

# Joint Source-Channel Secrecy Using Uncoded Schemes: Towards Secure Source Broadcast

Lei Yu, Houqiang Li, *Senior Member, IEEE*, and Weiping Li, *Fellow, IEEE*

## Abstract

This paper investigates a joint source-channel secrecy problem for the Shannon cipher broadcast system. We suppose list secrecy is applied, i.e., a wiretapper is allowed to produce a list of reconstruction sequences and the secrecy is measured by the minimum distortion over the entire list. For discrete communication cases, we propose a permutation-based uncoded scheme, which cascades a random permutation with a symbol-by-symbol mapping. Using this scheme, we derive an inner bound for the admissible region of secret key rate, list rate, wiretapper distortion, and distortions of legitimate users. For the converse part, we easily obtain an outer bound for the admissible region from an existing result. Comparing the outer bound with the inner bound shows that the proposed scheme is optimal under certain conditions. Besides, we extend the proposed scheme to the scalar and vector Gaussian communication scenarios, and characterize the corresponding performance as well. For these two cases, we also propose another uncoded scheme, orthogonal-transform-based scheme, which achieves the same performance as the permutation-based scheme. Interestingly, by introducing the random permutation or the random orthogonal transform into the traditional uncoded scheme, the proposed uncoded schemes, on one hand, provide a certain level of secrecy, and on the other hand, do not lose any performance in terms of the distortions for legitimate users.

## Index Terms

Uncoded scheme, secrecy, permutation, orthogonal transform, Shannon cipher system.

## I. INTRODUCTION

Investigations on joint source-channel coding (JSCC) could trace back to Shannon's pioneering work [1], where a geometric method was developed to design a communication system. For the JSCC of transmitting a Gaussian source over a Gaussian broadcast channel, Goblick observed [2] that when the source and channel bandwidths are matched (i.e., one channel use per source sample), directly sending a scaled version of the source samples on the channel (i.e., linear scheme) is in fact optimal; while for this case the separation scheme that cascades source coding with channel coding indeed suffers a performance loss [3]. For vector Gaussian communication cases, the optimal linear coding was studied in [4]. In general, the schemes that consist of symbol-by-symbol mappings (not limited to the linear one) are named *uncoded schemes*. The optimality of uncoded schemes for the general source-channel pair

L. Yu is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore (e-mail: leiyu@nus.edu.sg). This work was done when he was at University of Science and Technology of China. H. Li and W. Li are with the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China (e-mail: {lihq,wpli}@ustc.edu.cn).

has been investigated in [3], which showed that the Shannon limit can be achieved by uncoded schemes only when the source and channel satisfy a certain probabilistic matching condition. To further improve the performance for mismatched source-channel pairs, the hybrid coding (or hybrid digital-analog coding) has been studied in [5]-[10], which combines the traditional digital coding and symbol-by-symbol mapping together. As for the converse part of JSCC problem, Reznic *et al.* [11] and Tian *et al.* [12] derived some nontrivial converse results for Gaussian source broadcast problem. Besides, Yu *et al.* [9], [10] generalized the achievability and converse results for the Gaussian communication to the general source-channel case.

On information-theoretic security, the Shannon cipher system (the noisy broadcast version depicted in Fig. 1) was first investigated in Shannon's pioneering work [13], where a sender A communicates with a legitimate receiver B secretly by exploiting a shared secret key. For lossy source communication, wiretapper might only want to decrypt a lossy version of the source. Schieler *et al.* [14] studied a distortion-based secrecy measure in the Shannon cipher system around the assumption that the wiretapper has ability to conduct list decoding with fixed list size, and the secrecy is measured by the minimum distortion over the entire list. Yu *et al.* [15] showed that the systems with this secrecy measure are equivalent to those with secrecy measured by a new quantity *lossy-equivocation*, which could be considered as a lossy extension of the traditional equivocation. Hence the list secrecy is closely related to the traditional equivocation as well. Furthermore, Yu *et al.* used this secrecy measure to study the problem of *source-channel secrecy* for the Shannon cipher system, and showed that for the source-channel pair satisfying certain conditions, an uncoded scheme could outperform the separate one.

JSCC improves the robustness of communication or the performance of broadcast, while secrecy coding improves the security of communication by exploiting the secret key and/or the wiretap channel. Therefore, intuitively the robustness and the security could be obtained simultaneously if we combine JSCC and secrecy coding together. This *joint source-channel secrecy* (JSCS) problem has been considered in several works already. Yamamoto in [16] studied the secure lossy transmission over the noisy wiretap channel with secrecy measured by the wiretapper's best reconstruction distortion. However, it is shown in [14] this secrecy measure is cheap and fragile, since only one bit of secret key suffices to achieve the optimality of secrecy, and meanwhile, only one bit of additional information for the wiretapper suffices to decrypt this optimal encryption scheme. A different formulation of the problem was considered in [17], where the authors assumed there is a fixed information leakage to the wiretapper and wish to minimize the distortion at the legitimate receiver, while at the same time providing a graceful distortion degradation when there is an SNR (Signal Noise Ratio) mismatch. They showed that, for a positive leakage, this can be achieved by a hybrid digital-analog coding. This scenario was extended to consider side information at the receiver in [18] or side information at the sender in [19].

Analog encryption (or analog scrambling) technologies, e.g., sign-change based scheme [20], permutation based scheme [20] and bandwidth-keeping scheme [21], can be seen as uncoded JSCS schemes as well, although they are not designed for a specified source-channel pair. Sign-change based scheme improves secrecy by changing the sign of each sample according to the secret key. But owing to at most one bit secret key used per sample, this scheme could not provide higher secrecy even with a higher key rate available. The permutation based scheme improves secrecy by shuffling the positions of samples. Unlike the sign-change based scheme, it supports any arbitrarily high

key rate. Furthermore, Kang and Liu [23] recently applied the permutation operation in a digital encryption scheme, and showed that the permutation is another powerful encryption technique (besides the one-time pad) to achieve the optimality of secrecy.

#### A. Contributions

In this paper, we consider the joint source-channel secrecy problem of secure source broadcast in the bandwidth-matched Shannon cipher system (see Fig. 1). The list secrecy [14] is used to measure secrecy, that is, the wiretapper is allowed to conduct list decoding with fixed list size, and the secrecy is measured by the minimum distortion over the entire list. We study an achievable region of secret key rate, list rate, wiretapper distortion, and distortions of all legitimate users and show optimality under certain conditions. Our contributions are as follows:

- 1) For the discrete source case, we propose a permutation-based uncoded scheme, which cascades a random permutation with a symbol-by-symbol mapping. Our scheme differs from the permutation based scheme proposed in [23] in two main aspects: 1) our scheme, coupling a permutation operation with a traditional *uncoded scheme*, is designed for the *source-channel secrecy problem*, however, the scheme in [23] couples a permutation operation with a *digital scheme*, and is designed for the *source-secrecy coding problem*; 2) in addition to the finite alphabet case, we also extend the scheme to source-channel pairs with countably infinite alphabets and Gaussian source-channel pairs, which require us to use some more powerful techniques, including unified typicality, d-tilted information, and geometric analysis. By analyzing the proposed scheme, we provide an inner bound for the admissible region. For the converse part, we give an outer bound by using our recent result [15]. Comparing the outer bound with the inner bound shows that the proposed scheme is optimal under certain conditions.
- 2) We extend the proposed scheme to scalar and vector bandwidth-matched Gaussian communication scenarios. For these two cases, we also propose another uncoded scheme, orthogonal-transform-based scheme, which achieves the same inner bounds as the one achieved by the permutation-based scheme. Interestingly, by introducing the random permutation or the random orthogonal transform into the traditional uncoded scheme, the proposed uncoded schemes, no matter for the discrete source case or the Gaussian source-channel case, on one hand, provide a certain level of secrecy, and on the other hand, do not lose any performance in terms of the distortions for legitimate users.

Schieler and Cuff [14] studied the list secrecy problem for the *noiseless point-to-point*<sup>1</sup> version of Shannon cipher system, and showed a digital scheme, in which the secret key is used to choose a source codebook to code the source sequence, is optimal. For this problem, a separate coding, cascading source coding and one-time pad, has been proven optimal as well [15]. Yu *et al.* [15] extended this problem to the *noisy* channel case, and showed that the separate strategy (cascading source coding, one-time pad, and channel coding) is suboptimal in general and a single-letter uncoded scheme could outperform the separate scheme. In this paper we extend

<sup>1</sup>Here the word *noiseless* means the wiretap channel is noiseless, and the word *point-to-point* means there is only one legitimate user in the system.

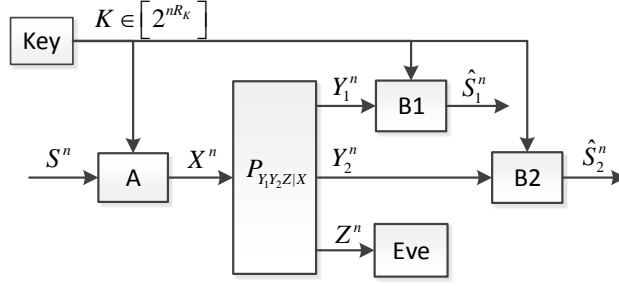


Fig. 1. The Shannon cipher broadcast system.

the problem to *noisy broadcast* scenarios, and propose two kind of uncoded schemes that adopt two different encryption strategies—random permutation and random orthogonal-transform (instead of the traditional one-time pad encryption). We show the proposed uncoded schemes could achieve the optimality under certain cases.

The rest of the paper is organized as follows. Section II formulates the joint source-channel secrecy problem. Section III proposes a permutation-based uncoded scheme for the discrete communication, and analyzed the corresponding performance. Sections IV and V extend the proposed scheme to the scalar and vector Gaussian communications respectively, and another scheme, orthogonal-transform based scheme, is also proposed in these two sections. Finally, Section VI concludes the paper.

## II. PROBLEM FORMULATION

### A. Problem setup

Consider a bandwidth-matched<sup>2</sup> Shannon cipher broadcast system with two legitimate users<sup>3</sup> shown in Fig. 1, where a sender A and two legitimate receivers B1 and B2 share a secret key  $K$  that is uniformly distributed over  $[2^{nR_K}]$ <sup>4</sup> and independent of a source  $S^n$ . The sender A observes the discrete memoryless (DM) source sequence  $S^n$  with each element i.i.d. (independent and identically distributed) according to  $P_S$ , and then transmits it to the legitimate users B1 and B2 over a DM wiretap broadcast channel  $P_{Y_1 Y_2 Z | X}$  confidentially by utilizing the secret key and the wiretap channel. Finally, the legitimate users B1 and B2 produce source reconstructions  $\hat{S}_1^n$  and  $\hat{S}_2^n$ , respectively.

**Definition 1.** An  $(n, R_K)$  block code consists of

- 1) Encoder:  $\varphi : \mathcal{S}^n \times [2^{nR_K}] \mapsto \mathcal{X}^n$ ;

<sup>2</sup>Although here we consider a bandwidth-matched communication system, our results in this section are easy to be extended to any bandwidth-mismatched system since any system with source-channel bandwidth ratio  $\frac{n_s}{n_c}$  can be converted into a bandwidth-matched system, by considering  $n_s$  source symbols and  $n_c$  channel symbols as a source supersymbol and a channel supersymbol, respectively.

<sup>3</sup>Although we only consider the system with two legitimate users, our results derived in this paper can be easily extended to the cases with more legitimate users.

<sup>4</sup>In this paper, the set  $\{1, \dots, m\}$  is sometimes denoted by  $[m]$ .

2) Decoders:  $\psi_i : \mathcal{Y}_i^n \times [2^{nR_K}] \mapsto \hat{\mathcal{S}}_i^n$ ,  $i = 1, 2$ .

The encoder and decoders can be stochastic.

Another output  $Z^n$  of the channel is accessed by a wiretapper Eve. Based on  $Z^n$ , the wiretapper produces a list  $\mathcal{L}(Z^n) \subseteq \check{\mathcal{S}}^n$  and the induced distortion is set to the minimum one over the entire list, i.e.,  $\min_{\check{s}^n \in \mathcal{L}(Z^n)} d_E(S^n, \check{s}^n)$ , where  $d_E(s^n, \check{s}^n) \triangleq \frac{1}{n} \sum_{t=1}^n d_E(s_t, \check{s}_t)$  is a distortion measure for the wiretapper. For given distortion levels  $D_0, D_1, D_2$ , Nodes A and B1, B2 want to communicate the source within distortions  $D_1, D_2$  (for B1 and B2 respectively) by exploiting the secret key and the wiretap channel, while ensuring that the wiretapper's strategy always suffers distortion above  $D_0$  with high probability.

**Definition 2.** The tuple  $(R_K, R_L, D_0, D_1, D_2)$  is achievable if there exists a sequence of  $(n, R_K)$  codes such that  $\forall \epsilon > 0$ ,

1) Distortion constraint:

$$\mathbb{P}\left[d_B(S^n, \hat{S}_i^n) \leq D_i + \epsilon\right] \xrightarrow{n \rightarrow \infty} 1, \quad i = 1, 2; \quad (1)$$

where  $d_B(s^n, \hat{s}^n) \triangleq \frac{1}{n} \sum_{t=1}^n d_B(s_t, \hat{s}_t)$ <sup>5</sup> is a distortion measure for the legitimate users;

2) Secrecy constraint:

$$\limsup_{n \rightarrow \infty} \min_{\substack{\mathcal{L}_n(z^n): \\ \frac{1}{n} \log |\mathcal{L}_n| \leq R_L - \epsilon}} \mathbb{P}\left[\min_{\check{s}^n \in \mathcal{L}(Z^n)} d_E(S^n, \check{s}^n) \geq D_0 - \epsilon\right] \xrightarrow{n \rightarrow \infty} 1. \quad (2)$$

**Definition 3.** The admissible region  $\mathcal{R} \triangleq \{\text{Achievable } (R_K, R_L, D_0, D_1, D_2)\}$ .

We assume that the wiretapper knows the  $(n, R_K)$  code and the distributions  $P_S$  and  $P_{Y_1 Y_2 Z|X}$ .

### B. Henchman problem

The problem above is equivalent to the henchman problem [14], in which wiretapper reconstructs a single sequence with the help of a rate-limited henchman who can access to the source  $S^n$  and the wiretapper's observation  $Z^n$ . As depicted in Fig. 2, the wiretapper receives the best possible  $nR_n$  bits from the henchman to assist in producing a reconstruction sequence  $\check{S}^n$ .

**Definition 4.** The  $R_n$  henchman code (Hcode) of a  $(n, R_K)$  block code consists of

- 1) Encoder:  $\varphi_H : \mathcal{S}^n \times \mathcal{Z}^n \mapsto [2^{nR_n}]$ ;
- 2) Decoder:  $\psi_H : [2^{nR_n}] \times \mathcal{Z}^n \mapsto \check{\mathcal{S}}^n$ .

We assume that the wiretapper and henchman are aware of the  $(n, R_K)$  block code adopted by Nodes A and B, and they cooperate to design a henchman code based on the  $(n, R_K)$  block code.

**Definition 5.** The tuple  $(R_K, R_L, D_0, D_1, D_2)$  is achievable in the henchman problem if there exists a sequence of  $(n, R_K)$  codes such that  $\forall \epsilon > 0$ ,

<sup>5</sup>For simplicity, we only consider the legitimate users have the same distortion measure. Note that our results derived in this paper still hold for the case with different distortion measures.

