Secret Key Agreement from Vector Gaussian Sources by Rate Limited Public Communication

Shun Watanabe, Member, IEEE, and Yasutada Oohama Member, IEEE,

Abstract—We investigate the secret key agreement from correlated vector Gaussian sources in which the legitimate parties can use the public communication with limited rate. For the class of protocols with the one-way public communication, we show that the optimal trade-off between the rate of key generation and the rate of the public communication is characterized as an optimization problem of a Gaussian random variable. The characterization is derived by using the enhancement technique introduced by Weingarten et. al. for MIMO Gaussian broadcast channel.

Index Terms—Enhancement Technique, Entropy Power Inequality, Extremal Inequality, Key Agreement, Rate Limited Public Communication Privacy Amplification, Vector Gaussian Sources,

I. INTRODUCTION

Key agreement is one of the most important problems in the cryptography, and it has been extensively studied in the information theory for discrete sources (e.g. [1], [2], [3]) since the problem formulation by Maurer [4]. Recently, the confidential message transmission [5], [6] in the MIMO wireless communication has attracted considerable attention as a practical problem setting (e.g. [7], [8], [9], [10], [11], [12], [13], [14], [15]). Although the key agreement in the context of the wireless communication has also attracted considerable attention recently [16], the key agreement from analog sources has not been studied sufficiently compared to the confidential message transmission. As a fundamental case of the key agreement from analog sources, we consider the key agreement from correlated vector Gaussian sources in this paper. More specifically, we consider the problem in which the legitimate parties, Alice and Bob, and an eavesdropper, Eve, have correlated vector Gaussian sources respectively, and Alice and Bob share a secret key from their sources by using the public communication. Recently, the key agreement from Gaussian sources has attracted considerable attention in the context of the quantum key distribution [17], which is also a motivation to investigate the present problem. Fig. 1 illustrates a scenario we are considering.

Typically, the first step of the key agreement protocol from analog sources is the quantization of the sources. In literatures (e.g. see [16], [18], [19]), the authors used the scalar quantizer, i.e., the observed source is quantized in each time instant. Using the finer quantization, we can expect the higher key rate in the protocol, where the key rate is the ratio between the length of the shared key and the block length of the sources that are used in the protocol. However, there is a problem such that the finer quantization might increase the rate of the public communication in the protocol. Although the public communication is usually regarded as a cheap resource in the context of the key agreement problem, it is limited by a certain amount in practice. Therefore, we consider the key agreement protocols with the rate limited public communication in this paper. The purpose of this paper is to clarify the trade-off between the key rate and the public communication rate of the key agreement protocol from vector Gaussian sources.

The key agreement by rate limited public communication was first considered by Csiszár and Narayan for discrete sources [2]. For the class of protocols with one-way public communication, they characterized the optimal trade-off between the key rate and the public communication rate in terms of the information theoretic quantities, i.e., they derived the so-called single letter characterization. However, there are two difficulties to extend their result to the vector Gaussian sources.

First, the direct part of the proof in [2] heavily relies on the finiteness of the alphabets of the sources, and cannot be applied to continuous sources. This difficulty was solved by
the authors in [20], and this result will also be used in this paper.

Second, although the converse part of Csiszár and Narayan’s characterization can be easily extended to continuous sources, the characterization is not computable because the characterization involves auxiliary random variables and the ranges of those random variables are unbounded for continuous sources.

In [20] for scalar Gaussian sources, the authors showed that Gaussian auxiliary random variables suffice, and derived a closed form expression of the optimal trade-off. In the problem for scalar Gaussian sources, we first solved the problem in which the sources are degraded, i.e., Alice’s source, Bob’s source, and Eve’s source form a Markov chain in this order. Then, we reduced the general case to the degraded case by using the fact that scalar Gaussian correlated sources are stochastically degraded [21].

In this paper for vector Gaussian sources, we show that Gaussian auxiliary random variables suffice, and characterize the optimal trade-off in terms of the (covariance) matrix optimization problem. One of difficulties to show our result is that vector Gaussian sources are not stochastically degraded in general, and cannot be reduced to the degraded case in the same manner as scalar Gaussian sources. To circumvent this difficulty, we utilize the enhancement technique introduced by Weingarten et al. [22].

The rest of the paper is organized as follows: In Section II we explain our problem formulation. In Section III we show our main results and some numerical examples. In Sections IV and V our main results are proved. Finally, in Section VI the conclusion and the future research agenda are discussed.

II. PROBLEM FORMULATION

Let $X, Y$, and $Z$ be correlated vector Gaussian sources on $\mathbb{R}^{m_x}$, $\mathbb{R}^{m_y}$, and $\mathbb{R}^{m_z}$ respectively, where $\mathbb{R}$ is the set of real numbers. Then, let $X^n$, $Y^n$, and $Z^n$ be i.i.d. copies of $X$, $Y$, and $Z$ respectively. Throughout the paper, upper case letters indicate random variables, and the corresponding lower case letters indicate their realizations. We also use the following notations throughout the paper: $\Sigma$ designates the covariance matrix of $(X, Y, Z)$, $\Sigma_x$, $\Sigma_{xy}$, and $\Sigma_{yx}$ designate $\mathbb{E}[X^T X]$, $\mathbb{E}[X^T Y]$, and the conditional covariance of $Y$ given $X$ etc.. $N \sim \mathcal{N}(0, A)$ means that the random variable $N$ is a Gaussian vector with zero mean and covariance matrix $A$. We use $|A|$ to denote the determinant of the matrix $A$, $\frac{A^2}{2}$ to denote $\frac{|A|}{2}$, and we denote $A \succeq B$ ($A \preceq B$) if the matrix $B - A$ is positive semidefinite (definite). Throughout the paper, we assume that $\Sigma > 0$.

