scenario proposed in [4]. Alice sends optical signal in free space to Bob where the beam center is aligned to Bob’s receiver. There is a third receiver, Charlie, who is under control of Alice and Bob. Eve, who may detect the signal to eavesdrop the key, is assumed to be at the outside of the triangle by Alice, Bob, and Charlie (shade in the figure) such that her detected signal is always weaker than that of Charlie. In practice, this can be guaranteed by monitoring the FSO channel by an additional sensor. The legitimate users estimate the key length by treating the channel between Alice, Bob, and Charlie as a wiretap channel.
given by interval estimation and suppose its confidence level \(1 - \epsilon\) satisfies \(\epsilon < \epsilon'\). Then the secure key length (in bits) is given by

\[
n = \left(\min_{P_{XZ} \in \Gamma_{XZ}} H(X|Z) - \text{IR}_{XZ}\right)
\]

plus negligible terms (see [5, Eq. (5)] for exact expressions), where \(\text{IR}_{XZ}\) is the number of bits for information reconciliation (error correction) transmitted over the public channel. As described above, [5] reduces the problem of determination of secure key length to the interval estimation of the joint probability distribution \(P_{XZ}\).

### 2.2 Proposed method of interval estimation to determine the key length

In order to determine the secure key length, the remaining task is to compute

\[
\min_{P_{XZ} \in \Gamma_{XZ}} H(X|Z)
\]

We have

\[
\min_{P_{XZ} \in \Gamma_{XZ}} H(X|Z) \geq \min_{P_{XZ} \in \Gamma_{XZ}} H(X, Z) - \max_{P_{Z} \in \Gamma_{Z}} H(Z), \quad (1)
\]

where \(\Gamma_{Z}\) is a confidence set of \(P_{Z}\). Each term in Eq. (1) is a (one-sided) interval estimate of entropy. To connect them with the variational distance, we employ the recent result on the interval estimation of entropy by Ho and Yeung [6, Section 4].

The proposed method of the secure key length derivation is as follows. Suppose \(\epsilon'\) is the maximum allowable value of the variational distance between the ideal key and the actual key. Let \(X\) and \(Z\) be the alphabets of random variables.

1. Choose a confidence level \(1 - \epsilon\) greater than \(1 - \epsilon'\).
2. By using Theorem 11 of [6], one obtains

\[
\min_{P_{XZ} \in \Gamma_{XZ}} H(X, Z) \geq H(S(Q_{XZ}, \delta)), \quad (2)
\]

and

\[
\max_{P_{Z} \in \Gamma_{Z}} H(Z) \leq H(R(Q_{Z}, |Z|, \delta)), \quad (3)
\]

with confidence level of \(1 - \epsilon\), where

\[
\delta = \sqrt{\frac{2 \ln 2}{\ln m} \log_2 \left(\frac{2^{|X \times Z|} - 2}{\epsilon}\right)}, \quad (4)
\]

and \(S(\cdot)\) and \(R(\cdot)\) are defined in Theorem 2 and 3 of [6], respectively.
3. By using [5, Eq. (5)] and an estimate of \(H(X|Z)\) given by Eqs. (1)–(4), one can determine secret key length with which the generated key has variational distance less than \(\epsilon'\) from the ideal secret key.

### Acknowledgment

This work was funded by ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan). The authors deeply thank Prof. Tomohiko Uyematsu for informing us of [6], and Dr. Masahide Sasaki, Dr. Mikio Fujiwara and Dr. Hiroyuki Endo for helpful discussions. The authors also thank the associate editor Prof. Kazunori Hayashi for detailed reading and helpful comments.