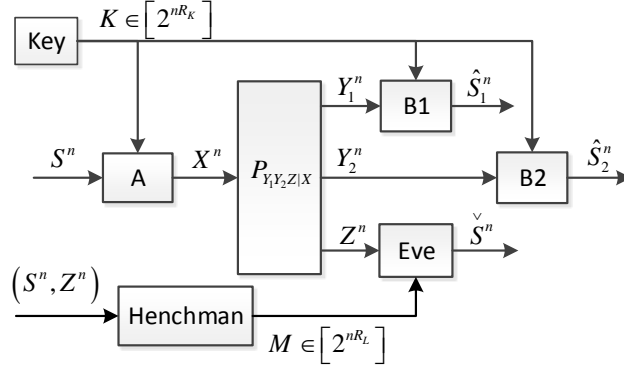


Fig. 2. Henschman problem.

- 1) Distortion constraint: (1);
- 2) Secrecy constraint:

$$\limsup_{n \rightarrow \infty} \min_{\substack{R_n \text{ Hcodes:} \\ R_n \leq R_L - \epsilon}} \mathbb{P} \left[ d_E(S^n, \check{S}^n) \geq D_0 - \epsilon \right] \xrightarrow{n \rightarrow \infty} 1. \quad (3)$$

The equivalence between the list secrecy problem and the henschman problem, shown in the following proposition, has been proven by Schieler and Cuff [14, Prop. 1].

**Proposition 1.** [14] *The tuple  $(R_K, R_L, D_0, D_1, D_2)$  is achievable in the list reconstruction problem if and only if it is achievable in the henschman problem.*

Furthermore, the list secrecy problem and henschman problem are also equivalent to a lossy-equivocation secrecy problem; see [15].

In addition to the DM system, we also consider the Shannon cipher system with a Gaussian source  $S \sim \mathcal{N}(0, \lambda)$  transmitted over a power-constrained Gaussian wiretap broadcast channel

$$Y_i = X + V_i, i = 1, 2, \quad (4)$$

$$Z = X + V_0, \quad (5)$$

where  $V_i, i = 0, 1, 2$  are zero-mean additive Gaussian noises with variances  $N_i, i = 0, 1, 2$ , independent of  $X$ . For this case, the constraint on channel input power

$$\mathbb{P}[\rho(X^n) \leq P + \epsilon] \xrightarrow{n \rightarrow \infty} 1, \forall \epsilon > 0, \quad (6)$$

should be added to Definitions 2 and 5, where  $\rho(x^n) = \frac{1}{n} \sum_{i=1}^n x_i^2$ . For the system involving the channel power constraint, Proposition 1 still holds.

### III. DISCRETE COMMUNICATION

#### A. Permutation based Scheme (Finite Alphabets)

In this section, we propose a secure uncoded scheme by coupling the permutation operation with the traditional uncoded JSCC scheme. The uncoded scheme for JSCC system (with two receivers) consists of three symbol-by-

symbol mappings:  $x(s), \hat{s}_1(y_1), \hat{s}_2(y_2)$ . The induced distortions are  $D_i = \mathbb{E}d_B(S, \hat{s}_i(Y_i)), i = 1, 2$ . It is easy to show that we can benefit from replacing the encoder  $x(s)$  with a stochastic one  $P_{X|S}$  when secrecy is considered for the system. On the other hand, observe that  $d_B(s^n, \hat{s}_i^n) = \frac{1}{n} \sum_{t=1}^n d_B(s_t, \hat{s}_{i,t}) = \mathbb{E}_{T_{s^n, \hat{s}_i^n}} d_B(s, \hat{s}_i)$ , where  $T_{s^n, \hat{s}_i^n}$  denotes the joint type (empirical distribution) of  $(s^n, \hat{s}_i^n)$ . That is, the induced distortions only depend on the joint type of source and reconstruction sequences. Therefore, if we want to improve the secrecy of a scheme and at the same time retain the induced distortions unchanged, we only need to require the encryption and decryption operations does not change the joint type of source and reconstruction sequences. That is, for the encryption  $s'^n(s^n, k)$  and the decryption  $\hat{s}_i^n(\hat{s}_i'^n, k)$ , we require  $T_{s'^n, \hat{s}_i'^n} = T_{s^n, \hat{s}_i^n}$ . To that end, here we consider a random permutation as the encryption operation, and the inverse permutation as the decryption operation. Obviously, the permutation and its inverse operation do not change the joint type of the source sequence and its reconstructions.

*Codebook (Public Key) Generation:* Generate a permutation set  $\mathcal{C} = \{\Psi_k, k \in [2^{nR_k}]\}$  with each element uniformly at random and independently selected from the set of permutations of  $[n]$  (which is denoted as  $\mathfrak{S}_n$ ). As a public key, the codebook  $\mathcal{C}$  is revealed to the sender and all the receivers (including the wiretapper).

*Encoding:* Upon observing a source sequence  $s^n$  and a key  $k$ , the encoder first generates

$$s'^n = \Psi_k(s^n), \quad (7)$$

and then generates  $x^n$  according to  $\prod_{t=1}^n P_{X|S}(x_t|s'_t)$ . Here for a permutation sequence  $\Psi_k = \sigma^n$ ,  $\Psi_k(s^n) \triangleq (s_{\sigma_1}, s_{\sigma_2}, \dots, s_{\sigma_n})$  denotes the permutation operation on  $s^n$  (more precisely, on the indices of  $s^n$ ) respect to the permutation sequence  $\Psi_k$ .<sup>6</sup>

*Decoding (for Legitimate Users):* For the legitimate user  $B_i$ ,  $i = 1, 2$ , upon the received sequence  $y_i^n$  and the key  $k$ , the decoder first reconstructs  $s'^n$  as

$$\hat{s}'_{i,t} = \hat{s}_i(y_{i,t}), t \in [n], \quad (8)$$

by using the symbol-by-symbol mapping  $\hat{s}_i(\cdot)$ , and then produces

$$\hat{s}_i^n = \Psi_k^{-1}(\hat{s}_i'^n), \quad (9)$$

by using the inverse permutation operation  $\Psi_k^{-1}(\cdot)$  of  $\Psi_k(\cdot)$ .

The proposed scheme above cascades a random permutation operation with the traditional uncoded JSCC scheme. The uncoded JSCC part provide a graceful degradation of the source for legitimate users with different channel qualities. The random permutation operation part that shifts the sequence in the same type provides a certain level of secrecy. Next we will analyze the asymptotic performance of the proposed scheme as blocklength  $n$  goes to infinity. At first, we need introduce some basic properties of the random codebook  $\mathcal{C}$ .

Observe that for any permutation sequence  $\Psi$ , the mapping between  $\Psi(\cdot)$  and  $\Psi^{-1}(\cdot)$  is bijective, hence we have the following lemma.

<sup>6</sup>In this paper, the permutation sequence is termed as *permutation sequence*, and to distinguish from it, the permutation mapping from one sequence to another sequence is termed as *permutation operation*. When there is no disambiguation, we call both of them *permutation*.

**Lemma 1.** Suppose  $\Psi$  is a permutation sequence uniformly at random selected from  $\mathfrak{S}_n$ , the set of permutations of  $[n]$ . Then  $\Psi^{-1}$  is also uniformly distributed on  $\mathfrak{S}_n$ , and moreover for any permutation sequence  $\psi \in \mathfrak{S}_n$ , both  $\Psi(\psi)$  and  $\Psi^{-1}(\psi)$  also have the uniform distribution on  $\mathfrak{S}_n$ .

Utilizing Lemma 1, we can establish the following lemma.

**Lemma 2.** Suppose  $\Psi$  is a permutation sequence uniformly at random selected from  $\mathfrak{S}_n$ . Then  $\Psi(s^n)$  transforms an arbitrary sequence  $s^n \in \mathcal{S}^n$  into a random sequence that is uniformly distributed on the set of sequences of type  $T_{s^n}$  (the type class of  $T_{s^n}$ ). Moreover, for finite  $\mathcal{S}$ , the set of sequences of type  $T_{s^n}$  has cardinality  $2^{n(H(T_{s^n})-o(1))}$ , and hence

$$\mathbb{P}[\Psi(s^n) = s'^n] = 2^{-n(H(T_{s^n})-o(1))} 1\{T_{s'^n} = T_{s^n}\}, \quad s'^n \in \mathcal{S}^n, \quad (10)$$

where  $o(1)$  denotes a term tending zero as  $n \rightarrow \infty$ .

*Proof:*

$$\mathbb{P}[\Psi(s^n) = s'^n] = \sum_{\psi \in \mathfrak{S}_n: \psi(s^n) = s'^n} \mathbb{P}(\Psi = \psi) \quad (11)$$

$$= \sum_{\psi \in \mathfrak{S}_n: \psi(s^n) = s'^n} \frac{1}{n!} \quad (12)$$

$$= \frac{\prod_{s \in \mathcal{S}} (nT_{s^n}(s))!}{n!} 1\{T_{s'^n} = T_{s^n}\} \quad (13)$$

$$= \frac{1\{T_{s^n} = T_{s'^n}\}}{|\{s'^n \in \mathcal{S}^n : T_{s'^n} = T_{s^n}\}|}, \quad (14)$$

where (14) follows from  $|\{s'^n \in \mathcal{S}^n : T_{s'^n} = T_{s^n}\}| = \frac{n!}{\prod_{s \in \mathcal{S}} (nT_{s^n}(s))!}$ . This implies  $\Psi(s^n)$  transforms an arbitrary sequence  $s^n \in \mathcal{S}^n$  into a random sequence uniformly distributed on the set of sequences of type  $T_{s^n}$ .

From the type counting lemma [29, Lem. 2.3], we have that for finite  $\mathcal{S}$ ,

$$(n+1)^{-|\mathcal{S}|} 2^{nH(T_{s^n})} \leq |\{s'^n \in \mathcal{S}^n : T_{s'^n} = T_{s^n}\}| \leq 2^{nH(T_{s^n})}. \quad (15)$$

Hence  $|\{s'^n \in \mathcal{S}^n : T_{s'^n} = T_{s^n}\}| = 2^{n(H(T_{s^n})-o(1))}$ . Combining it with (14) gives us

$$\mathbb{P}[\Psi(s^n) = s'^n] = 2^{-n(H(T_{s^n})-o(1))} 1\{T_{s'^n} = T_{s^n}\}. \quad (16)$$

■

Lemma 2 shows a nice property of the random permutation operation: The resulting sequence will be uniformly distributed on the set of sequences of type  $T_{s^n}$  for the input sequence  $s^n$ , if the permutation is randomly and uniformly chosen from the set of permutations of  $[n]$ . Utilizing this property, we can characterize the performance of the proposed scheme, as shown in the following theorem. The proof of this theorem is given in Appendix A.



**Theorem 1** (Permutation based Scheme for Finite Alphabets). *For DM communication with finite alphabets  $(S, \check{S}, \mathcal{X}, \mathcal{Z}, \mathcal{Y}_i, \hat{S}_i, i = 1, 2$  are all finite), the permutation based scheme above achieves the region  $\mathcal{R}^{(i)} \subseteq \mathcal{R}$ , where*

$$\mathcal{R}^{(i)} \triangleq \bigcup_{P_{X|S}} \left\{ \begin{array}{l} (R_K, R_L, D_0, D_1, D_2) : \\ D_i \geq \min_{\hat{s}_i} \mathbb{E} d_B(S, \hat{s}_i(Y_i)), i = 1, 2, \\ R_L \leq \min \{R_K + R_{S|Z}(D_0), R_S(D_0)\} \end{array} \right\}, \quad (17)$$

with

$$(S, Y_1, Y_2, Z) \sim \sum_x P_S P_{X|S} P_{Y_1 Y_2 Z|X}, \quad (18)$$

$$R_S(D) = \min_{P_{\check{S}|S} : \mathbb{E} d_E(S, \check{S}) \leq D} I(S; \check{S}) \quad (19)$$

denoting the rate-distortion function of  $S$ , and

$$R_{S|Z}(D) = \min_{P_{\check{S}|SZ} : \mathbb{E} d_E(S, \check{S}) \leq D} I(S; \check{S}|Z) \quad (20)$$

denoting the conditional rate-distortion function of  $S$  given two-sided information  $Z$ .

Note that for the  $\mathcal{R}^{(i)}$  above, the components  $(D_1, D_2)$  and the components  $(R_K, R_L, D_0)$  depend on each other through  $P_{X|S}$ . Observe that for a given  $P_{X|S}$ ,  $\min_{\hat{s}_i} \mathbb{E} d_B(S, \hat{s}_i(Y_i)), i = 1, 2$  are the minimal distortions that the legitimate users can achieve even for the non-secrecy communication case. On the other hand,  $\min \{R_K + R_{S|Z}(D_0), R_S(D_0)\}$  is larger than  $R_{S|Z}(D_0)$ , the optimal  $R_L$  can be achieved by uncoded schemes when there is no key. Hence compared with traditional uncoded schemes, the proposed scheme, on one hand, improves the performance of secrecy to a certain extent, and on the other hand, does not lose any performance in terms of the distortions of legitimate users.

The first constraint of  $\mathcal{R}^{(i)}$  is consistent with the performance of traditional uncoded schemes. The second constraint of  $\mathcal{R}^{(i)}$ , roughly speaking, follows from the following argument. On one hand, the henchman and the wiretapper can ignore the signal  $Z^n$  altogether and use a  $R_S(D_0)$ -rate point-to-point source code to describe  $S^n$  within distortion  $D_0$ . On the other hand, the proposed scheme forces the wiretapper's optimal strategy to be an indirect guessing strategy: First, the wiretapper decrypts the secret key by using  $R_K$  rate; then upon the observation  $Z^n$ , the wiretapper reconstructs the sequence  $S'^n$  within distortion  $D_0$  by using rate  $R_{S|Z}(D_0)$  (denote the reconstruction as  $\check{S}'^n$ ); finally, upon the secret key and  $\check{S}'^n$ , the wiretapper reconstructs the source  $S^n$  as  $\check{S}^n = \Psi_k^{-1}(\check{S}'^n)$ . Obviously the distortion between  $S^n$  and  $\check{S}^n$  is the same as that between  $S'^n$  and  $\check{S}'^n$ , since the average distortion only depends the joint type of the sequences. Hence the wiretapper needs rate  $R_K + R_{S|Z}(D_0)$  to achieve the distortion  $D_0$ .

Now we consider a special case: sending a binary source over a binary wiretap broadcast channel. For the binary communication, the source is a Bernoulli source  $S \sim \text{Bern}(\frac{1}{2})$  with the Hamming distortion measure  $d_B(s, \hat{s}) = d_E(s, \hat{s}) \triangleq 0$ , if  $s = \hat{s}$ ; 1, otherwise. The binary wiretap broadcast channel is  $Y_i = X \oplus V_i, i = 1, 2, Z = X \oplus V_0$  with  $V_i \sim \text{Bern}(p_i), V_0 \sim \text{Bern}(p_0), 0 \leq p_0, p_1, p_2 \leq \frac{1}{2}$ . Set  $X = S \oplus E$  with  $E \sim \text{Bern}(p')$ . Then from Theorem 1, we get the following corollary.

**Corollary 1** (Binary Communication). *For the binary communication above, we have  $\mathcal{R}^{(i)} \subseteq \mathcal{R}$ , where*

$$\mathcal{R}^{(i)} \triangleq \bigcup_{0 \leq p' \leq \frac{1}{2}} \left\{ \begin{array}{l} (R_K, R_L, D_0, D_1, D_2) : \\ D_i \geq p' \star p_i, i = 1, 2, \\ R_L \leq \min \left\{ R_K + [H_2(p' \star p_0) - H_2(D_0)]^+, [1 - H_2(D_0)]^+ \right\} \end{array} \right\},$$

with  $[x]^+ \triangleq \max\{0, x\}$ ,  $\star$  denoting the binary convolution, i.e.,

$$x \star y = (1 - x)y + x(1 - y), \quad (21)$$

and  $H_2$  denoting the binary entropy function, i.e.,

$$H_2(p) = -p \log p - (1 - p) \log(1 - p). \quad (22)$$

### B. Permutation based Scheme (More General Alphabets)

Theorem 1 can be extended to more general alphabets cases, as shown in the following theorem. The proof of this theorem is given in Appendix C.