Although Alice and Bob can use public communication interactively in general, we concentrate on the class of key agreement protocols in which only Alice sends a message to Bob over the public channel. First, Alice computes the message $C_n$ from $X^n$ and sends the message to Bob over the public channel. Then, she also compute the key $S_n$. Bob compute the key $S'_n$ from $Y^n$ and $C_n$. Fig. 2 illustrates the protocol with one-way public communication.

The error probability of the protocol is defined by

$$\varepsilon_n := \Pr\{S_n \neq S'_n\}.$$
Theorem 1 Let $\mathcal{R}_G(X,Y,Z)$ be the set of all rate pairs $(R_p, R_k)$ satisfying
\[
R_p \geq \frac{1}{2} \log \left| \frac{\Sigma_x}{\Sigma_x} - \frac{1}{2} \log \left| \frac{B \Sigma_u B^T + I}{B \Sigma_u B^T + I} \right| \right|
\]
\[
R_k \leq \frac{1}{2} \log \left| \frac{2 \Sigma_x + 2 \Sigma_{xy}}{2 \Sigma_x + 2 \Sigma_{xy}} - \frac{1}{2} \log \left| \frac{E \Sigma_u E^T + I}{E \Sigma_u E^T + I} \right| \right|
\]

for some $0 < \Sigma_{xy} < \Sigma_x$. Then, we have
\[
\mathcal{R}(X,Y,Z) = \mathcal{R}_G(X,Y,Z).
\]

We are also interested in the asymptotic behavior of the function
\[
R_k(R_p) := \sup \{ R_k : (R_p, R_k) \in \mathcal{R}(X,Y,Z) \}. \tag{1}
\]

Following the approach in [II], we can obtain a closed form expression of $\lim_{R_p \to \infty} R_k(R_p)$ as follows. Let $\phi_i, i = 1, \ldots, m_e$ be the generalized eigenvalues [23 Chapter 6.3] of the matrices
\[
\left( \Sigma_x B^T B \Sigma_x + I_{m_x}, \Sigma_x E^T E \Sigma_x + I_{m_x} \right).
\]
Without loss of generality, we may assume that these generalized eigenvalues are ordered as
\[
\phi_1 \geq \cdots \geq \phi_{\rho} > 1 \geq \phi_{\rho+1} \geq \cdots \geq \phi_{m_e}, \tag{2}
\]
i.e., a total of $\rho$ of them are assumed to be greater than 1. Then, we have
\[
\lim_{R_p \to \infty} R_k(R_p) = \max_{0 < \Sigma_{xy} < \Sigma_x} \left\{ \frac{1}{2} \log \left| \frac{B \Sigma_u B^T + I}{B \Sigma_u B^T + I} \right| - \frac{1}{2} \log \left| \frac{E \Sigma_u E^T + I}{E \Sigma_u E^T + I} \right| \right\}
\]
\[
= \frac{1}{2} \sum_{i=1}^{\rho} \log \phi_i. \tag{3}
\]

Since Eq. (3) can be proved almost in the same manner as [II] Theorem 3], we omit a proof.

When $m_x = m_y = m_z$ and both $B$ and $E$ are invertible, it suffices to consider the case in which
\[
Y = X + W_y, \tag{4}
\]
\[
Z = X + W_z, \tag{5}
\]
where the covariance matrices $\Sigma_{W_y}$ and $\Sigma_{W_z}$ are not necessarily identity but are invertible. Following [22], we call this case the aligned case. As is usual with the vector Gaussian problems (e.g. [22]), the general statement (Theorem 1) is shown by detouring the statement for the aligned case.

Theorem 2 Let $\mathcal{R}_G^a(X,Y,Z)$ be the set of all rate pairs $(R_p, R_k)$ satisfying
\[
R_p \geq I_p(\Sigma_{xy}), \tag{6}
\]
\[
r_k \leq I_k(\Sigma_{xy}), \tag{6}
\]

for some $0 < \Sigma_{xy} < \Sigma_x$. Then, we have
\[
\mathcal{R}(X,Y,Z) = \mathcal{R}_G(X,Y,Z).
\]

Theorem 2 is shown in Section VI and Theorem 1 is shown in Section V by using Theorem 2.

B. Numerical Examples

In this section, we show some numerical example to illustrate Theorem 1. In general, calculation of $\mathcal{R}_G(X,Y,Z)$ involves a nonconvex optimization problem and is not tractable. However for $m_x \geq 2$ and $m_y = m_z = 1$, following the method in [24] (see also [10]), we can transform the calculation of $\mathcal{R}_G(X,Y,Z)$ into tractable form. For $m_x \geq 2$ and $m_y = m_z = 1$, we have
\[
I_p(\Sigma_{xy}) = \frac{1}{2} \log \left| \frac{\Sigma_x}{\Sigma_{xy}} \right| - \frac{1}{2} \log \left( B \Sigma_{xy} B^T + I \right),
\]
\[
I_k(\Sigma_{xy}) = \frac{1}{2} \log \left| \frac{\Sigma_x}{\Sigma_{xy}} \right| - \frac{1}{2} \log \left( E \Sigma_{xy} E^T + I \right),
\]
where $b, e \in \mathbb{R}^{m_x}$. Noting the relation
\[
\frac{e \Sigma_{xy} e^T + 1}{b \Sigma_{xy} b^T + 1} = 1 + \frac{e \Sigma_{xy} e^T - b \Sigma_{xy} b^T + 1}{b \Sigma_{xy} b^T + 1},
\]
we set
\[
\begin{align*}
\phi_1 \geq & \cdots \geq \phi_{\rho} > 1 \geq \phi_{\rho+1} \geq \cdots \geq \phi_{m_e}, \tag{2}
\end{align*}
\]

Then we can easily find that
\[
\mathcal{R}_G(X,Y,Z) = \{(R_p, R_k) : R_p \geq I_p(\Sigma_{xy}), \quad R_k \leq I_k(\Sigma_{xy}), \quad 0 < \Sigma_{xy} \leq \Sigma_x \}
\]

For fixed $(s, t)$, the optimization problem

\[
\begin{align*}
\text{minimize} & & I_p(\Sigma_{xy}), \\
\text{subject to} & & t(b \Sigma_{xy} b^T + 1) \leq e \Sigma_{xy} e^T - b \Sigma_{xy} b^T, \\
& & b \Sigma_{xy} b^T \leq s, \\
& & 0 \leq \Sigma_{xy} \leq \Sigma_x
\end{align*}
\]

is a convex problem. By sweeping $(s, t)$, we can calculate the region $\mathcal{R}_G(X,Y,Z)$. For
\[
\Sigma_x = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 & 0.5 \\ 0.7 & 0.35 \end{bmatrix},
\]

(6)
In this case, the function \( \mathcal{Z} \) is defined in Eq. (3) instead of \( \mathcal{G} \).

The region \( \mathcal{R}_G(X, Y, Z) \) is plotted in Fig. 3. Note that this case is degraded in the sense of [25] Definition 1, i.e., \( X \leftrightarrow Y \leftrightarrow Z \) by appropriately choosing the correlation between \( (Y, Z) \).

In this example is not degraded. Although \( I(X; Y) - I(X; Z) = 0 \) in this example, Fig. 4 clarifies that appropriate quantization enables Alice and Bob to share a secret key at positive key rate.

For non-degraded case, \( R_k(R_p) \) converges to the quantity given by Eq. 5 instead of \( I(X; Y) - I(X; Z) \) as \( R_p \) increases, and it is also plotted in Fig. 4.

IV. PROOF OF THEOREM 2

A. Direct Part

In [20], the present authors proved the following proposition, which is an extension of [2, Theorem 2.6] to continuous sources.

Proposition 3 For an auxiliary random variable \( U \) satisfying the Markov chain

\[ U \leftrightarrow X \leftrightarrow (Y, Z), \]

let \( (R_p, R_k) \) be a rate pair such that

\[ R_p \geq I(U; X) - I(U; Y), \]
\[ R_k \leq I(U; Y) - I(U; Z). \]

Then, we have \( (R_p, R_k) \in \mathcal{R}(X, Y, Z) \).

The direct part of Theorem 2 is shown by taking Gaussian auxiliary random variable \( U \) such that the conditional covariance matrix of \( X \) given \( U \) is \( \Sigma_{x|u} \) in Proposition 3.

B. Converse Part

In the converse proof, we will use the following Proposition and Corollary. The proposition was shown for discrete sources in [2, Theorem 2.6], and it can be shown almost in the same manner for continuous sources.

Proposition 4 (2) Suppose that a rate pair \( (R_p, R_k) \) is included in \( \mathcal{R}(X, Y, Z) \). Then, there exist auxiliary random variables \( U \) and \( V \) satisfying

\[ R_p \geq I(U; X|Y), \] \( (8) \)
\[ R_k \leq I(U; Y|V) - I(U; Z|V), \] \( (9) \)

and the Markov chain

\[ V \leftrightarrow U \leftrightarrow X \leftrightarrow (Y, Z). \] \( (10) \)

For degraded sources, we can simplify the above proposition (see [20, Appendix B] for a proof).

Corollary 5 Suppose that \((X, Y, Z)\) is degraded, i.e., \( X \leftrightarrow Y \leftrightarrow Z \). If \((R_p, R_k) \in \mathcal{R}(X, Y, Z)\), then there exists an auxiliary random variable \( U \) satisfying

\[ R_p \geq I(U; X|Y) = I(U; X) - I(U; Y), \] \( (11) \)
\[ R_k \leq I(U; Y|Z) = I(U; Y) - I(U; Z), \] \( (12) \)

and the Markov chain

\[ U \leftrightarrow X \leftrightarrow Y \leftrightarrow Z. \] \( (13) \)

We show a converse proof of Theorem 2 by contradiction. Suppose that there exists a rate pair such that \((R_p^0, R_k^0) \in \mathcal{R}(X, Y, Z) \) and \((R_p^0, R_k^0) \notin \mathcal{R}_G(X, Y, Z)\), where we assume \( R_k^0 > 0 \) to avoid the trivial case. Then, there exists \( 0 < \Sigma_{x|u}^0 \leq \Sigma_x \) such that \( I_p(\Sigma_{x|u}^0) \leq R_p^0 \). Therefore, we can write

\[ R_k^0 = R_k^0 + \delta \] \( (14) \)
for some $\delta > 0$, where $R^*_k$ is given by the optimal value of

$$\text{maximize} \quad I_k(\Sigma_{x|u}^*)$$

$$\text{subject to} \quad I_p(\Sigma_{x|u}^*) \leq R^*_p,$$

$$0 < \Sigma_{x|u}^* \leq \Sigma_x.$$  \hspace{1cm} (15)

An optimal solution $\Sigma_{x|u}^*$ of this optimization problem satisfies the Karash-Kuhn-Tucker (KKT) condition (see Appendix A for the derivation)

$$\mu(\Sigma_{x|u}^*)^{-1} + (\Sigma_{x|u}^* + \Sigma_{W_y})^{-1} = (1 + \mu)(\Sigma_{x|u}^* + \Sigma_{W_y})^{-1} + M,$$

$$M(\Sigma_x - \Sigma_{x|u}^*) = 0,$$

$$\mu(R^*_p - I_p(\Sigma_{x|u}^*)) = 0,$$  \hspace{1cm} (16-18)

where $\mu \geq 0$ and $M \geq 0$. From Eqs. (16) and (17), we have

$$R^*_k - \mu R^*_p = I_k(\Sigma_{x|u}^*) - \mu I_p(\Sigma_{x|u}^*) + \delta.$$  \hspace{1cm} (19)

We shall find a contradiction to Eq. (19) by showing that for any $(R_p, R_k) \in \mathcal{R}(X, Y, Z)$

$$R_k - \mu R_p \leq I_k(\Sigma_{x|u}^*) - \mu I_p(\Sigma_{x|u}^*).$$  \hspace{1cm} (20)

The proof of Eq. (20) roughly consists of three steps: In the first step, we reduce the proof for the non-degraded sources to that for the degraded sources by using the enhancement technique introduced by Weingarten et al. \cite{22}. In the second step, we change the variable so that we can use the entropy power inequality (EPI). In the last step, we derive an upper bound on $R_k - \mu R_p$ by using the EPI, which turn out to be tight.