**Theorem 2** (Permutation based Scheme for More General Alphabets). *Assume  $\mathcal{S}$  is countable,  $\check{\mathcal{S}}$  is finite, and  $\mathcal{X}, \mathcal{Z}, \mathcal{Y}_i, \hat{\mathcal{S}}_i, i = 1, 2$  are general<sup>7</sup>. Assume  $H(S)$  is finite, and  $P_S$  satisfies*

$$N_{P_S} \left( \frac{1}{n} \right) = o \left( \frac{n}{\log n} \right), \quad (23)$$

$$\Phi_{P_S} \left( \frac{1}{n} \right) = o \left( \frac{1}{\log n} \right), \quad (24)$$

$$\tilde{N}_{P_S} \left( \frac{\delta}{\log n} \right) = o \left( \frac{n}{\log^2 n} \right), \forall 0 < \delta \leq 1, \quad (25)$$

where  $N_{P_S}(\alpha) \triangleq |\{s : P_S(s) \geq \alpha\}|$  denotes the number of probability values that are not smaller than  $\alpha$ ,

$\Phi_{P_S}(\alpha) \triangleq \sum_{s: P_S(s) < \alpha} P_S(s)$  denotes the sum of probability values that are smaller than  $\alpha$ ,  $\tilde{N}_{P_S}(\beta) \triangleq \min_{\alpha: \Phi_{P_S}(\alpha) \leq \beta} N_{P_S}(\alpha)$

denotes the minimum number  $N$  such that the sum of the probability values except  $N$  largest ones is not larger than  $\beta$ . Then Theorem 1 still holds.

**Remark 1.** The conditions (23)-(25) is equivalent to as  $x \downarrow 0$ ,<sup>8</sup>

$$N_{P_S}(x) = o \left( \frac{1}{x \log \frac{1}{x}} \right), \quad (26)$$

$$\Phi_{P_S}(x) = o \left( \frac{1}{\log \frac{1}{x}} \right), \quad (27)$$

$$\tilde{N}_{P_S}(x) = o \left( x^2 e^{\frac{\delta}{x}} \right), \forall 0 < \delta \leq 1. \quad (28)$$

**Remark 2.** The conditions (23)-(25) require that the sequence  $P_S(s), s \in \mathcal{S}$  should vanish as fast as possible. Obviously, (23)-(25) hold for any finite  $\mathcal{S}$ . Besides, for any countably infinite  $\mathcal{S}$ , it is easy to verify that any distribution  $P_S$  such that  $P_S(s) = o(s^{-\alpha}), s = 1, 2, \dots$ <sup>9</sup> for some  $\alpha > 1$  satisfies (23)-(25) as well. However, if

<sup>7</sup>An alphabet is countable means that it is either finite or countably infinite. An alphabet is general means that it is either countable or uncountable (e.g., continuous).

<sup>8</sup>This claim holds when we ignore  $n$  is an integer in (23)-(25).

<sup>9</sup>Without loss of generality, any countably infinite  $\mathcal{S}$  can be converted into  $\{1, 2, 3, \dots\}$  by some bijective mapping.

$P_S(s)$  converges slower than or as slow as  $\frac{1}{s}$ , then  $\sum_{s \geq 1} P_S(s)$  does not converge, and hence  $P_S$  cannot be a probability distribution. This implies Theorem 2 holds for almost all probability distributions.

Note that for a countably infinite alphabet  $\mathcal{S}$ , we need the conditions (23)-(25) to guarantee the existence of a high-probability set (unified typicality set), for each sequence of which, Lemma 2 still holds. This further makes Theorem 2 hold, just as done for the finite alphabets case.

### C. Outer Bound

For the system with a single legitimate user (remove the legitimate user B2 from the system in Fig. 1), the following outer bound for the admissible region of  $(R_K, R_L, D_0, D_1)$  has been proven by us recently [15].

**Lemma 3.** [15] *For the DM communication with only one legitimate user,*

$$\mathcal{R} \subseteq \mathcal{R}^{(o)} \triangleq \bigcup_{P_{\hat{S}_1|S}} \left\{ \begin{array}{l} (R_K, R_L, D_0, D_1) : C_1 \geq I(S; \hat{S}_1), \\ D_1 \geq \mathbb{E}d_B(S, \hat{S}_1), \\ R_L \leq \min \left\{ R_K + \Gamma \left( I(S; \hat{S}_1), P_{Y_1|X}, P_{Z|X} \right) + R_{S|\hat{S}_1}(D_0), \right. \\ \left. R_S(D_0) \right\} \end{array} \right\},$$

where  $C_1$  denotes the channel capacity for the legitimate user, and

$$\Gamma(R, P_{Y|X}, P_{Z|X}) \triangleq \min_{\substack{Q_{Y|Z|X}: Q_{Y|X}=P_{Y|X}, \\ Q_{Z|X}=P_{Z|X}}} \max_{Q_X: I_Q(X; Y) \geq R} I_Q(X; Y|Z) \quad (29)$$

with  $I_Q(\cdot)$  denoting the mutual information under distribution  $Q_X Q_{Y|Z|X}$ , is a function specified by the wiretap channel.

The first two constraints of  $\mathcal{R}^{(o)}$  follow from the source-channel coding theorem [30], and the last constraint follows from an indirect decryption strategy for the wiretapper: Roughly speaking, the wiretapper first reconstructs  $\hat{S}_1^n$  using rate  $\Gamma(I(S; \hat{S}_1), P_{Y_1|X}, P_{Z|X})$ , next decrypts the secret key using rate  $R_K$ , then upon  $Y_1^n$  and secret key, produces the legitimate user's reconstruction  $\hat{S}_1^n$ , and finally upon  $\hat{S}_1^n$  produces a final reconstruction  $\tilde{S}^n$  using rate  $R_{S|\hat{S}_1}(D_0)$ . The details can be seen in [15].

By applying this lemma to the system with two legitimate users (the system considered in this paper), the following outer bound is immediate.

**Theorem 3 (Outer Bound).** *For the DM communication (with two legitimate users),*

$$\mathcal{R} \subseteq \mathcal{R}^{(o)} \triangleq \bigcup_{P_{\hat{S}_1 \hat{S}_2|S}} \left\{ \begin{array}{l} (R_K, R_L, D_0, D_1, D_2) : C_i \geq I(S; \hat{S}_i), \\ D_i \geq \mathbb{E}d_B(S, \hat{S}_i), i = 1, 2, \\ R_L \leq \min \{R_1, R_2, R_S(D_0)\} \end{array} \right\},$$

where  $C_i$  denotes the channel capacity for the legitimate user  $i$ , and

$$R_i = R_K + \Gamma \left( I(S; \hat{S}_i), P_{Y_i|X}, P_{Z|X} \right) + R_{S|\hat{S}_i}(D_0), i = 1, 2. \quad (30)$$

When specialized to the binary communication, we have the following corollary.

**Corollary 2** (Binary Communication). *For binary communication,*

$$\mathcal{R} \subseteq \mathcal{R}^{(o)} \triangleq \left\{ (R_K, R_L, D_0, D_1, D_2) : \begin{array}{l} D_i \geq p_i, i = 1, 2, \\ R_L \leq \min \{ R_1, R_2, [1 - H_2(D_0)]^+ \} \end{array} \right\}.$$

where

$$R_i = R_K + [H_2(p_0) - H_2(p_i)]^+ + [H_2(D_i) - H_2(D_0)]^+, i = 1, 2. \quad (31)$$

Comparing Theorem 1 and Corollary 2, we can identify the optimality of the proposed scheme for binary communication.

**Theorem 4** (Optimality of the Proposed Scheme). *For the binary communication (with 2 legitimate users), the proposed uncoded scheme is optimal if  $p_0 \leq p_i \leq D_i \leq D_0$  or  $p_0 \geq p_i = D_i \geq D_0$  holds for  $i = 1$  or  $2$ .*

*Remark 3.* Theorem 4 implies under conditions that compared with one of legitimate users, the wiretapper has a better channel and wants to produce a worse reconstruction, or the legitimate user's distortion is restricted to be the Shannon limit and meanwhile the wiretapper has a worse channel and wants to produce a better reconstruction, the proposed uncoded scheme is optimal. It is worth noting that these optimality conditions do not include the practical scenario in which the wiretapper has a worse channel than the legitimate users and a higher distortion requirement. But it does not mean our scheme is not optimal for the practical scenario. We believe that for the binary broadcast communication without secrecy requirement, the proposed uncoded scheme with  $p' = 0$  and with no permutation operation is the *unique* scheme to achieve the Shannon limits for both the legitimate users. If so, when the secrecy requirement is involved, the proposed scheme is optimal as well, no matter what the wiretapper's channel condition is and what his desired distortion level is. This is because  $R_K$  rate of secret key could increase  $R_L$  at most by  $R_K$ , and our scheme satisfies this point. Of course, we need a rigorous proof about this claim, but unfortunately, now we have no idea how to prove it.

We know that when there is no secrecy constraint, the traditional uncoded scheme could outperform the separate scheme for broadcast communication scenarios. It is not surprising that when secrecy constraint is involved, the proposed uncoded scheme still could outperform the separate scheme. However, surprisingly, the example given in [15] shows the proposed uncoded scheme may strictly outperform the separate coding even for the secure *point-to-point* communication (with only one legitimate user).

#### IV. SCALAR GAUSSIAN COMMUNICATION

In this section, we consider a Gaussian source  $S \sim \mathcal{N}(0, \lambda)$  transmitted over a bandwidth-matched<sup>10</sup> and power-constrained Gaussian wiretap broadcast channel (the average input power is constrained by  $P$ ). The distortion measures are set to  $d_B(s, \hat{s}) = d_E(s, \hat{s}) = d(s, \hat{s}) \triangleq (s - \hat{s})^2$ .

<sup>10</sup>Although we can also convert a bandwidth-mismatched Gaussian system into a bandwidth-matched system, just as done in Remark 2, our results in this section cannot be easily extended to the bandwidth-mismatched system since the linear coding used in our schemes is specified for the bandwidth-matched one.

For this communication system, we provide two uncoded schemes. The first one is just the scheme proposed in previous section. Next we will show that the permutation based scheme also works in the Gaussian communication case. The other one is an orthogonal-transform based scheme, which cascades a random orthogonal transform (instead of random permutation operation) with a symbol-by-symbol mapping.

#### A. Permutation based Scheme

It has been shown that linear coding is optimal for the bandwidth-matched Gaussian broadcast communication when there is no secrecy requirement [2]. Hence we set  $P_{X|S}$  and  $\hat{s}_i(y_i), i = 1, 2$  to the linear functions  $x = \alpha s$ ,  $\hat{s}_i = \beta_i y_i, i = 1, 2$  in the proposed scheme for DM communications, where  $\alpha = \sqrt{\frac{P'}{\lambda}}$  with  $0 \leq P' \leq P$  and  $\beta_i = \frac{\sqrt{\lambda P'}}{P' + N_i}$ . Then we apply this permutation based scheme to the Gaussian communication. The performance of this scheme is provided in the following theorem, the proof of which is given in Appendix E.

**Theorem 5** (Permutation based Scheme). *For the Gaussian communication, the proposed permutation based scheme achieves the region  $\mathcal{R}^{(i)} \subseteq \mathcal{R}$ , where*

$$\mathcal{R}^{(i)} \triangleq \bigcup_{0 \leq P' \leq P} \left\{ \begin{array}{l} (R_K, R_L, P, D_0, D_1, D_2) : \\ D_i \geq \frac{\lambda N_i}{P' + N_i}, i = 1, 2, \\ R_L \leq \min \left\{ R_K + \frac{1}{2} \log^+ \left( \frac{\lambda N_0}{D_0(P' + N_0)} \right), \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right) \right\} \end{array} \right\}, \quad (32)$$

with  $\log^+ x \triangleq \max \{0, \log x\}$ .

*Remark 4.* The  $\mathcal{R}^{(i)}$  here is just the one given in Theorem 1 with  $P_{X|S}$  and  $\hat{s}_i(y_i), i = 1, 2$  set to  $x = \alpha s$  and  $\hat{s}_i = \beta_i y_i, i = 1, 2$ , respectively. This is because they are achieved by the same scheme.

*Remark 5.* The first constraint of  $\mathcal{R}^{(i)}$  is consistent with the performance of linear coding [2]. The second constraint of  $\mathcal{R}^{(i)}$  follows from the similar argument to the DM case.

Note that for  $\mathcal{R}^{(i)}$ ,  $P'$  is a variable. Moreover, the region of  $(D_1, D_2)$  and the region of  $(R_K, R_L, D_0)$  depend on each other through  $P'$  which satisfies  $0 \leq P' \leq P$ . This finding is similar to the discrete communication case. Given  $(R_K, D_0)$ , the minimum of achievable  $D_1$  (or  $D_2$ ) and the maximum of achievable  $R_L$  are both decreasing in  $P'$ , which implies for the proposed scheme, transmitting the source using a larger power results in smaller distortions for legitimate users, but also leads to decrypting the source more easily for the wiretapper. The proposed scheme with  $P' = P$ , on one hand, provides a certain level of secrecy, and on the other hand, it achieves the Shannon's distortion limits for both legitimate users. The region in Theorem 5 with  $\lambda = 1$  and  $P' = 1$  is illustrated in Fig. 3. Given  $P'$ ,  $(D_1, D_2)$  has no effect on the  $(R_K, R_L, D_0)$  tradeoff.

#### B. Orthogonal-Transform based Scheme

The proposed scheme above uses a random permutation operation (which shuffles the sequence within the same type class) to improve the level of secrecy. It works not only for the discrete communication but also for the continuous communication, such as the Gaussian communication. In this subsection we propose another secure uncoded scheme for the Gaussian communication case which is designed from a geometric point of view.

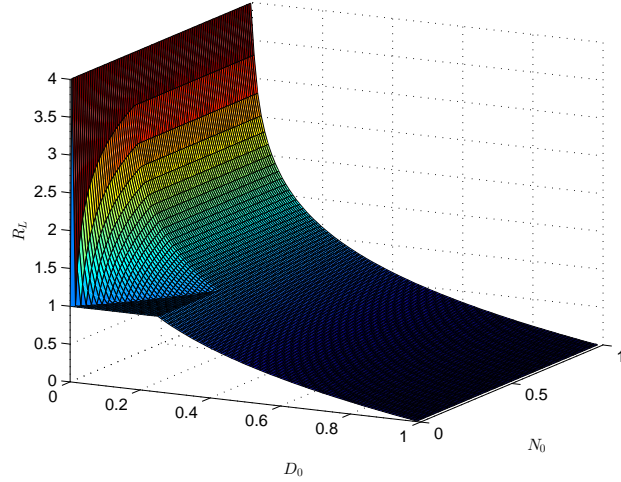


Fig. 3. The region in Theorem 5 with  $\lambda = 1$  and  $P' = 1$ .

To give an interpretation for the motivation of our proposed scheme, we consider a special case where the wiretapper has a noiseless channel. Apply linear coding to the Gaussian communication, then we know that given the Euclidean norm, the sequence of a Gaussian source uniformly distributed on some sphere, and so are the sequences of channel input, outputs, and source reconstructions. Assume we generate a set of bijective transforms (as a codebook), and randomly choose one of them (according to the key) to transform the source sequence before applying linear coding on it. To keep the power unchanged, these transforms are required to map a sphere into itself. On the other hand, by using the secret key the legitimate users could transform it back. Hence the induced distortions of legitimate users do not change as well. Furthermore, without knowing the secret key but with knowing the norm of the source sequence and the codebook, in the view of the wiretapper, the source sequence is uniformly distributed over the vectors that are possible to generate the channel output (wiretapper's observation) through some key values. To make the wiretapper guess the source as difficultly as possible, these vectors should be “uniformly” (at equal distance) located on the sphere. This is because if so, the wiretapper has to cover either all these vectors or the whole sphere to meet the decryption requirement. It can be shown the orthogonal transform is one of such transforms. Hence it is adopted in our second scheme.