**Step 1:** In this step, in order to reduce the proof for the non-degraded sources to that for the degraded sources, we introduce the covariance matrix $\Sigma_{\tilde{W}_y}$ satisfying

$$(1 + \mu)(\Sigma_{x|u}^* + \Sigma_{\tilde{W}_y})^{-1} = (1 + \mu)(\Sigma_{x|u}^* + \Sigma_{W_y})^{-1} + M.$$  \hspace{1cm} (21)

Then, we have (see Appendix B for a proof)

$$0 < \Sigma_{\tilde{W}_y} \leq \Sigma_{W_y},$$

$$\Sigma_{\tilde{W}_y} \leq \Sigma_{W_y}.$$  \hspace{1cm} (22-23)

Let $\tilde{W}_y$ be the Gaussian random vector whose covariance matrix is $\Sigma_{\tilde{W}_y}$, and let

$$\tilde{Y} = X + \tilde{W}_y.$$  \hspace{1cm} (24)

From Eq. (24), we can find that the sources $(X, \tilde{Y}, Z)$ satisfy $X \leftrightarrow \tilde{Y} \leftrightarrow Z$. Furthermore, from Eq. (22), we can also find that $X \leftrightarrow \tilde{Y} \leftrightarrow Y$, which implies

$$\mathcal{R}(X, Y, Z) \subset \mathcal{R}(X, \tilde{Y}, Z).$$

Thus, it suffice to show that Eq. (20) holds for any $(R_p, R_k) \in \mathcal{R}(X, Y, Z)$. In steps 2 and 3, we will show that

$$R_k - \mu R_p \leq I_k(\Sigma_{x|u}^*) - \mu I_p(\Sigma_{x|u}^*)$$  \hspace{1cm} (25)

for any $(R_p, R_k) \in \mathcal{R}(X, \tilde{Y}, Z)$, where

$$\hat{I}_k(\Sigma_{x|u}^*) := \frac{1}{2} \log \frac{\Sigma_x}{\Sigma_{x|u}^*} - \frac{1}{2} \log \frac{\Sigma_x + \Sigma_{W_y}}{\Sigma_{x|u}^* + \Sigma_{W_y}},$$

$$\hat{I}_p(\Sigma_{x|u}^*) := \frac{1}{2} \log \frac{\Sigma_x + \Sigma_{\tilde{W}_y}}{\Sigma_{x|u}^* + \Sigma_{\tilde{W}_y}} - \frac{1}{2} \log \frac{\Sigma_x + \Sigma_{W_y}}{\Sigma_{x|u}^* + \Sigma_{W_y}}.$$  \hspace{1cm} (26)

Then, by using the relation (see Appendix C for a proof)

$$(\Sigma_x + \Sigma_{W_y})(\Sigma_{x|u}^* + \Sigma_{\tilde{W}_y})^{-1} = (\Sigma_x + \Sigma_{W_y})(\Sigma_{x|u}^* + \Sigma_{W_y})^{-1},$$  \hspace{1cm} (26)

we have $I_k(\Sigma_{x|u}^*) = \hat{I}_k(\Sigma_{x|u}^*)$ and $I_p(\Sigma_{x|u}^*) = \hat{I}_p(\Sigma_{x|u}^*)$. Thus, Eq. (25) implies that Eq. (20) holds for any $(R_p, R_k) \in \mathcal{R}(X, \tilde{Y}, Z)$.

**Step 2:** First, we show Eq. (25) for $\mu = 0$. In this case, from Eqs. (16) and (21), we have $\Sigma_{\tilde{W}_y} = \Sigma_{W_y}$. Thus, from Corollary 5 we have

$$R_k - \mu R_p \leq I(U; Y) - I(U; Z) = 0 = \hat{I}_k(\Sigma_{x|u}^*).$$

Thus, we have the assertion.

In order to prove Eq. (25) for $\mu > 0$, we change the variable as follows. Since $(X, Y, Z)$ is jointly Gaussian, we can write

$$X = K_{xz}Z + N_1,$$

$$\tilde{Y} = K_{\tilde{y}x}X + K_{\tilde{y}z}Z + N_2$$

for Gaussian random vectors $N_1, N_2$ with covariance matrices

$$\Sigma_{N_1} = \Sigma_{xz} := \Sigma_x - K_{xz} \Sigma_{xx},$$

$$\Sigma_{N_2} = \Sigma_{\tilde{y}xz} := \Sigma_{\tilde{y}x} - K_{\tilde{y}x} \Sigma_{xz} - K_{\tilde{y}z} \Sigma_{\tilde{y}z},$$

where the coefficients are given by

$$K_{xz} = \Sigma_{xz} \Sigma_z^{-1}$$

and

$$[ K_{\tilde{y}z} \quad K_{\tilde{y}x} ] = [ \Sigma_{\tilde{y}z} \quad \Sigma_{\tilde{y}x} ] [ \Sigma_{xz} \Sigma_{xx} ]^{-1}. \hspace{1cm} (27)$$

By noting the relations

$$I(U; X|\tilde{Y}) = I(U; X) - I(U; \tilde{Y}),$$

$$I(U; \tilde{Y}|Z) = I(U; \tilde{Y}) - I(U; Z),$$

$$I(U; X|Z) = I(U; X) - I(U; Z),$$

$$I(U; X|\tilde{Y}) = I(U; X|Z) - I(U; \tilde{Y}|Z)$$
for random variables satisfying $U \leftrightarrow X \leftrightarrow Y \leftrightarrow Z$, we have

$$
\tilde{I}_k(\Sigma_{x|u}) - \mu \tilde{I}_p(\Sigma_{x|u}) \\
= I(U; Y|Z) - \mu I(U; X|Y) \\
= [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] \\
+ [\mu h(X|U, Z) - (1 + \mu)h(\tilde{Y}|U, Z)] \\
= [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] - \frac{1 + \mu}{2} \log |K_{gx}K_{gy}^\top| \\
+ \mu [h(X|U, Z) - \gamma h(X + N_3|U, Z)] \\
= [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] - \frac{1 + \mu}{2} \log |K_{gx}K_{gy}^\top| \\
+ \mu [h(X|U, Z) - \gamma h(X + N_3|U, Z)] \\
= [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] - \frac{1 + \mu}{2} \log |K_{gx}K_{gy}^\top| \\
+ \mu [\frac{1}{\gamma} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma_{N_3}] - \frac{\gamma}{2} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma_{N_3}] \\
= (\Sigma_{x|u} + \Sigma_{N_3})^{-1} \gamma - \frac{\gamma}{2} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma_{N_3}] \\
= (\Sigma_{x|u} + \Sigma_{N_3})^{-1}
$$

(28)

where we set $\gamma := \frac{1 + \mu}{\mu} > 1$ and $N_3 := K_{gx}^{-1} N_2$. It should be noted that

$$
|K_{gx}| \neq 0
$$

(29)

for $\mu > 0$, which will be proved in Appendix [D]

For the change of variable

$$
\phi : \Sigma_{x|u} \mapsto \Sigma_{x|u} = (\Sigma_{x|u}^{-1} + \Sigma_{W_2}^{-1})^{-1},
$$

let $\Sigma_{x|u}^* := \phi(\Sigma_{x|u}^*)$. From Eqs. (16) and (21) and the relation

$$
\tilde{I}_k(\Sigma_{x|u}) - \mu \tilde{I}_p(\Sigma_{x|u}) \\
= \frac{\mu}{2} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma_{W_2}| \\
- \frac{1 + \mu}{2} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma_{W_2}| \\
+ (1 + \mu)h(\tilde{Y}) - h(Z) - h(X),
$$

we have

$$
\nabla_{\Sigma_{x|u}} \left[ \tilde{I}_k(\Sigma_{x|u}^*) - \mu \tilde{I}_p(\Sigma_{x|u}^*) \right] = 0.
$$

By the chain rule for the derivative, we have

$$
\nabla_{\Sigma_{x|u}} \left[ \tilde{I}_k(\phi^{-1}(\Sigma_{x|u}^*)) - \mu \tilde{I}_p(\phi^{-1}(\Sigma_{x|u}^*)) \right] \\
= \nabla_{\Sigma_{x|u}} \phi^{-1}(\Sigma_{x|u}^*) \cdot \nabla_{\Sigma_{x|u}} \left[ \tilde{I}_k(\Sigma_{x|u}^*) - \mu \tilde{I}_p(\Sigma_{x|u}^*) \right] \\
= 0.
$$

Thus, from Eq. (28), we have

$$
(\Sigma_{x|u}^*)^{-1} = \gamma(\Sigma_{x|u}^* + \Sigma_{N_3})^{-1}.
$$

(30)

Step 3: By noting that $(X, \tilde{Y}, Z)$ is degraded, from Corollary [5] for any $(R_p, R_k) \in \mathcal{R}(X, \tilde{Y}, Z)$ we have

$$
R_k - \mu R_p \\
\leq I(U; Y|Z) - \mu I(U; X|Y) \\
= [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] - \frac{1 + \mu}{2} \log |K_{gx}K_{gy}^\top| \\
+ \mu [h(X|U, Z) - \gamma h(X + N_3|U, Z)],
$$

(31)

where $U$ is not necessarily Gaussian. By using the conditional version of EPI [26], we have

$$
\begin{align*}
&h(X|U, Z) - \gamma h(X + N_3|U, Z) \\
&\leq h(X|U, Z) \\
&\leq \frac{\gamma m}{2} \log \left( \exp \left[ \frac{2}{m} h(X|U, Z) \right] + \exp \left[ \frac{2}{m} h(N_3) \right] \right) \\
&\leq f \left( h(N_3) - \frac{m}{2} \log(\gamma - 1); h(N_3) \right),
\end{align*}
$$

(32)

where we set

$$
f(t; a) := t - \frac{\gamma m}{2} \log \left( \exp \left[ \frac{2}{m} t \right] + \exp \left[ \frac{2}{m} a \right] \right).
$$

Note that the function $f(t; a)$ is concave function of $t$ and takes the maximum at $t = a - \frac{m}{2} \log(\gamma - 1)$ [27]. From Eq. (30), we have

$$
(\gamma - 1)^{-1} \Sigma_{N_3} = \Sigma_{x|u}^*,
$$

which implies

$$
\begin{align*}
&h(N_3) - \frac{m}{2} \log(\gamma - 1) = \frac{1}{2} \log(2\pi e)^m(\gamma - 1)^{-m} |\Sigma_{N_3}| \\
&= \frac{1}{2} \log(2\pi e)^m |\Sigma_{x|u}^* + \Sigma_{N_3}|
\end{align*}
$$

Furthermore, since $\Sigma_{x|u}^*$ and $\Sigma_{N_3}$ are proportional to each other, we have

$$
|\Sigma_{x|u}^*|^{1/m} + |\Sigma_{N_3}|^{1/m} = |\Sigma_{x|u}^* + \Sigma_{N_3}|^{1/m}.
$$

Thus, from Eqs. (31) and (33), we have

$$
R_k - \mu R_p \\
\leq [(1 + \mu)h(\tilde{Y}|Z) - \mu h(X|Z)] - \frac{1 + \mu}{2} \log |K_{gx}K_{gy}^\top| \\
+ \mu \left[ \frac{1}{2} \log(2\pi e)^m|\Sigma_{x|u}^*| - \frac{\gamma}{2} \log(2\pi e)^m|\Sigma_{x|u}^* + \Sigma_{N_3}| \right] \\
= \tilde{I}_k(\Sigma_{x|u}^*) - \mu \tilde{I}_p(\Sigma_{x|u}^*).
$$