*Codebook (Public Key) Generation:* Generate random  $n \times n$  matrices  $Q_k, k \in [2^{nR_k}]$  independently whose elements are generated i.i.d. according to  $\mathcal{N}(0, 1)$ . Then apply Gram-Schmidt orthonormalization process to the columns of each matrix, hence all the resulting matrices are orthogonal and they constitute a subset of orthogonal matrices  $\mathcal{C} = \{\Psi_k, k \in [2^{nR_k}]\}$ . As a public key, the codebook  $\mathcal{C}$  is revealed to the sender and all the receivers (including the wiretapper).

*Encoding:* Upon observing a source sequence  $s^n$  and a key  $k$ , the encoder generates  $x^n$  as follows.

$$x^n = \alpha \Psi_k s^n, \quad (33)$$

where  $\alpha = \sqrt{\frac{P'}{\lambda}}$  with  $0 \leq P' \leq P$ .

*Decoding (for Legitimate Users):* For legitimate user  $B_i$ ,  $i = 1, 2$ , upon the received sequence  $y_i^n$  and the key  $k$ , the decoder reconstructs the source as follows.

$$\hat{s}_i^n = \beta_i \Psi_k^T y_i^n, \quad (34)$$

where  $\beta_i = \frac{\sqrt{\lambda P'}}{P' + N_i}$ , and  $\Psi_k^T$  denotes the transpose of the matrix  $\Psi_k$ .

Next we will analyze the asymptotic performance of this scheme. Similar to the case of permutation based scheme, we need first introduce some basic properties of the random codebook  $\mathcal{C}$ .

**Lemma 4.** [31] Suppose  $Q$  is a random  $n \times n$  matrix with each element independently distributed according to Gaussian distribution  $\mathcal{N}(0, 1)$ . Let  $Q_1, Q_2, \dots, Q_n$  be the columns of  $Q$  and let  $\Psi$  be the random matrix whose columns are obtained by applying the Gram-Schmidt orthonormalization procedure to  $Q_1, Q_2, \dots, Q_n$ . Then both  $\Psi$  and  $\Psi^T$  have the uniform distribution (Haar measure under orthogonal transform) on the set of  $n \times n$  orthogonal matrices  $\mathcal{F}(n)$ , and moreover for any orthogonal matrix  $A$ , both  $A\Psi$  and  $\Psi A$  also have the uniform distribution on  $\mathcal{F}(n)$ .

Utilizing Lemma 4, we can establish the following lemma.

**Lemma 5.** Random orthogonal transform  $x^n = \Psi s^n$  with  $\Psi$  uniformly distributed on orthogonal matrices set  $\mathcal{F}(n)$ , transforms an arbitrary vector  $s^n \in \mathbb{R}^n$  into a random vector that is uniformly distributed on the  $(n-1)$ -sphere with radius  $\|s^n\|$ .

*Proof:* From Lemma 4, without loss of generality we can assume  $\Psi$  is obtained in the manner described in Lemma 4. Let  $\Psi_1, \Psi_2, \dots, \Psi_n$  be the columns of  $\Psi$ . From Gram-Schmidt orthonormalization, we know that  $\Psi_1 = \frac{Q_1}{\|Q_1\|}$ , and for any rotation matrix (or more generally, orthogonal matrix)  $A$ ,  $A\Psi_1 = \frac{AQ_1}{\|Q_1\|} = \frac{AQ_1}{\|AQ_1\|}$ . On the other hand,  $Q_1$  is a random vector with each element i.i.d.  $\sim \mathcal{N}(0, 1)$ , and it is easy to verify that for any rotation matrix  $A$ ,  $AQ_1$  has the same distribution as  $Q_1$ , i.e., a normally distributed random vector is invariant to rotation. Therefore,  $A\Psi_1$  has the same distribution as  $\Psi_1$ , i.e.,  $\Psi_1$  is also invariant to rotation. This implies  $\Psi_1$  is uniformly distributed on the unit  $(n-1)$ -sphere. In addition, observe  $\Psi(1, 0, \dots, 0)^T = \Psi_1$ . Hence the random matrix  $\Psi$  transforms vector  $(1, 0, \dots, 0)^T$  to a random vector uniformly distributed on the  $(n-1)$ -sphere. For arbitrary vector  $s^n \in \mathbb{R}^n$ , we can easily find an orthogonal matrix  $B$  with the first column  $\frac{s^n}{\|s^n\|}$ . Hence  $s^n$  can be expressed as  $s^n = \|s^n\| B(1, 0, \dots, 0)^T$ . Then we have  $\Psi s^n = \|s^n\| \Psi B(1, 0, \dots, 0)^T$ . From Lemma 4,  $\Psi B$  has the same distribution as  $\Psi$ . Hence  $\Psi B(1, 0, \dots, 0)^T$  is also uniformly distributed on the unit  $(n-1)$ -sphere, which implies  $\Psi s^n$  is uniformly distributed on the  $(n-1)$ -sphere with radius  $\|s^n\|$ . ■

Lemma 5 implies the resulting vector will be uniformly distributed on the sphere where the input vector is, if the transform matrix is randomly and uniformly chosen from the set of orthogonal matrices. This is a nice property of the random orthogonal transform, similar to the property of the random permutation operation. Utilizing the properties, we can establish the following theorem, the proof of which is given in Appendix F.

**Theorem 6** (Orthogonal-Transform based Scheme). *For the Gaussian communication, the inner bound  $\mathcal{R}^{(i)}$  given in Theorem 5 can be achieved by the scheme above as well.*

The inner bound  $\mathcal{R}^{(i)}$  can be understood from a geometric point of view. The random orthogonal transform in the proposed scheme guarantees that given  $Z^n, S^n$  has a uniform distribution on  $2^{nR_K}$  small  $(n-2)$ -spheres with radius  $r_2 = \sqrt{\frac{n\lambda N_0}{P' + N_0}}$  whose centers are uniformly distributed on the  $(n-1)$ -sphere with center  $O$  (the origin) and radius  $r_1 = \sqrt{\frac{n\lambda P'}{P' + N_0}}$ . However, owing to the uniform conditional distribution of the source given  $Z^n$  and the lack of secret key, the wiretapper needs at least  $2^{nR_K} (\frac{r_2}{\sqrt{nD_0}})^n$  balls with radius  $\sqrt{nD_0}$  to cover these  $(n-2)$ -spheres. On the other hand, under the unconditional case, the source has a uniform distribution on the  $(n-1)$ -sphere with center  $O$  and radius  $r_0 = \sqrt{n\lambda}$ . Hence if ignoring  $Z^n$ , the wiretapper needs at least  $(\frac{r_0}{\sqrt{nD_0}})^n$  balls with radius  $\sqrt{nD_0}$  to cover the sphere. This results in the inner bound  $\mathcal{R}^{(i)}$ .

It seems somewhat counterintuitive that the permutation based scheme achieves the same performance as the orthogonal-transform based scheme, as shown by Theorems 5 and 6; it is easy to observe that for low-dimension cases, e.g., 2-dimension case (see Fig. 4), permutations cannot always transform a source sequence into vectors “uniformly” (at equal distance) distributed over a sphere, so why does this property hold (with high probability) when the dimension goes to infinity? Actually, it indeed does. This is because as the dimension increases, such “bad”<sup>11</sup> source sequences will occur with vanishing probability. This can be seen from that<sup>12</sup>  $\mathbb{P}([S]^n \in \mathcal{U}_\delta^n([S])) \rightarrow 1$  as  $n \rightarrow \infty$  (i.e., besides on the sphere, the source sequence should also with high probability appear the neighborhoods of the vectors in  $\mathcal{U}_\delta^n([S])$ ), and moreover,  $\mathcal{U}_\delta^n([S])$  consists of a set of “good” source sequences. Hence the “good” source sequences will occur with high probability as the dimension increases, that is, permutations will transform an arbitrary source sequence from a high probability set into vectors “uniformly” distributed over a sphere.

### C. Comparison with Sign-Change Based Scheme

In previous two subsections, we give an analysis of the asymptotic performance of permutation based scheme or orthogonal-transform based scheme. However, is it necessary to let the blocklength  $n$  go to infinity? What if  $n$  is set to be a finite value? In this subsection, we study the simplest finite blocklength case:  $n = 1$  (single-letter codes). For this case, the permutation based scheme is obviously inferior to the asymptotic case, since for 1 dimension case no permutation exists except for the source sequence itself. Hence in the following, we mainly consider the orthogonal-transform based scheme.

For  $n = 1$ , the orthogonal-transform based scheme reduces to a sign-change based scheme [20], [15]. Next we compare the proposed schemes with this sign-change based scheme [20], [15]. Assume  $R_K = 1$ , and the secret key is uniformly distributed on  $\{0, 1\}$ .

*Encoding:* Upon observing a source sequence  $s$  and a key  $k$ , the encoder generates  $x$  as follows.

$$x = \alpha \Psi_k s, \quad (35)$$

<sup>11</sup>Here a source sequence is said to be “good” if its permutations are “uniformly” distributed over a sphere; otherwise it is “bad”. Obviously, the permutations of a “good” source sequence are also “good”.

<sup>12</sup>Here  $[S] = \Delta \cdot \text{Round}\left(\frac{S}{\Delta}\right)$  and  $\mathcal{U}_\delta^n([S])$  is the  $\delta$ -unified typical set for  $P_{[S]}$ ; see the proof in E.



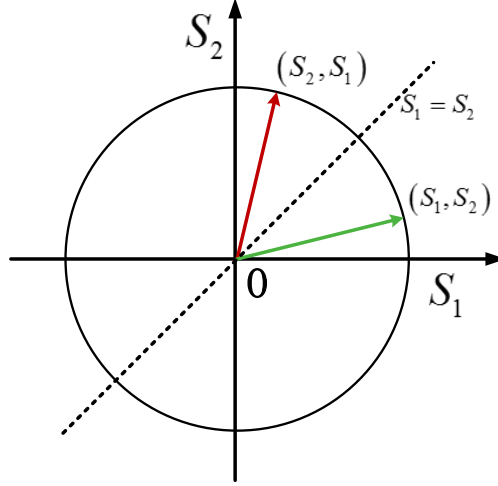


Fig. 4. Illustration of permutations of a source sequence for  $n = 2$  case.

where  $\alpha = \sqrt{\frac{P'}{\lambda}}$  with  $0 \leq P' \leq P$ , and

$$\Psi_k = \begin{cases} -1, & \text{if } k = 0; \\ 1, & \text{if } k = 1. \end{cases} \quad (36)$$

*Decoding (for Legitimate Receivers):* For legitimate receiver  $B_i$ ,  $i = 1, 2$ , upon the received sequence  $y_i$  and the key  $k$ , the decoder reconstructs the source as follows.

$$\hat{s}_i = \beta_i \Psi_k y_i, \quad (37)$$

where  $\beta_i = \frac{\sqrt{\lambda P'}}{P' + N_i}$ .

It is easy to verify that  $(S_t, K_t, X_t, Y_{1,t}, Y_{2,t}, Z_t, \hat{S}_{1,t}, \hat{S}_{2,t})_{t=1}^{\infty}$  are i.i.d. and

$$f_{S,Z}(s, z) = f_S(s) \cdot \frac{1}{2} [f_{V_0}(z - \alpha s) + f_{V_0}(z + \alpha s)] \quad (38)$$

$$= f_Z(z) \cdot \frac{1}{2} [f_{V'_0}(s - \beta_0 z) + f_{V'_0}(s + \beta_0 z)], \quad (39)$$

where  $\beta_0 = \frac{\sqrt{\lambda P'}}{P' + N_0}$ ,  $f_{V_0}$  denotes the probability distribution function (pdf) of the wiretapper's channel noise  $V_0$ , and  $f_{V'_0}$  denotes the pdf of  $V'_0 \sim \mathcal{N}(0, \frac{\lambda N_0}{P' + N_0})$ . Given  $Z$ ,  $S$  can be regarded as a Gaussian mixture with two components of equal weight and variance. For such single-letter scheme, in [15] we have shown the maximum achievable  $R_L$  (or equivalently the minimum rate needed to code  $S$  within distortion  $D_0$  with two-sided information  $Z$ ) equals the conditional rate-distortion function  $R_{S|Z}(D_0)$ . The performance of the sign-change based scheme is given by the following theorem.

**Theorem 7** (Sign-Change based Scheme). [15] *For the Gaussian communication with  $R_K = 1$ , the sign-change based scheme above achieves the region  $\mathcal{R}_{\text{sign}}^{(i)} \subseteq \mathcal{R}$ , where*

$$\mathcal{R}_{\text{sign}}^{(i)} \triangleq \bigcup_{0 \leq P' \leq P} \left\{ (R_K, R_L, P, D_0, D_1, D_2) : \begin{array}{l} D_i \geq \frac{\lambda N_i}{P' + N_i}, i = 1, 2, \\ R_L \leq R_{S|Z}(D_0) \end{array} \right\},$$

with  $R_{S|Z}(D_0)$  denoting the conditional rate-distortion function of  $S$  given two-sided information  $Z$ , defined in (20).

Since it is hard (even if possible) to express  $R_{S|Z}(D_0)$  in closed form, for ease of comparison, we will derive a closed-form upper bound for  $R_{S|Z}(D_0)$ . The result is shown in the following lemma, and the proof is given in Appendix I.

**Lemma 6.** *If  $(S, Z)$  follows the distribution (38) or (39), then*

$$R_{S|Z}(D_0) \leq \min \left\{ R_{S|Z}^{(\text{UB})}(D_0), \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right) \right\}, \quad (40)$$

where

$$R_{S|Z}^{(\text{UB})}(D_0) \triangleq \begin{cases} \frac{(\lambda - D_0)(P' + N_0)}{\lambda P'}, & \text{if } \frac{\lambda N_0}{P' + N_0} < D_0 \leq \lambda; \\ 1 + \frac{1}{2} \log \left( \frac{\lambda N_0}{D_0(P' + N_0)} \right), & \text{if } 0 \leq D_0 \leq \frac{\lambda N_0}{P' + N_0}. \end{cases} \quad (41)$$

Since  $R_{S|Z}(D_0)$  denotes the minimum rate needed to code  $S$  within distortion  $D_0$  when  $Z$  is available at both encoder and decoder, we can give an interpretation for the upper bound from the perspective of source coding. First, by ignoring the side information, we have  $R_{S|Z}(D_0) \leq \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right)$ , where  $\frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right)$  is the minimum rate needed to code  $S$  without any side information. Second, if  $\frac{\lambda N_0}{P' + N_0} \leq D_0 \leq \lambda$ , then consider the following timesharing coding strategy.<sup>13</sup> If we code the secret key  $K$  (1 bit per symbol), then using a linear decoder (similar to those of legitimate users), we can reconstruct the source within distortion  $\frac{\lambda N_0}{P' + N_0}$ . On the other hand, if we do not code anything, then it results in rate 0 and distortion  $\lambda$ . By using a timesharing strategy between these two schemes, we need  $\frac{(\lambda - D_0)(P' + N_0)}{\lambda P'}$  rate to reconstruct the source within distortion  $D_0$ . Finally, if  $0 \leq D_0 \leq \frac{\lambda N_0}{P' + N_0}$ , then we reconstruct the source within distortion  $\frac{\lambda N_0}{P' + N_0}$  by using rate 1 to code the secret key, and upon the reconstruction, we further code the residual error within distortion  $D_0$  by using rate  $\frac{1}{2} \log \left( \frac{\lambda N_0}{D_0(P' + N_0)} \right)$ .

Combining Theorem 7 and Lemma 6 gives us the following result.