Remark 6 One of the difficulties in the above proof is that, after Step 1, we have to show the extremal inequality of the form

$$
\mu h(X|U) + h(X + W_2|U) - (1 + \mu)h(X + \tilde{W}_2|U) \\
\leq \frac{\mu}{2} \log|\Sigma_{x|u}^*| + \frac{1}{2} \log|\Sigma_{x|u}^* + \Sigma_{W_2}| \\
- \frac{1 + \mu}{2} \log|\Sigma_{x|u}^* + \Sigma_{W_2}|.
$$

(34)

This type of extremal inequality has appeared in [28] Corollary 2 (scalar version has appeared in [29] Lemma 1). In [14], the extremal inequality was proved by using a vector generalization of Costa’s entropy power inequality [30]. On the otherhand, we showed Eq. (34) by using the change of variable in Step 2 and by reducing to more tractable form (Eq. 32), which has appeared in the literature [27]. By this reduction, we only need the standard EPI in our proof instead of Costa’s type EPI, and our proof seems more elementary.
V. PROOF OF THEOREM 7

In this section, we show Theorem 7 by using Theorem 2. We follow a similar approach as in [10] Section 4. Since the direct part can be proved by taking a Gaussian auxiliary random variable \( U \) in Proposition 3 (see Section IV-A), we concentrate on the converse part. Without loss of generality, we can assume that the matrices \( B \) and \( E \) are square (but not necessarily invertible). If that is not the case, we can apply singular value decomposition (SVD) to show equivalent sources \((X', Y', Z')\) on \( \mathbb{R}^{m_x} \times \mathbb{R}^{m_y} \times \mathbb{R}^{m_z} \) such that \( \mathcal{R}(X', Y', Z') = \mathcal{R}(X, Y, Z) \) in a similar manner as [22] Section 5-B.

By using SVD, we can write the matrices as
\[
B = U_y A_y V_y, \quad E = U_z A_z V_z,
\]
where \( U_y, V_y, U_z, \) and \( V_z \) are \( m_x \times m_x \) orthogonal matrices, and \( A_y \) and \( A_z \) are diagonal matrices. Let
\[
\tilde{B} = U_y (A_y + \alpha I) V_y, \quad \tilde{E} = U_z (A_z + \alpha I) V_z
\]
for some \( \alpha > 0 \). Then, let
\[
\tilde{Y} = \tilde{B} X + W_y, \quad \tilde{Z} = \tilde{E} X + W_z.
\]

Since \( \tilde{B} \) and \( \tilde{E} \) are invertible, Theorem 2 implies
\[
\mathcal{R}(X, \tilde{Y}, \tilde{Z}) = \mathcal{R}_G(X, \tilde{Y}, \tilde{Z}). \tag{35}
\]
In the following, we will show the following lemma.

**Lemma 7** We have
\[
\mathcal{R}(X, Y, Z) \subset \mathcal{R}(X, \tilde{Y}, \tilde{Z}) + \mathcal{O}(X, \tilde{Y}, \tilde{Z}),
\]
where
\[
\mathcal{O}(X, \tilde{Y}, \tilde{Z}) = \left\{ (0, R_k) : 0 \leq R_k \leq \frac{1}{2} \log |E \Sigma X E^T | + I \right\}.
\]

By letting \( \alpha \to 0 \), \( \mathcal{R}_G(X, \tilde{X}, \tilde{Z}) \) converges to \( \mathcal{R}_G(X, Y, Z) \) and \( \mathcal{O}(X, \tilde{Y}, \tilde{Z}) \) converges to \( \{(0, 0)\} \). Thus, Eq. (35) and Lemma 7 imply \( \mathcal{R}(X, Y, Z) \subset \mathcal{R}_G(X, Y, Z) \).

**Proof of Lemma 7**

Let
\[
C_y = U_y A_y (A_y + \alpha I)^{-1} V_y, \quad C_z = U_z A_z (A_z + \alpha I)^{-1} V_z.
\]
Then, we have \( C_y C_y^T < I \) and \( C_z C_z^T < I \). Thus, we can write
\[
Y = C_y \tilde{Y} + W_y', \quad Z = C_z \tilde{Z} + W_z'
\]
for \( W_y' \sim \mathcal{N}(0, I - C_y C_y^T) \) and \( W_z' \sim \mathcal{N}(0, I - C_z C_z^T) \), i.e., we have
\[
X \leftrightarrow \tilde{Y} \leftrightarrow Y, \quad X \leftrightarrow \tilde{Z} \leftrightarrow Z. \tag{36} \tag{37}
\]

From Proposition 3 for any \((R_p, R_k) \in \mathcal{R}(X, Y, Z)\), there exist \((U, V)\) satisfying
\[
R_p \geq I(U; Y) - I(U; X), \quad R_k \leq I(U; Z | V) - I(U; Y | V),
\]
and \((U, V) \leftrightarrow X \leftrightarrow (Y, Z)\). Let
\[
\tilde{R}_p = I(U; X) - I(U; \tilde{Y}), \quad \tilde{R}_k = I(U; \tilde{Y} | V) - I(U; \tilde{Z} | V).
\]
Then, we have
\[
R_p - \tilde{R}_p \geq I(U; X) - I(U; Y) - I(U; X) + I(U; \tilde{Y}) = I(U; \tilde{Y}) - I(U; Y) \geq 0,
\]
where the second inequality follows from Eq. (36). On the other hand, we have
\[
R_k - \tilde{R}_k \leq I(U; Y | V) - I(U; Z | V) - I(U; \tilde{Y} | V) - I(U; Y | V) \leq I(U; \tilde{Y} | V) - I(U; Y | V) \leq I(U; V) - I(U; V) \leq I(U; V) - I(U; V)
\]
\[
= I(U; X) - I(U; X) - [I(U; \tilde{Z} | U, V) - I(X; Z | U, V)] \leq I(U; \tilde{Z}) - I(U; Z) \leq [I(U; Z | U, V) - I(X; Z)] \leq I(X; Z) - I(X; Z) \leq \frac{1}{2} \log |E \Sigma_x E^T | + I - \frac{1}{2} \log |E \Sigma_x E^T | + I,
\]
where the second, third, and forth inequalities follow from the Markov relations in Eqs. (36) and (37). □

VI. CONCLUSION

In this paper, we investigated the secret key agreement from vector Gaussian sources by rate limited public communication. We characterized the optimal trade-off between the key rate and the public communication rate as a (covariance) matrix optimization problem. Investigating an efficient method to solve the optimization problem is a future research agenda.

**APPENDIX A DERIVATION OF THE KKT CONDITION**

We first rewrite the optimization problem in Eq. (15) as a standard form
\[
\begin{align*}
\text{minimize} & \quad -I_k(\Sigma_x | u) \\
\text{subject to} & \quad I_p(\Sigma_x | u) - R_p^o \leq 0 \quad (38) \\
& \quad 0 \prec \Sigma_x | u \prec \Sigma_x.
\end{align*}
\]

Let \( \Sigma_x^* | u \) be an optimal solution for this problem, which is also an optimal solution of Eq. (15). Then, we have \( \Sigma_x^* | u \succ 0 \) because of the constraint \( I_p(\Sigma_x | u) - R_p^o \leq 0 \). Thus, there exists a positive definite matrix \( L \) satisfying \( L \prec \Sigma_x^* | u \).
Let us consider another optimization problem

\[
\begin{align*}
\text{minimize} & \quad -I_k(\Sigma_{x|u}) \\
\text{subject to} & \quad I_p(\Sigma_{x|u}) - R_p^o \leq 0 \\
& \quad L \preceq \Sigma_{x|u} \preceq \Sigma_x.
\end{align*}
\]

(39)

Obviously, \(\Sigma_{x|u}^*\) is also an optimal solution for the problem in Eq. (39), and the optimal values for Eqs. (38) and (39) are the same. Although the optimization problem in Eq. (39) is not convex, there exist Lagrange multipliers \(M_1 \geq 0, M_2 \geq 0,\) and \(\mu \geq 0\) satisfying

\[
-\left( -\nabla \Sigma_{x|u} I_k(\Sigma_{x|u}) + \mu \nabla \Sigma_{x|u} (I_p(\Sigma_{x|u}) - R_p^o) \right) = M_2 - M_1, \quad (40)
\]

\[
M_1(\Sigma_{x|u} - L) = 0, \quad (41)
\]

\[
M_2(\Sigma_{x} - \Sigma_{x|u}^*) = 0, \quad (42)
\]

\[
\mu(R_p - I_p(\Sigma_{x|u}^*)) = 0. \quad (43)
\]

if the set of constraint qualifications (CQs) shown below are satisfied (see (22) Appendix 4 for the detail). Since \(\Sigma_{x|u}^* \succ L,\) Eq. (41) implies \(M_1 = 0.\) Thus, by noting the relation

\[
I_k(\Sigma_{x|u}) - \mu I_p(\Sigma_{x|u})
\]

\[
= \frac{\mu}{2} \log(2\pi e)^m|\Sigma_{x|u}| + \frac{1}{2} \log(2\pi e)^m|\Sigma_{x|u} + \Sigma W_s|
\]

\[
- \left[ 1 + \frac{\mu}{2} \right] \log(2\pi e)^m|\Sigma_{x|u} + \Sigma W_s|
\]

\[
+ \left[ 1 + \mu \right] h(Y) - h(Z) - \mu h(X),
\]

(44)

and by setting \(M = 2M_2,\) we have the KKT conditions in Eqs. (16)–(18).

The CQs shown in (22) Appendix 4], which is an interpretation of (31) CQ5a of Section 5.4] are the following: There exists a matrix \(A\) satisfying

1) For any \(u \neq 0\) in the null space of \(\Sigma_{x|u}^* - L,\) we have \(u^T A u > 0.\)

2) For any \(v \neq 0\) in the null space of \(\Sigma_{x} - \Sigma_{x|u}^*,\) we have \(v^T A v < 0.\)

3) \(\text{Tr} \left[ \nabla \Sigma_{x|u} (I_p(\Sigma_{x|u}) - R_p^o) A^T \right] > 0.\)

To check whether the above CQs are satisfied, we suggest \(A\) given by

\[
A = \alpha (L - \Sigma_{x|u}^*) + (\Sigma_{x} - \Sigma_{x|u}^*)
\]

for \(\alpha > 0.\) First we check (1). For any \(u \neq 0\) in the null space of \(\Sigma_{x|u}^* - L,\) we have

\[
u^T A u = u^T (\Sigma_{x} - \Sigma_{x|u}^*) u.
\]

Suppose that \(u^T (\Sigma_{x} - \Sigma_{x|u}^*) u = 0.\) Then we have

\[
0 = u^T (\Sigma_{x|u} - L + (\Sigma_{x} - \Sigma_{x|u})) u
\]

\[
= u^T (\Sigma_{x} - L) u,
\]

which is a contradiction because \(\Sigma x \succ L.\) Thus the condition (1) is satisfied.