**Theorem 8** (Outer Bound of  $\mathcal{R}_{\text{sign}}^{(i)}$ ). *For the Gaussian communication with  $R_K = 1$ , the region achieved by the sign-change based scheme satisfies  $\mathcal{R}_{\text{sign}}^{(i)} \subseteq \mathcal{R}_{\text{sign}}^{(o)}$ , where*

$$\mathcal{R}_{\text{sign}}^{(o)} \triangleq \bigcup_{0 \leq P' \leq P} \left\{ (R_K, R_L, P, D_0, D_1, D_2) : \begin{aligned} & D_i \geq \frac{\lambda N_i}{P' + N_i}, i = 1, 2, \\ & R_L \leq \min \left\{ R_{S|Z}^{(\text{UB})}(D_0), \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right) \right\} \end{aligned} \right\}. \quad (42)$$

*Remark 6.* Observe that only 1 bit/symbol of key can be exploited by the sign-change based scheme even when  $R_K > 1$ . Hence for that case, its performance is still that given by Theorem 7 and outer bounded by 42.

From Lemma 6, it can be observed that when  $R_K = 1$ ,  $R_{S|Z}^{(\text{UB})}(D_0) = 1 + \frac{1}{2} \log^+ \left( \frac{\lambda N_0}{D_0(P' + N_0)} \right)$  for  $0 \leq D_0 \leq \frac{\lambda N_0}{P' + N_0}$ , and  $R_{S|Z}^{(\text{UB})}(D_0) < 1 = 1 + \frac{1}{2} \log^+ \left( \frac{\lambda N_0}{D_0(P' + N_0)} \right)$  for  $\frac{\lambda N_0}{P' + N_0} < D_0 \leq \lambda$ . Hence  $\mathcal{R}_{\text{sign}}^{(o)} \subsetneq \mathcal{R}^{(i)}$ , where  $\mathcal{R}^{(i)}$

<sup>13</sup>Note that the argument here is only available for the inequality (40), and does not apply to the secrecy problem considered in this paper. For the secrecy problem the wiretapper and henchman cannot benefit from adopting a timesharing strategy since the constraint (2) or (3) is to restrict the excess-distortion probability, instead of the average distortion.

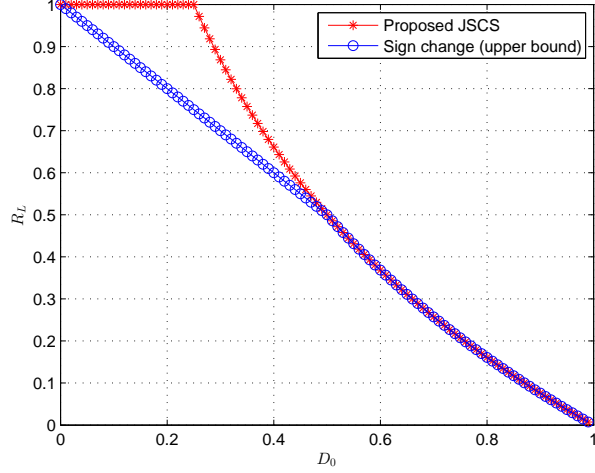


Fig. 5. Comparison of the achievable  $R_L$  by the proposed (infinite blocklength) schemes and that by the sign-change based scheme.  $\lambda = 1$ ,  $N_0 = 0$  (noiseless wiretap channel) and  $R_K = 1$ .

given in Theorem 5 denotes the achievable region by the permutation based scheme or orthogonal-transform based scheme. This implies for the same  $P'$ , the sign-change based scheme is strictly inferior to the proposed schemes under the condition  $\frac{\lambda N_0}{P' + N_0} < D_0 \leq \lambda$ . That is, the single-letter version of orthogonal-transform based scheme is inferior to the corresponding infinite blocklength version. To see it clearer, the  $R_L$  achieved by the proposed (infinite blocklength) schemes (given in Theorem 5) and the upper bound of  $R_L$  achieved by the sign-change based scheme (given in Theorem 8) are illustrated in Fig. 5.

#### D. Outer Bound

For the Gaussian communication, the following outer bound has been proven for the system with only one legitimate user [15].

**Lemma 7.** [15] *For the Gaussian communication with only one legitimate user,*

$$\mathcal{R} \subseteq \mathcal{R}^{(\circ)} \triangleq \left\{ (R_K, R_L, P, D_0, D_1) : \begin{array}{l} D_1 \geq \frac{\lambda N_1}{P + N_1}, \\ R_L \leq \min \left\{ R_1, \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right) \right\} \end{array} \right\}, \quad (43)$$

where

$$R_1 = R_K + \frac{1}{2} \log^+ \left( \frac{1 + P/N_1}{1 + P/N_0} \right) + \frac{1}{2} \log^+ \left( \frac{D_1}{D_0} \right). \quad (44)$$

Using this result, we have the following outer bound for the system with 2 legitimate users (the system considered in this paper).

**Theorem 9** (Outer Bound). *For the Gaussian communication (with 2 legitimate users),*

$$\mathcal{R} \subseteq \mathcal{R}^{(o)} \triangleq \left\{ (R_K, R_L, P, D_0, D_1, D_2) : \begin{array}{l} D_i \geq \frac{\lambda N_i}{P+N_i}, i = 1, 2, \\ R_L \leq \min \left\{ R_1, R_2, \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right) \right\} \end{array} \right\}, \quad (45)$$

where

$$R_i = R_K + \frac{1}{2} \log^+ \left( \frac{1 + P/N_i}{1 + P/N_0} \right) + \frac{1}{2} \log^+ \left( \frac{D_i}{D_0} \right), i = 1, 2. \quad (46)$$

Comparing Theorem 6 and Corollary 9, we can identify the optimality of the proposed schemes for the Gaussian communication. This result is similar to Theorem 4 for the binary communication.

**Theorem 10** (Optimality of the Proposed Schemes). *For the Gaussian communication (with 2 legitimate users), the proposed scheme is optimal if  $N_0 \leq N_i, D_0 \geq D_i$  or  $N_0 \geq N_i, D_0 \leq D_i = \frac{\lambda N_i}{P+N_i}$  holds for  $i = 1$  or  $2$ .*

A similar remark to Remark 3 applies to this theorem.

## V. VECTOR GAUSSIAN COMMUNICATION

The proposed schemes are easily extended to vector Gaussian communication scenarios. Consider an  $m$ -vector Gaussian source  $\mathbf{S} \sim \mathcal{N}(\mathbf{0}, \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m))$ <sup>14</sup> transmitted over an  $m$ -vector Gaussian broadcast channel

$$\mathbf{Y}_i = \mathbf{X} + \mathbf{V}_i, i = 1, 2, \quad (47)$$

where  $\mathbf{Y}_i$  is the channel output vector observed by the  $i$ -th legitimate user, and  $\mathbf{V}_i \sim \mathcal{N}(\mathbf{0}, \text{diag}(N_{i,1}, N_{i,2}, \dots, N_{i,m}))$  is an additive Gaussian noise vector. A wiretapper Eve accesses to another channel output  $\mathbf{Z}$  through a channel

$$\mathbf{Z} = \mathbf{X} + \mathbf{V}_0, \quad (48)$$

where  $\mathbf{V}_0 \sim \mathcal{N}(\mathbf{0}, \text{diag}(N_{0,1}, N_{0,2}, \dots, N_{0,m}))$  is an additive Gaussian noise vector as well. The distortion measures are set to  $d_B(\mathbf{s}, \hat{\mathbf{s}}) = d_E(\mathbf{s}, \hat{\mathbf{s}}) = \sum_{j=1}^m (s_j - \hat{s}_j)^2$ , and the channel cost function is set to  $\rho(\mathbf{x}) = \sum_{j=1}^m x_j^2$ .

Consider the vectors  $\mathbf{S}, \mathbf{X}, \mathbf{Y}_i, \mathbf{Z}, \hat{\mathbf{S}}_i, \check{\mathbf{S}}_i$  as super-symbols, then the proposed permutation based scheme can be applied to the vector Gaussian case directly. The performance of this scheme can be proven by following similar steps to the proof for the scalar Gaussian case.

Furthermore, we can apply the proposed orthogonal-transform based scheme to each subsource-subchannel pair, as shown in the following.

**Codebook (Public Key) Generation:** Generate  $m \cdot 2^{nR_K}$  random  $n \times n$  matrices  $Q_{j,k}, j \in [m], k \in [2^{nR_K}]$  independently whose elements are generated i.i.d. according to  $\mathcal{N}(0, 1)$ . Then we apply Gram-Schmidt orthonormalization process on every matrix, hence all the resulting matrices are orthogonal, and constitute a subset of orthogonal matrices  $\mathcal{C} = \{\Psi_{j,k}, j \in [m], k \in [2^{nR_K}]\}$ . As a public key, the codebook  $\mathcal{C}$  is revealed to the sender and all the receivers (including the wiretapper).

<sup>14</sup>In this paper, we use bold font to denote vector or matrix, e.g.,  $(S_1, \dots, S_m)$  and  $(s_1, \dots, s_m)$  are denoted by  $\mathbf{S}$  and  $\mathbf{s}$ , respectively.

*Encoding:* Upon observing a source sequence  $\mathbf{s}^n = (s_1^n, s_2^n, \dots, s_m^n)$  and a key  $k$ , the encoder generates  $\mathbf{x}^n = (x_1^n, x_2^n, \dots, x_m^n)$  as follows.

$$x_j^n = \alpha_j \Psi_{j,k} s_j^n, j \in [m], \quad (49)$$

where  $\alpha_j = \sqrt{\frac{P_j}{\lambda_j}}$  with transmitting power  $P_1, P_2, \dots, P_m$  such that  $0 \leq \sum_{j=1}^m P_j \leq P$ .

*Decoding (for Legitimate Users):* For the legitimate user Bi,  $i = 1, 2$ , upon the received sequence  $\mathbf{y}_i^n$  and the key  $k$ , the decoder reconstructs the source as follows.

$$\hat{s}_{i,j}^n = \beta_{i,j} \Psi_{j,k}^T y_{i,j}^n, j \in [m], \quad (50)$$

where  $\beta_{i,j} = \frac{\sqrt{\lambda_j P_j}}{P_j + N_i}$ .

The achievable regions by the proposed schemes (permutation based scheme and orthogonal-transform based scheme) are given in the following theorem, the proof of which is given in Appendix J.

**Theorem 11** (Performance of the Proposed Schemes). *For the vector Gaussian communication, the permutation based scheme or the orthogonal-transform based scheme above achieves the same region  $\mathcal{R}^{(i)} \subseteq \mathcal{R}$ , where*

$$\mathcal{R}^{(i)} \triangleq \bigcup_{\substack{P_1, P_2, \dots, P_m \geq 0, \\ 0 \leq \sum_{j=1}^m P_j \leq P}} \left\{ \begin{array}{l} (R_K, R_L, P, D_0, D_1, D_2) : \\ D_i \geq \sum_{j=1}^m \frac{\lambda_j N_{i,j}}{P_j + N_{i,j}}, i = 1, 2, \\ R_L \leq \min \{R_K + R_{S|Z}(D_0), R_S(D_0)\} \end{array} \right\},$$

with

$$R_S(D_0) = \sum_{j=1}^m \frac{1}{2} \log^+ \left( \frac{\lambda_j}{\mu} \right) \quad (51)$$

$$R_{S|Z}(D_0) = \sum_{j=1}^m \frac{1}{2} \log^+ \left( \frac{\lambda_j N_{0,j}}{\theta (P_j + N_{0,j})} \right) \quad (52)$$

and with  $\mu$  and  $\theta$  such that

$$D_0 = \sum_{j=1}^m \min \{ \mu, \lambda_j \}, \quad (53)$$

$$D_0 = \sum_{j=1}^m \min \left\{ \theta, \frac{\lambda_j N_{0,j}}{P_j + N_{0,j}} \right\}. \quad (54)$$

*Remark 7.* Actually, in Theorem 11,  $R_S(D_0)$  denotes the rate-distortion function of the source  $\mathbf{S}$ , and  $R_{S|Z}(D_0)$  denotes the rate-distortion function of the source  $\mathbf{S}$  with the side information  $\mathbf{Z}$  available at both the encoder and decoder, where  $Z_j = \sqrt{\frac{P_j}{\lambda_j}} S_j + V_j, j \in [m]$  with  $\mathbf{V} \sim \mathcal{N}(0, \text{diag}(N_{0,1}, N_{0,2}, \dots, N_{0,m}))$  independent of  $\mathbf{S}$ .

## VI. CONCLUDING REMARKS

In this paper, we studied the joint source-channel secrecy problem for secure source broadcast in the Shannon cipher system, in which the list secrecy is used to measure the secrecy of communication. We proposed two secure uncoded schemes: a permutation based scheme for discrete, scalar Gaussian, and vector Gaussian communications, and an orthogonal-transform based scheme for the latter two communications. In these two uncoded schemes, a

random permutation or a random orthogonal transform is cascaded with the traditional uncoded JSCC scheme. The analysis showed that the proposed schemes outperform the sign-change based scheme. Interestingly, by adding the random permutation operation or the random orthogonal transform into the traditional uncoded scheme, the proposed uncoded schemes, on one hand, provide a certain level of secrecy, and on the other hand, do not lose any performance in terms of the distortions for legitimate users.

Although the proposed schemes adopt two different random transforms, permutation operation and orthogonal transform, they are consistent in two aspects: First, actually the permutation operation is one kind of orthogonal transform; second, for the Gaussian communication, the orthogonal transform can be also considered as a shift operation that shifts a sequence to another in the same “type”, if we treat the Euclidean norm of the source sequence as its “type”<sup>15</sup>. Furthermore, it is worth noting that different from the common construction of codebook in information theory (including *spherical codes* such as the one used in [27]), the codebooks in the proposed schemes are constructed by generating a sequence of i.i.d. random permutations or random matrices, instead of a sequence of i.i.d. random samples. In other words, the codebooks used here specify a sequence of bijective operations or transforms and hence they apply to uncoded schemes; while the common codebooks in information theory only specify a sequence of samples and hence can only be used in quantization operation (or digital schemes). Furthermore, such random-permutation or random-matrix based codebook construction can be also found in [22], [23], [24], [25], [26], where they were used to design digital schemes for communication, secrecy communication, and antijamming communication problems. But different from those works, in our case they were used to design uncoded schemes, instead of digital schemes.

It is worth noting that the proofs used in this paper follow basic outline of the proofs in [14]. But different from [14], besides the finite alphabet case, we also considered the countably infinite alphabet and continuous (Gaussian) alphabet cases. Hence some powerful techniques, including unified typicality, d-tilted information, geometric analysis, and discretization, are used in our proofs. Furthermore, the unified typicality used in our proofs is different from the existing one defined in [34]. The unified typical set defined by us has a good property that the sequences in it only have (nearly) sub-exponential number of types. This property coincides with the finite alphabet case, and is of crucial importance to our proofs. We believe our definition of unified typicality could be used to further extend the method of types to countably infinite alphabet cases (besides the extension in [34]).

## APPENDIX A

### PROOF OF THEOREM 1

Denote

$$Z'^n \triangleq \Psi_K^{-1}(Z^n), \quad (55)$$

$$X'^n \triangleq \Psi_K^{-1}(X^n), \quad (56)$$

$$Y_i'^n \triangleq \Psi_K^{-1}(Y_i^n). \quad (57)$$

<sup>15</sup>This kind of type can be called “weak type”, since the relationship of it and the weak typicality is similar to that of the traditional type (empirical distribution) and strong typicality.

Then from the fact that the permutation operation is bijective, we have that

$$P_{CS^n KS'^n X^n Y_i^n Z^n X'^n Y_i'^n Z'^n \hat{S}_i'^n \hat{S}_i^n} = P_K P_C P_{S^n} P_{S'^n|S^n \Psi_K} P_{X^n|S'^n \Psi_K} P_{Y_i^n Z^n|X^n \Psi_K} P_{\hat{S}_i'^n|Y_i^n \Psi_K} P_{\hat{S}_i^n|\hat{S}_i'^n} P_{X'^n|X^n \Psi_K} P_{Y_i'^n|Y_i^n \Psi_K} P_{Z'^n|Z^n \Psi_K} \quad (58)$$

$$= P_K P_C P_{S^n} P_{S'^n|S^n \Psi_K} P_{X'^n|S'^n \Psi_K} P_{Y_i'^n Z'^n|X'^n \Psi_K} P_{\hat{S}_i'^n|Y_i'^n \Psi_K} P_{\hat{S}_i^n|\hat{S}_i'^n} P_{X^n|X'^n \Psi_K} P_{Y_i^n|Y_i'^n \Psi_K} P_{Z^n|Z'^n \Psi_K} \quad (59)$$

$$= P_K P_C P_{S^n} P_{X'^n|S^n} P_{Y_i'^n Z'^n|X'^n} P_{\hat{S}_i'^n|Y_i'^n} P_{S'^n|S^n \Psi_K} P_{\hat{S}_i^n|Y_i'^n \Psi_K} P_{X^n|X'^n \Psi_K} P_{Y_i^n|Y_i'^n \Psi_K} P_{Z^n|Z'^n \Psi_K}, \quad (60)$$

and similarly,  $P_{CS^n KS'^n X^n Y_i^n Z^n X'^n Y_i'^n Z'^n \hat{S}_i'^n \hat{S}_i^n}$  can be also expressed as

$$P_{CS^n KS'^n X^n Y_i^n Z^n X'^n Y_i'^n Z'^n \hat{S}_i'^n \hat{S}_i^n} = P_K P_C P_{S'^n} P_{X^n|S'^n} P_{Y_i^n Z^n|X^n} P_{\hat{S}_i'^n|Y_i^n} P_{S^n|S'^n \Psi_K} P_{\hat{S}_i^n|\hat{S}_i'^n} P_{X'^n|X^n \Psi_K} P_{Y_i'^n|Y_i^n \Psi_K} P_{Z'^n|Z^n \Psi_K}. \quad (61)$$

Hence  $(\Psi_K, Z^n) \rightarrow Z'^n \rightarrow S^n$  forms a Markov chain. Furthermore, since the permutation operation does not change the joint distribution of the sequences, we have  $P_{S^n} P_{X'^n|S^n} P_{Y_i'^n Z'^n|X'^n} P_{\hat{S}_i'^n|Y_i'^n} = P_{S'^n} P_{X^n|S'^n} P_{Y_i^n Z^n|X^n} P_{\hat{S}_i^n|Y_i^n} = \prod [P_S P_{X|S} P_{Y_i Z|X} P_{\hat{S}_i|Y_i}]$ , where  $P_{\hat{S}_i|Y_i}(\hat{s}|y) \triangleq 1\{\hat{s} = \hat{s}_i(y)\}$  denotes the conditional distribution induced by the decoder  $i$ , and  $P_S P_{X|S} P_{Y_i Z|X} P_{\hat{S}_i|Y_i}$  is the distribution given in (18).

Since  $(S^n, \hat{S}_i^n)$  is an i.i.d. sequence, by the law of large numbers,

$$\mathbb{P}[d_B(S^n, \hat{S}_i^n) \leq \mathbb{E}d_B(S, \hat{S}_i) + \epsilon] \xrightarrow{n \rightarrow \infty} 1, \quad (62)$$

for any  $\epsilon > 0$ . Hence the distortion constraints for legitimate users are satisfied.

Next we prove the secrecy constraint is also satisfied, i.e., if

$$\limsup_{n \rightarrow \infty} R_n < \min\{R_K + R_{S|Z}(D_0), R_S(D_0)\}, \quad (63)$$

then  $\lim_{n \rightarrow \infty} \mathbb{E}_{CZ^n} [\max_{R_n \text{ Hcodes}} \mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0]] = 0$ . To that end, we need the following lemma.

**Lemma 8.** [15] *For a sequence of random variables  $\{X_n\}$ , and a sequence of events  $\{\mathcal{A}_n\}$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{A}_n) = 0$ , if and only if  $\lim_{n \rightarrow \infty} \mathbb{P}[\mathbb{P}(\mathcal{A}_n|X_n) > \tau_n] = 0$  for some sequence  $\{\tau_n\}$  with  $\tau_n > 0$  and  $\lim_{n \rightarrow \infty} \tau_n = 0$ .*

From Lemma 8, to prove the secrecy constraint we only need to show that if  $R_n$  satisfies (63), then

$$\lim_{n \rightarrow \infty} \mathbb{P}_{CZ^n} \left[ \max_{R_n \text{ Hcodes}} \mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0] > \tau_n \right] = 0, \quad (64)$$

for some sequence  $\{\tau_n\}$  with  $\tau_n > 0$  and  $\lim_{n \rightarrow \infty} \tau_n = 0$ . Next we prove this.

Define event

$$\mathcal{A} \triangleq \{(S^n, Z'^n) \in \mathcal{T}_\delta^n(S, Z')\}, \quad (65)$$

for  $\delta > 0$ . The  $\delta$ -typical set is defined according to the notion of strong typicality, see [30]:

$$\mathcal{T}_\delta^n(S) \triangleq \{s^n \in \mathcal{S}^n : \sum_{s \in S} |T_{s^n}(s) - P_S(s)| \leq \delta\}, \quad (66)$$

where  $T_{s^n}$  denotes the type (or empirical distribution) of  $s^n$ . For simplicity,  $\mathcal{T}_\delta^n(S)$  is also shortly denoted as  $\mathcal{T}_\delta^n$ .

Since  $(S^n, Z'^n)$  is an i.i.d. sequence, from the fact that the typical set has total probability close to one [30], we have the following lemma.

**Lemma 9.** [30] For any  $\delta > 0$ ,  $\mathbb{P}[\mathcal{A}] \rightarrow 1$ , as  $n \rightarrow \infty$ .

Consider that for each  $n$ , the optimal  $R_n$ -rate henchman code that maximizes  $\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n]$  is adopted, then we only need to show  $\lim_{n \rightarrow \infty} \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n] = 0$  for these codes. By utilizing Lemmas 8 and 9, we have

$$\begin{aligned} & \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n] \\ & \leq \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n, \mathbb{P}[\mathcal{A}^c | \mathcal{C}Z^n] \leq \epsilon_n] + \mathbb{P}[\mathbb{P}[\mathcal{A}^c | \mathcal{C}Z^n] > \epsilon_n] \end{aligned} \quad (67)$$

$$\leq \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] + \mathbb{P}[\mathcal{A}^c | \mathcal{C}Z^n] > \tau_n, \mathbb{P}[\mathcal{A}^c | \mathcal{C}Z^n] \leq \epsilon_n] + \epsilon'_n \quad (68)$$

$$\leq \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] > \tau'_n] + \epsilon'_n, \quad (69)$$

for some  $\epsilon_n$  and  $\epsilon'_n$  that both vanish as  $n \rightarrow \infty$ , where  $\tau'_n = \tau_n - \epsilon_n$ . By choosing proper  $\tau_n$ ,  $\tau'_n$  can be set to some sequence that converges to zero sub-exponentially fast (i.e.,  $\tau'_n = 2^{-o(n)}$ ). Since  $\epsilon_n$  vanishes as  $n \rightarrow \infty$ , this guarantees that  $\tau_n$  also vanishes as  $n \rightarrow \infty$ .

Owing to the rate constraint, given  $(\mathcal{C}, Z^n)$ , the reconstruction  $\check{S}^n$  cannot take more than  $R_n$  values. Denote the set of possible values as  $c(\mathcal{C}, Z^n)$ , then

$$\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] = \mathbb{P}\left[\min_{\check{s}^n \in c(\mathcal{C}, Z^n)} d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n\right]. \quad (70)$$

Now we apply a union bound to the right-hand side of (70) and write

$$\begin{aligned} & \mathbb{P}\left[\min_{\check{s}^n \in c(\mathcal{C}, Z^n)} d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n\right] \\ & \leq \sum_{\check{s}^n \in c(\mathcal{C}, Z^n)} \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] \end{aligned} \quad (71)$$

$$\leq 2^{nR_n} \max_{\check{s}^n \in c(\mathcal{C}, Z^n)} \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] \quad (72)$$

$$\leq 2^{nR_n} \max_{\check{s}^n \in \mathcal{S}^n} \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n] \quad (73)$$

$$= 2^{nR_n} \max_{\check{s}^n \in \mathcal{S}^n} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[K = k | \mathcal{C}Z^n] \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \mathcal{C}Z^n, K = k] \quad (74)$$

$$= 2^{n(R_n - R_K)} \max_{\check{s}^n \in \mathcal{S}^n} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \Psi_k, Z^n], \quad (75)$$

where (75) follows from the Markov chain  $\mathcal{C}KZ^n \rightarrow \Psi_K Z^n \rightarrow S^n Z^n \mathcal{A}$  and  $\mathbb{P}[K = k | \mathcal{C} = c, Z^n = z^n] = 2^{-nR_K}$  (see (61)).

Combine (69), (70), and (75), then we have

$$\begin{aligned} & \mathbb{P}_{\mathcal{C}Z^n}[\mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n] \\ & \leq \mathbb{P}_{\mathcal{C}Z^n}\left[\max_{\check{s}^n \in \mathcal{S}^n} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[d_E(S^n, \check{s}^n) \leq D_0, \mathcal{A} | \Psi_k, Z^n] > \tau'_n 2^{-n(R_n - R_K)}\right] + \epsilon'_n \end{aligned} \quad (76)$$

$$\leq |\check{\mathcal{S}}^n| \max_{\check{s}^n \in \mathcal{S}^n} \mathbb{P}_{\mathcal{C}Z^n}\left[\sum_{k=1}^{2^{nR_K}} \xi_{k, z^n}(\check{s}^n) > \tau'_n 2^{-n(R_n - R_K)}\right] + \epsilon'_n, \quad (77)$$



where

$$\xi_{k,z^n}(\tilde{s}^n) \triangleq \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, \mathcal{A} | \Psi_k, Z^n], \quad (78)$$

Therefore, if we can show that the probability in (77) decays doubly exponentially fast with  $n$ , then the proof will be complete.

Consider that given  $\tilde{s}^n$  and  $z^n$ ,  $\xi_{k,z^n}(\tilde{s}^n)$ ,  $k \in [2^{nR_K}]$  are i.i.d. random variables, with mean

$$\mathbb{E}_C \xi_{k,z^n}(\tilde{s}^n) = \mathbb{E}_C \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, \mathcal{A} | \Psi_k, z^n] \quad (79)$$

$$= \mathbb{E}_{\Psi_k} \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, \mathcal{A} | \Psi_k, z^n] \quad (80)$$

To complete the proof, we need introduce the following lemmas. The proof of Lemma 10 is given in Appendix B.

**Lemma 10.** Assume  $S^n$  is i.i.d. according to  $P_S$ , then for any type  $t$  of sequences in  $\mathcal{S}^n$  and any  $\tilde{s}^n$ ,

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n | T_{S^n} = t] \leq 2^{-n(R_S(D) - o(1))}, \quad (81)$$

where  $T_{S^n}$  denotes the type of  $S^n$ , and  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

**Lemma 11.** [14] Fix  $P_{S|Z}$  and  $z^n$ . If  $S^n$  is distributed according to  $\prod_{i=1}^n P_{S|Z=z_i}$ , then for any  $\tilde{s}^n$ ,

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, (S^n, z^n) \in \mathcal{T}_\delta^n | z^n] \leq 2^{-n(R_{S|Z}(D) - o(1))}, \quad (82)$$

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

**Lemma 12.** [14] If  $X^m$  is a sequence of i.i.d. random variables on the interval  $[0, a]$  with  $\mathbb{E}[X_i] = p$ , then

$$\mathbb{P}\left[\sum_{i=1}^m X_i > k\right] \leq \left(\frac{e \cdot m \cdot p}{k}\right)^{k/a}. \quad (83)$$

From (60), we have

$$\mathbb{P}[S^n = s^n | \Psi_k, z^n] = \prod P_{S|Z}(s_i | z'_i). \quad (84)$$

Hence Lemma 11 implies

$$\xi_{k,z^n}(\tilde{s}^n) = \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, \mathcal{A} | z'^n] \quad (85)$$

$$\leq 2^{-n(R_{S|Z}(D_0) - o(1))}. \quad (86)$$

On the other hand,

$$\mathbb{E}_C \xi_{k,z^n}(\tilde{s}^n) \leq \mathbb{E}_{\Psi_k} \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, S^n \in \mathcal{T}_\delta^n | \Psi_k, z^n] \quad (87)$$

$$= \sum_{s'^n} \mathbb{P}[S'^n = s'^n | z^n] \mathbb{E}_{\Psi_k} \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, S^n \in \mathcal{T}_\delta^n | S'^n = s'^n, \Psi_k] \quad (88)$$

$$= \sum_{s'^n} \mathbb{P}[S'^n = s'^n | z^n] \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, S^n \in \mathcal{T}_\delta^n | S'^n = s'^n] \quad (89)$$

$$= \sum_{s'^n} \mathbb{P}[S'^n = s'^n | z^n] \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D_0, S^n \in \mathcal{T}_\delta^n | T_{S^n} = T_{s'^n}] \quad (90)$$

$$\leq 2^{-n(R_S(D_0) - o(1))}. \quad (91)$$

Using these bounds, we apply Lemma 12 to the probability in (77) by identifying

$$m = 2^{nR_K}, \quad (92)$$

$$a = 2^{-n(R_{S|Z}(D_0) - o(1))}, \quad (93)$$

$$p \leq 2^{-n(R_S(D_0) - o(1))}, \quad (94)$$

$$k = \tau'_n 2^{-n(R_n - R_K)}. \quad (95)$$

Then we have

$$\mathbb{P}\left[\sum_{k=1}^{2^{nR_K}} \xi_{k,z^n}(\tilde{s}^n) > \tau'_n 2^{-n(R_n - R_K)}\right] \leq 2^{-n\alpha 2^{n\beta}}, \quad (96)$$

where

$$\begin{aligned} \alpha &= R_S(D_0) - R_n - o(1) \\ \beta &= R_K + R_{S|Z}(D_0) - R_n - o(1). \end{aligned} \quad (97)$$

For small enough  $\delta$  and large enough  $n$ , both  $\alpha$  and  $\beta$  are positive and bounded away from zero, and (96) vanishes doubly exponentially fast. Therefore, the expression in (77) vanishes. This completes the proof of Theorem 1.

## APPENDIX B

### PROOF OF LEMMA 10

If  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| > \delta$ , then  $\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n | T_{S^n} = t] = 0$ . Hence we only need to consider the  $t$ 's such that  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| \leq \delta$ .

Consider

$$\mathbb{P}[T_{S^n} = t] = |\{s'^n \in \mathcal{S}^n : T_{s'^n} = t\}| 2^{-n(D(t||P_S) + H(t))} \quad (98)$$

$$= 2^{-n(D(t||P_S) + o(1))} \quad (99)$$

for any type  $t$  of sequences in  $\mathcal{S}^n$ , where  $D(t||P_S)$  denotes the relative entropy between  $t$  and  $P_S$ , and (99) follows from (15). Moreover, from [32, Thm. 25] we have

$$D(t||P_S) \leq \log \left( 1 + \frac{(\sum_s |t(s) - P_S(s)|)^2}{2P_{S,\min}} \right) \leq \log \left( 1 + \frac{\delta^2}{2P_{S,\min}} \right) \rightarrow 0, \quad (100)$$

as  $\delta \rightarrow 0$ , where  $P_{S,\min} = \min_{s \in \mathcal{S}} P_S(s)$ . Therefore,

$$\mathbb{P}[T_{S^n} = t] \geq 2^{-no(1)}. \quad (101)$$

Utilizing (101), we get

$$\begin{aligned} &\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n | T_{S^n} = t] \\ &= \frac{\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n, T_{S^n} = t]}{\mathbb{P}[T_{S^n} = t]} \end{aligned} \quad (102)$$

$$\leq \frac{\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n]}{2^{-no(1)}}. \quad (103)$$

To complete the proof, we need the following lemma.