Next, we check (2). For any \(v \neq 0\) in the null space of \(\Sigma_{x} - \Sigma_{x|u}^*,\) we have

\[
v^T A v = v^T (L - \Sigma_{x|u}^*) v < 0
\]

because \(L \prec \Sigma_{x|u}^*.\)

Finally, we check (3). By noting

\[
\nabla \Sigma_{x|u} (I_p(\Sigma_{x|u}) - R_p^o) A
\]

\[
= \alpha \frac{1}{2} \text{Tr} \left[ \left( (\Sigma_{x|u}^* + \Sigma W_s)^{-1} - (\Sigma_{x|u}^*)^{-1} \right) (L - \Sigma_{x|u}^*) \right]
\]

\[
+ \frac{1}{2} \text{Tr} \left[ \left( (\Sigma_{x|u}^* + \Sigma W_s)^{-1} - (\Sigma_{x|u}^*)^{-1} \right) (\Sigma x - \Sigma_{x|u}^*) \right].
\]

Since \(L - \Sigma_{x|u}^* < 0,\) by taking \(\alpha > 0\) to be sufficiently large, the condition (3) is satisfied.

Remark 8 We need to introduce the optimization problem in Eq. (39) because the arguments in (22) Appendix 4] is guaranteed only under the condition such that the range of the variable \(\Sigma_{x|u}\) is a closed set.

APPENDIX B

PROOF OF EQUATIONS (22) AND (23)

By noting \(M \geq 0,\) we have

\[
(\Sigma_{x|u}^* + \Sigma W_s)^{-1} = (\Sigma_{x|u}^* + \Sigma W_s)^{-1} + M \succeq (\Sigma_{x|u}^* + \Sigma W_s)^{-1}.
\]

Thus we have

\[
\Sigma W_s \preceq \Sigma W_s.
\]

Since \(\Sigma W_s \succ 0,\) by substituting Eq. (21) into Eq. (16), we have

\[
(\Sigma_{x|u}^* + \Sigma W_s)^{-1}
\]

\[
= \frac{\mu}{1 + \mu} (\Sigma_{x|u}^*)^{-1} + \frac{1}{1 + \mu} (\Sigma_{x|u}^* + \Sigma W_s)^{-1}
\]

\[
< (\Sigma_{x|u}^*)^{-1}
\]

when \(\mu > 0.\) Thus, we have

\[
\Sigma W_s > 0.
\]

Note that \(\Sigma W_s = \Sigma W_s < 0\) when \(\mu = 0.\)

From Eq. (45), we have

\[
(\Sigma_{x|u}^* + \Sigma W_s)^{-1} \succeq (\Sigma_{x|u}^* + \Sigma W_s)^{-1},
\]

where the strict inequality holds for \(\mu > 0.\) Thus we have

\[
\Sigma W_s \preceq \Sigma W_s
\]

and especially

\[
\Sigma W_s \prec \Sigma W_s
\]

(46)

for \(\mu > 0.\)
Appendix C
Proofs of Eq. (26)

Eq. (26) can be derived by the following sequence of equalities:

\[ (\Sigma_x + \Sigma_{Y'})/(\Sigma_{x|u} + \Sigma_{W_y})^{-1} = \left( \Sigma_x - \Sigma_{x|u} + (\Sigma_{x|u} + \Sigma_{W_y}) \right) \left( \Sigma_{x|u} + \Sigma_{W_y} \right)^{-1} = (\Sigma_x - \Sigma_{x|u})(\Sigma_{x|u} + \Sigma_{W_y})^{-1} + I \]

where Eq. (47) follows from Eq. (21) and Eq. (48) follows from Eq. (17).

Appendix D
Proof of Eq. (29)

From Eqs. (24), (5) and (46), we can write

\[ Z = X + \tilde{W}_y + W' \]

where \( W' \sim \mathcal{N}(0, \Sigma_{Y' W} - \Sigma_{W_y}) \). Thus, we have

\[ \Sigma_{yz} = \Sigma_{g}, \quad \Sigma_{zx} = \Sigma_{xz} = \Sigma_{x}. \]

Furthermore, we have

\[ \Sigma_{y} < \Sigma_{z}. \]

From the block inversion of the matrix (e.g. see Appendix 5.5) and Eq. (27), we have

\[ K_{\tilde{g}x} = \left[ \begin{array}{c|c} \Sigma_{\tilde{g}} & \Sigma_{x} \\ \hline \Sigma_{x} & S^{-1} \end{array} \right] = (I - \Sigma_{\tilde{g}} S^{-1})S^{-1}, \]

where

\[ S = \Sigma_x - \Sigma_z \Sigma_z^{-1} \Sigma_x \]

is the Schur complement.

From Eq. (51), we have

\[ I - \Sigma_{\tilde{g}} S^{-1} \Sigma_{\tilde{g}}^{-1} = I - \Sigma_{\tilde{g}} S^{-1} \Sigma_{\tilde{g}}^{-1} \neq 0. \]

Thus we have

\[ |I - \Sigma_{\tilde{g}} S^{-1} \Sigma_{\tilde{g}}^{-1}| = |I - \Sigma_{\tilde{g}} S^{-1} \Sigma_{\tilde{g}}^{-1}| \neq 0. \]

By combining Eqs. (52) and (53), we have Eq. (29). □

Acknowledgment

The first author would like to thank Prof. Ryutaroh Matsumoto for valuable discussions and comments. The authors also would like to thank Prof. Jun Chen for informing the literatures [28], [29]. This research is partly supported by Grant-in-Aid for Young Scientists (Start-up): KAKENHI 21860064.
[27] T. Liu and P. Viswanath, “An extremal inequality motivated by multi-
terminal information-theoretic problems,” IEEE Trans. Inform. Theory, vol. 53, no. 5, pp. 1839–1851, May 2007.
[28] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), “A vector generaliza-
tion of costa’s entropy-power inequality with applications,” IEEE Trans.
Inform. Theory, vol. 56, no. 4, pp. 1865–1879, April 2010.
[29] J. Chen, “Rate region of gaussian multiple description coding with indi-
vidual and central distortion constraints,” IEEE Trans. Inform. Theory, vol. 55, no. 9, pp. 3991–4005, September 2009.
[30] M. H. M. Costa, “A new entropy power inequality,” IEEE Trans. Inform.
Theory, vol. 31, no. 6, pp. 751–760, November 1985.
[31] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, Convex Analysis and
Optimization. Athena Scientific, 2003.
[32] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge
University Press, 2004.