**Lemma 13.** [14] Assume  $S^n$  is i.i.d. according to  $P_S$ , then for any  $\tilde{s}^n$ ,

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n] \leq 2^{-n(R_S(D)-o(1))}. \quad (104)$$

By the lemma above, (103) implies that

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{T}_\delta^n | T_{S^n} = t] \leq 2^{-n(R_S(D)-o(1))}. \quad (105)$$

## APPENDIX C

### PROOF OF THEOREM 2

Define  $X'^n, Y_i'^n, Z'^n$  same as (55)-(57), then the distribution  $P_{\mathcal{C}S^n K S'^n X^n Y_i^n Z^n X'^n Y_i'^n Z'^n \hat{S}_i'^n \hat{S}_i^n}$  also satisfies (60) and (61). Similar to the finite alphabet case, it is easy to show the distortion constraints for legitimate users are satisfied.

Next following similar steps to the proof for the finite alphabet case, we prove the secrecy constraint is also satisfied for this case. Before proving that, we need introduce d-tilted information and conditional d-tilted information first.

Let  $P_{\tilde{S}^*|S}$  be a distribution that achieves the rate-distortion function  $R_S(D)$  (which is not necessarily unique). Then d-tilted information is defined as follows.

**Definition 6** (d-tilted information [15]). For  $D > D_{\min} \triangleq \inf \{D: R_S(D) < \infty\}$ , the d-tilted information in  $s$  is defined as

$$J_S(s, D) = \log \frac{1}{\mathbb{E}[\exp(\lambda^* D - \lambda^* d(s, \tilde{S}^*))]}, \quad (106)$$

where the expectation is with respect to  $P_{\tilde{S}^*}$ , i.e. the unconditional distribution of the reproduction random variable that achieves  $R_S(D)$ , and

$$\lambda^* = -R'_S(D). \quad (107)$$

For  $(S, Z)$  that follow the distribution in (18), we define

$$R_{S|Z=z}(\beta) = \min_{P_{\tilde{S}|S, Z=z}: \mathbb{E}[d_E(S, \tilde{S})|Z=z] \leq \beta} I(S; \tilde{S}|Z=z). \quad (108)$$

Let  $P_{\tilde{S}^*|S, Z=z}$  be a distribution that achieves  $R_{S|Z=z}(\beta)$ . Define  $b^*(z) \triangleq \mathbb{E}_{S, \tilde{S}^*|Z=z} d(S, \tilde{S}^*)$  with the expectation taken with respect to  $P_{S|Z=z} P_{\tilde{S}^*|S, Z=z}$ .

**Definition 7** (Conditional d-tilted information [15]). For  $b^*(z) > \beta_{\min}(z) \triangleq \inf \{\beta: R_{S|Z=z}(\beta) < \infty\}$ , the conditional d-tilted information in  $s$  under condition  $Z = z$  is defined as

$$J_{S|Z=z}(s, b^*(z)) = \log \frac{1}{\mathbb{E}_{\tilde{S}^*|Z=z} \left[ \exp \left( \lambda^*(z) b^*(z) - \lambda^*(z) d(s, \tilde{S}^*) \right) \right]}, \quad (109)$$

where the expectation is with respect to  $P_{\tilde{S}^*|Z=z}$ , i.e. the margin distribution of  $P_{S|Z=z} P_{\tilde{S}^*|S, Z=z}$ , and

$$\lambda^*(z) = -R'_{S|Z=z}(b^*(z)). \quad (110)$$

Next we prove the secrecy constraint. To that end, we need re-define

$$\mathcal{A} \triangleq \left\{ S^n \in \mathcal{U}_\delta^n, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D_0) \geq R_S(D_0) - \delta, \right. \\ \left. \frac{1}{n} \sum_{i=1}^n J_{S|Z=Z'_i}(S_i, b^*(Z'_i)) \geq R_{S|Z}(D_0) - \delta, \frac{1}{n} \sum_{i=1}^n b^*(Z'_i) \geq D_0 - \delta \right\} \quad (111)$$

for  $\delta > 0$ . The  $\delta$ -unified typical set is defined as<sup>16</sup>

$$\mathcal{U}_\delta^n(S) \triangleq \mathcal{T}_{\frac{\delta}{\log n}}^n(S) \cap \mathcal{W}_\delta^n(S), \quad (112)$$

where  $\mathcal{T}_{\frac{\delta}{\log n}}^n(S)$  defined in (66), denotes the  $\frac{\delta}{\log n}$ -strongly typical set, and

$$\mathcal{W}_\delta^n(S) \triangleq \left\{ s^n \in \mathcal{S}^n : \left| -\frac{1}{n} \log P_{S^n}(s^n) - H(S) \right| \leq \delta \right\}, \quad (113)$$

denotes the  $\delta$ -weakly typical set [28]. For simplicity,  $\mathcal{U}_\delta^n(S)$  is also shortly denoted as  $\mathcal{U}_\delta^n$ .

Since  $(S^n, Z^n)$  is an i.i.d. sequence, we have the following lemma.

**Lemma 14.** [15, Lem. 18] [34, Lem. 2] Assume  $P_S$  satisfies  $\tilde{N}_{P_S}\left(\frac{\delta'}{\log n}\right) = o\left(\frac{n}{\log^2 n}\right)$ ,  $\forall 0 < \delta' \leq 1$ . Then for any  $\delta > 0$ ,  $\mathbb{P}[\mathcal{A}] \rightarrow 1$  as  $n \rightarrow \infty$ .

Then the derivation up to (77) still holds, i.e.,

$$\mathbb{P}_{\mathcal{C}Z^n} \left[ \mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n \right] \\ \leq \left| \check{\mathcal{S}}^n \right| \max_{\check{s}^n \in \check{\mathcal{S}}^n} \mathbb{P}_{\mathcal{C}Z^n} \left[ \sum_{k=1}^{2^{nR_K}} \xi_{k, z^n}(\check{s}^n) > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n, \quad (114)$$

Therefore, if we can show that the probability in (114) decays doubly exponentially fast with  $n$ , then the proof will be complete. To that end, we need introduce the following lemmas. The proof of Lemma 15 is given in Appendix D.

**Lemma 15.** Assume  $P_S$  satisfies  $N_{P_S}\left(\frac{1}{n}\right) = o\left(\frac{n}{\log n}\right)$ ,  $\Phi_{P_S}\left(\frac{1}{n}\right) = o\left(\frac{1}{\log n}\right)$ , and  $S^n$  is i.i.d. according to  $P_S$ , then for any type  $t$  of sequences in  $\mathcal{S}^n$  and any  $\check{s}^n \in \check{\mathcal{S}}^n$ ,

$$\mathbb{P}[d_E(S^n, \check{s}^n) \leq D, S^n \in \mathcal{U}_\delta^n, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta | T_{S^n} = t] \leq 2^{-n(R_S(D) - o(1))}, \quad (115)$$

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

**Lemma 16.** [15] Fix  $P_{SZ}$  and  $z^n \in \mathcal{Z}^n$ . Assume given  $Z^n = z^n$ ,  $S^n$  is distributed according to  $\prod_{i=1}^n P_{S|Z=z_i}$ , then for any  $\check{s}^n \in \check{\mathcal{S}}^n$ ,

$$\mathbb{P}\left[d_E(S^n, \check{s}^n) \leq D, \frac{1}{n} \sum_{i=1}^n J_{S|Z=z_i}(S_i, b^*(z_i)) \geq R_{S|Z}(D) - \delta, \frac{1}{n} \sum_{i=1}^n b^*(z_i) \geq D - \delta | Z^n = z^n\right] \\ \leq 2^{-n(R_{S|Z}(D) - o(1))}, \quad (116)$$

<sup>16</sup>Here the  $\delta$ -unified typical set is different from the one defined in [34]. Our definition has the benefit that it makes the following property hold: For each sequence  $s^n \in \mathcal{U}_\delta^n(S)$ ,  $|\{s'^n \in \mathcal{U}_\delta^n(S) : T_{s'^n} = T_{s^n}\}| = 2^{n(H(S) - o(1))}$ , or equivalently,  $\mathbb{P}[T_{S^n} = T_{s^n}] = 2^{-no(1)}$ , where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ . This property coincides with (101) for the finite alphabet case, and it is of crucial importance to our proof here (see (126)).

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

Apply Lemmas 12, 15 and 16, then we have that the probability in (114) decays doubly exponentially fast with  $n$ . This completes the proof of Theorem 2.

#### APPENDIX D

##### PROOF OF LEMMA 15

If  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| \leq \frac{\delta}{\log n}$  does not hold, then  $\mathbb{P}[d_E(S^n, \check{s}^n) \leq D, S^n \in \mathcal{U}_\delta^n | T_{S^n} = t] = 0$ . Hence we only need to consider the  $t$ 's satisfying  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| \leq \frac{\delta}{\log n}$ .

The Lemma 2.6 of [29] says that for any type  $t$  of sequences in  $S^n$ ,

$$\mathbb{P}[T_{S^n} = t] \geq (n+1)^{-|\text{supp}(t)|} 2^{-nD(t||P_S)}, \quad (117)$$

where  $\text{supp}(t) \triangleq \{s \in \mathcal{S} : t(s) > 0\}$  denotes the suppose of  $t$ .

Now we prove that for any  $\delta > 0$ ,  $|\text{supp}(t)| \leq \frac{n}{\log n} (\delta + \epsilon_n)$  holds, where  $\epsilon_n$  is a term that vanishes as  $n \rightarrow \infty$ . To that end, we divide  $\mathcal{S}$  into two parts:  $\{s : P_S(s) \geq \frac{1}{n}\}$  and  $\{s : P_S(s) < \frac{1}{n}\}$ . Then

$$|\text{supp}(t)| = \left| \left\{ s : t(s) > 0, P_S(s) \geq \frac{1}{n} \right\} \right| + \left| \left\{ s : t(s) > 0, P_S(s) < \frac{1}{n} \right\} \right| \quad (118)$$

$$\leq \left| \left\{ s : P_S(s) \geq \frac{1}{n} \right\} \right| + \left| \left\{ s : t(s) > 0, P_S(s) < \frac{1}{n} \right\} \right| \quad (119)$$

$$= N_{P_S} \left( \frac{1}{n} \right) + \sum_{s: P_S(s) < \frac{1}{n}} 1 \{t(s) > 0\} \quad (120)$$

$$\leq N_{P_S} \left( \frac{1}{n} \right) + n \sum_{s: P_S(s) < \frac{1}{n}} t(s), \quad (121)$$

where (120) follows from the definition of  $N_{P_S}(\frac{1}{n})$ , and (121) follows from the fact  $t(s) \geq \frac{1}{n}$  for any  $s$  such that  $t(s) > 0$ .

Since  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| \geq \sum_{s: P_S(s) < \frac{1}{n}} |t(s) - P_S(s)| \geq \sum_{s: P_S(s) < \frac{1}{n}} (t(s) - P_S(s))$  and  $\sum_{s \in \mathcal{S}} |t(s) - P_S(s)| \leq \frac{\delta}{\log n}$ , we have

$$\sum_{s: P_S(s) < \frac{1}{n}} t(s) \leq \sum_{s: P_S(s) < \frac{1}{n}} P_S(s) + \frac{\delta}{\log n} \quad (122)$$

$$= \Phi_{P_S} \left( \frac{1}{n} \right) + \frac{\delta}{\log n}. \quad (123)$$

Therefore,

$$|\text{supp}(t)| \leq N_{P_S} \left( \frac{1}{n} \right) + n \Phi_{P_S} \left( \frac{1}{n} \right) + \frac{\delta n}{\log n}. \quad (124)$$

Since  $N_{P_S}(\frac{1}{n}) = o\left(\frac{n}{\log n}\right)$ ,  $\Phi_{P_S}(\frac{1}{n}) = o\left(\frac{1}{\log n}\right)$ , we have  $|\text{supp}(t)| \leq \frac{n}{\log n} (\delta + \epsilon_n)$ . Therefore, (117) implies

$$\mathbb{P}[T_{S^n} = t] \geq 2^{-n(D(t||P_S) + \delta + \epsilon_n)}. \quad (125)$$

Furthermore,  $S^n \in \mathcal{U}_\delta^n$  implies  $D(t|P_S) \leq 2\delta$ , which is obtained by following part of proof steps of [34, Thm. 3] (but with  $\epsilon$  and  $\delta$  replaced with  $\delta$  and  $\frac{n}{\log n}$ , respectively). Hence it holds that

$$\mathbb{P}[T_{S^n} = t] \geq 2^{-no(1)}, \quad (126)$$

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ . Utilizing (126), we can get

$$\begin{aligned} & \mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{U}_\delta^n, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta | T_{S^n} = t] \\ &= \frac{\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{U}_\delta^n, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta, T_{S^n} = t]}{\mathbb{P}[T_{S^n} = t]} \end{aligned} \quad (127)$$

$$\leq \frac{\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta]}{2^{-no(1)}}. \quad (128)$$

To complete the proof, we need the following lemma.

**Lemma 17.** [15] Assume  $\mathcal{S}$  and  $\check{\mathcal{S}}$  are general (not necessarily countable) alphabets, and  $S^n$  is i.i.d. drawn from  $S^n$  according to  $P_S$ . Then for any  $D > D_{\min}$  ( $D_{\min}$  is defined in Definition 6) and any  $\tilde{s}^n \in \check{\mathcal{S}}^n$ ,

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta] \leq 2^{-n(R_S(D) - o(1))}. \quad (129)$$

Hence by the lemma above, (128) implies that

$$\mathbb{P}[d_E(S^n, \tilde{s}^n) \leq D, S^n \in \mathcal{U}_\delta^n, \frac{1}{n} \sum_{i=1}^n J_S(S_i, D) \geq R_S(D) - \delta | T_{S^n} = t] \leq 2^{-n(R_S(D) - o(1))}. \quad (130)$$

## APPENDIX E

### PROOF OF THEOREM 5

Similar to the DM cases, it is easy to show the distortion constraints for legitimate users are satisfied.

Next by following similar steps to the proof for the DM cases, we prove the secrecy constraint is also satisfied. To that end, we first need to discretize the source  $S$  and the reconstruction  $\check{S}$ . Let

$$[S], [\check{S}] \in \mathcal{N} \triangleq \{\dots, -2\Delta, -\Delta, 0, \Delta, 2\Delta, \dots\}, \quad (131)$$

be quantized versions of  $S$  and  $\check{S}$ , obtained by mapping  $S$  and  $\check{S}$  to the closest quantization point, i.e.,  $[S] = \Delta \cdot \text{Round}(\frac{S}{\Delta})$ ,  $[\check{S}] = \Delta \cdot \text{Round}(\frac{\check{S}}{\Delta})$ . Then we have for any  $s^n \in \mathbb{R}^n$ ,

$$0 \leq \frac{1}{n} \sum_{i=1}^n (s_i - [s_i])^2 \leq \frac{\Delta^2}{4}. \quad (132)$$

Furthermore, it holds that

$$\sqrt{nd(s^n, \tilde{s}^n)} = \|s^n - [\tilde{s}]^n + [\tilde{s}]^n - \tilde{s}^n\| \quad (133)$$

$$\geq \|s^n - [\tilde{s}]^n\| - \|[\tilde{s}]^n - \tilde{s}^n\| \quad (134)$$

$$\geq \|[\tilde{s}]^n - [\tilde{s}]^n\| - \|s^n - [\tilde{s}]^n\| - \|[\tilde{s}]^n - \tilde{s}^n\| \quad (135)$$

$$\geq \sqrt{nd([\tilde{s}]^n, [\tilde{s}]^n)} - \Delta \quad (136)$$

where (134) follows from triangle inequality. Utilizing this inequality, we have

$$\mathbb{P}[d(S^n, \check{s}^n) \leq D_0, \mathcal{A}|\Psi_k, Z^n] \leq \mathbb{P}\left[d([S]^n, [\check{s}]^n) \leq \left(\sqrt{D_0} + \Delta\right)^2, \mathcal{A}|\Psi_k, Z^n\right] \quad (137)$$

$$= \mathbb{P}[d([S]^n, [\check{s}]^n) \leq D'_0, \mathcal{A}|\Psi_k, Z^n], \quad (138)$$

where  $D'_0 \triangleq (\sqrt{D_0} + \Delta)^2$ .

Reorder the probabilities  $P_{[S]}([s]), [s] \in \mathcal{N}$  in decreasing order, and denote the result as  $P_i, i = 1, 2, \dots$ . Then  $P_1 = P_{[S]}(0)$ ,  $P_{2j} = P_{2j+1} = P_{[S]}(j\Delta), j \geq 1$ . Obviously,

$$\Delta f_S((j+1)\Delta) \leq P_{2j} = P_{2j+1} \leq \Delta f_S(j\Delta), \quad (139)$$

and hence for Gaussian sources,  $P_{2j} = P_{2j+1} = o(e^{-j^2})$ . From Remark 2, we have that  $P_{[S]}$  satisfies the conditions (23)-(25).

Define event

$$\begin{aligned} \mathcal{A} \triangleq & \left\{ S^n \in \mathcal{W}_\delta^n, [S]^n \in \mathcal{U}_\delta^n([S]), \frac{1}{n} \sum_{i=1}^n J_S([S]_i, D_0) \geq R_{[S]}(D_0) - \delta, \right. \\ & \left. \frac{1}{n} \sum_{i=1}^n J_{[S]|Z}(Z'_i, b^*(Z'_i)) \geq R_{[S]|Z}(D_0) - \delta, \frac{1}{n} \sum_{i=1}^n b^*(Z'_i) \geq D_0 - \delta \right\} \end{aligned} \quad (140)$$

for  $\delta > 0$ . Observe that the distribution  $P_{\mathcal{C}S^n K S'^n X^n Y^n Z^n X'^n Y'^n Z'^n \hat{S}_i^n \hat{S}'_i^n}$  also satisfies (60) and (61), which implies  $(S^n, Z'^n)$  is an i.i.d. sequence. Hence Lemma 14 still holds for this case. Following similar steps to the proof of Theorem 1, we can get

$$\begin{aligned} & \mathbb{P}_{\mathcal{C}Z^n} \left[ \mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n \right] \\ & \leq \mathbb{P}_{\mathcal{C}Z^n} \left[ \max_{[\check{s}]^n \in \mathcal{N}^n} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[d([S]^n, [\check{s}]^n) \leq D'_0, \mathcal{A}|\Psi_k, Z^n] > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n \end{aligned} \quad (141)$$

$$= \mathbb{P}_{\mathcal{C}Z^n} \left[ \max_{[\check{s}]^n \in \mathcal{B}^n} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[d([S]^n, [\check{s}]^n) \leq D'_0, \mathcal{A}|\Psi_k, Z^n] > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n \quad (142)$$

$$\leq |\mathcal{B}^n| \max_{[\check{s}]^n \in \mathcal{B}^n} \mathbb{P}_{\mathcal{C}Z^n} \left[ \sum_{k=1}^{2^{nR_K}} \xi_{k, z^n}([\check{s}]^n) > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n, \quad (143)$$

where

$$\mathcal{B}^n \triangleq \left\{ [\check{s}]^n \in \mathcal{N}^n : \|\check{s}^n\| \leq \sqrt{n\Gamma} \right\} \quad (144)$$

with  $\sqrt{\Gamma} \triangleq \sqrt{\lambda(1+\delta)} + \sqrt{\Delta} + \sqrt{D'_0}$ ,

$$\xi_{k, z^n}([\check{s}]^n) \triangleq \mathbb{P}[d([S]^n, [\check{s}]^n) \leq D'_0, \mathcal{A}|\Psi_k, Z^n], \quad (145)$$

and (142) follows that  $S^n$  only appears in the ball with radius  $\sqrt{n\lambda(1+\delta)}$  which implies  $[S]^n$  only appears in the ball with radius  $\sqrt{\lambda(1+\delta)} + \sqrt{\Delta}$ , hence it is sufficient to only consider  $[\check{s}]^n \in \mathcal{B}^n$  instead of the whole set  $\mathcal{N}^n$ .

Furthermore, observe that

$$|\mathcal{B}^n| \leq \frac{\text{Volume of } n\text{-ball with radius } \sqrt{n\Gamma} + \sqrt{n\Delta^2}}{\Delta^n} \quad (146)$$

$$= \frac{\pi^{n/2} \left( \sqrt{n\Gamma} + \sqrt{n\Delta^2} \right)^n}{\Delta^n \Gamma\left(\frac{n}{2} + 1\right)} \quad (147)$$

$$\leq 2^{\frac{n}{2} \log \pi + n \log(\sqrt{n\Gamma} + \sqrt{n\Delta^2}) - n \log \Delta - \frac{n}{2} \log \frac{n}{2e} - \frac{1}{2} \log \pi n + o(1)} \quad (148)$$

$$= 2^{\frac{n}{2} \log \frac{2\pi e (\sqrt{\Gamma} + \sqrt{\Delta^2})^2}{\Delta^2} - \frac{1}{2} \log \pi n + o(1)}. \quad (149)$$

Therefore, if we can show that the probability in (143) decays doubly exponentially fast with  $n$ , then the proof will be complete.

Apply Lemmas 12, 15 and 16, then we have the probability in (143) decays doubly exponentially fast with  $m$ . Hence  $\lim_{n \rightarrow \infty} \mathbb{E}_{CZ^n} \left[ \max_{R_n \text{ Hcodes}} \mathbb{P}[d_E(S^n, \check{S}^n) \leq D_0] \right] = 0$  if

$$\limsup_{n \rightarrow \infty} R_n < \min \{ R_K + R_{[S]|Z}(D'_0), R_{[S]}(D'_0) \}. \quad (150)$$

To complete the proof, we need to show  $R_{[S]|Z}(D'_0) \geq R_{S|Z}(D_0)$  and  $R_{[S]}(D'_0) \geq R_S(D_0)$  as  $\Delta \rightarrow 0$ . Suppose  $P_{[\check{S}]^*|[S]Z}$  achieves  $R_{[S]|Z}(D'_0)$ , then  $R_{[S]|Z}(D'_0) = I([S]; [\check{S}]^*|Z)$  and  $\mathbb{E}d([S], [\check{S}]^*) \leq D'_0$ . Since for  $P_{[S][\check{S}]^*|SZ} = P_{[S]|S}P_{[\check{S}]^*|[S]Z}$ ,

$$\mathbb{E}d(S, [\check{S}]^*) = \mathbb{E} \left( S - [S] + [S] - [\check{S}]^* \right)^2 \quad (151)$$

$$= \mathbb{E} \left( [S] - [\check{S}]^* \right)^2 + \mathbb{E} (S - [S])^2 + 2\mathbb{E} (S - [S]) \left( [S] - [\check{S}]^* \right) \quad (152)$$

$$\leq \mathbb{E}d([S], [\check{S}]^*) + \Delta \sqrt{\mathbb{E}d([S], [\check{S}]^*)} \quad (153)$$

$$\leq D'_0 + \Delta \sqrt{D'_0}, \quad (154)$$

where (153) follows from the Cauchy-Schwarz inequality. On the other hand,  $R_{S|Z}(D'_0 + \Delta \sqrt{D'_0})$  is defined as the minimum  $I(S; \check{S}|Z)$  such that  $\mathbb{E}d(S, \check{S}) \leq D'_0 + \Delta \sqrt{D'_0}$ . Hence  $R_{S|Z}(D'_0 + \Delta \sqrt{D'_0}) \leq I(S; [\check{S}]^*|Z) = I([S]; [\check{S}]^*|Z) = R_{[S]|Z}(D'_0)$ . Let  $\Delta \rightarrow 0$ , then we have  $R_{S|Z}(D_0) \leq \lim_{\Delta \rightarrow 0} R_{[S]|Z}(D'_0)$ . Similarly, we can prove  $R_S(D_0) \leq \lim_{\Delta \rightarrow 0} R_{[S]}(D'_0)$ . This completes the proof of Theorem 5.

## APPENDIX F

### PROOF OF THEOREM 6

Denote  $X'^n, Y_i'^n, Z'^n$  as (55)-(57), where  $\Psi_K(\cdot)$  denotes the orthogonal transform, instead of the permutation operation. Then it can be verified that for Gaussian source-channel pair, the distribution  $P_{C^{S^n} K^{S'^n} X^n Y_i^n Z^n X'^n Y_i'^n Z'^n \hat{S}_i'^n \hat{S}_i^n}$  also satisfies (60) and (61). Hence  $(\Psi_K, Z^n) \rightarrow Z'^n \rightarrow S^n$  forms a Markov chain.

Similar to the permutation-based scheme, it is easy to show the power constraint and the distortion constraints for legitimate users are satisfied. Next by following similar steps to the proof for the permutation-based scheme, we prove the secrecy constraint is also satisfied. But a slight difference is that here we use an argument from a geometric point of view, instead of the one from the view of rate-distortion theory (or method of types) used for Theorems



1, 2, and 5. Here we do not need to discretize the source. But we still need to discretize the reconstruction  $\tilde{S}$  as (131), since it will enable us to take the maximizing operation out of the probability, just as done in (141)-(143).

Define event

$$\mathcal{A} \triangleq \{(S^n, Z'^n) \in \mathcal{W}_\delta^n(S, Z)\}, \quad (155)$$

for  $\delta > 0$ . For jointly Gaussian variables  $X$  and  $Z$ , where  $Z = X + U$  and  $U$  is independent of  $X$ , the  $\delta$ -weakly typical set and the  $\delta$ -weakly jointly typical set become

$$\mathcal{W}_\delta^n(X) \triangleq \left\{x^n \in \mathcal{X}^n : \left| \frac{\|x^n\|^2}{nN_X} - 1 \right| \leq \delta \right\}, \quad (156)$$

and

$$\mathcal{W}_\delta^n(X, Z) \triangleq \left\{(x^n, z^n) \in \mathbb{R}^{2n} : \left| \frac{\|x^n\|^2}{nN_X} - 1 \right| \leq \delta, \right. \quad (157)$$

$$\left. \left| \frac{\|z^n\|^2}{nN_Z} - 1 \right| \leq \delta, \right. \quad (158)$$

$$\left. \left| \frac{\|x^n\|^2}{nN_X} + \frac{\|z^n - x^n\|^2}{nN_U} - 2 \right| \leq \delta \right\}, \quad (159)$$

respectively, where  $N_Z, N_X$  and  $N_U$  denote the variances of  $Z, X$  and  $U$ , respectively.

Since  $(S^n, Z'^n)$  is an i.i.d. sequence, from the fact that (weakly) typical set has total probability close to one [28], we have the following lemma.

**Lemma 18.** [28] For any  $\delta > 0$ ,  $\mathbb{P}[\mathcal{A}^c] \rightarrow 0$ , as  $n \rightarrow \infty$ .

Following similar steps to the proof of Theorem 5, we can get

$$\begin{aligned} & \mathbb{P}_{\mathcal{C}Z^n} \left[ \mathbb{P}[d(S^n, \tilde{S}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n \right] \\ & \leq |\mathcal{B}^n| \max_{[\tilde{s}]^n \in \mathcal{B}^n} \mathbb{P}_{\mathcal{C}Z^n} \left[ \sum_{k=1}^{2^{nR_K}} \xi_{k,z^n}([\tilde{s}]^n) > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n, \end{aligned} \quad (160)$$

where  $\mathcal{B}^n$  is given by (144), and

$$\xi_{k,z^n}([\tilde{s}]^n) \triangleq \mathbb{P}[d(S^n, [\tilde{s}]^n) \leq D'_0, \mathcal{A} | \Psi_k, Z^n]. \quad (161)$$

Since as shown in (149),  $|\mathcal{B}^n|$  is upper bounded by an exponential function, we only need to show that the probability in (160) decays doubly exponentially fast with  $n$ . To that end, we need introduce the following lemmas. The proofs of Lemmas 20 and 19 are given in Appendixes G and H, respectively.

**Lemma 19.** Assume  $S^n = Z^n + U^n$ , where  $Z^n \sim \mathcal{N}(\mathbf{0}, N_Z \mathbf{I})$  and  $U^n \sim \mathcal{N}(\mathbf{0}, N_U \mathbf{I})$ <sup>17</sup> are independent, then for any  $z^n, \bar{s}^n \in \mathbb{R}^n$ ,

$$\mathbb{P}[d(S^n, \bar{s}^n) \leq D, (S^n, z^n) \in \mathcal{W}_\delta^n | z^n] \leq 2^{-n(R_{S|Z}(D) - o(1))}, \quad (162)$$

where  $R_{S|Z}(D) = \frac{1}{2} \log^+ \left( \frac{N_U}{D} \right)$ , and  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

<sup>17</sup>Here  $\mathbf{0}$  denotes an all-zero vector and  $\mathbf{I}$  denotes an identity matrix

**Lemma 20.** Assume  $S^n \sim \mathcal{N}(\mathbf{0}, N_S \mathbf{I})$  and  $S'^n = \Psi S^n$ , with  $\Psi$  uniformly distributed on orthogonal matrices set and independent of  $S^n$ , then for any  $s'^n, \bar{s}^n \in \mathbb{R}^n$ ,

$$\mathbb{P}[d(S^n, \bar{s}^n) \leq D, S^n \in \mathcal{W}_\delta^n | s'^n] \leq 2^{-n(R_S(D) - o(1))}, \quad (163)$$

where  $R_S(D) = \frac{1}{2} \log^+ \left( \frac{N_S}{D} \right)$ , and  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ .

Then we have

$$\mathbb{E}_{\mathcal{C}} \xi_{k,z^n}([\tilde{s}]^n) \leq \mathbb{E}_{\Psi_k} \mathbb{P}[d(S^n, [\tilde{s}]^n) \leq D'_0, S^n \in \mathcal{W}_\delta^n | \Psi_k, z^n] \quad (164)$$

$$= \mathbb{E}_{\Psi_k} \int \mathbb{P}[d(S^n, [\tilde{s}]^n) \leq D'_0, S^n \in \mathcal{W}_\delta^n | \Psi_k, z^n, s'^n] f_{S'^n | Z^n}(s'^n | \Psi_k, z^n) ds'^n \quad (165)$$

$$= \int \mathbb{P}_{\Psi_k}[d(S^n, [\tilde{s}]^n) \leq D'_0, S^n \in \mathcal{W}_\delta^n | s'^n] f_{S'^n | Z^n}(s'^n | z^n) ds'^n \quad (166)$$

$$\leq 2^{-n(R_S(D'_0) - o(1))}, \quad (167)$$

where (167) follows from Lemma 20. Furthermore, Lemma 19 implies

$$\xi_{k,z^n}([\tilde{s}]^n) = \mathbb{P}[d(S^n, [\tilde{s}]^n) \leq D'_0, \mathcal{A} | z'^n] \quad (168)$$

$$\leq 2^{-n(R_{S|Z}(D'_0) - o(1))}, \quad (169)$$

where (167) follows from  $\Psi_K Z^n \rightarrow Z'^n \rightarrow S^n$ .

Applying Lemmas 12, we have that the probability in (160) decays doubly exponentially fast with  $n$ . This completes the proof of Theorem 6.

## APPENDIX G

### PROOF OF LEMMA 19

Consider that

$$\begin{aligned} & \mathbb{P}[d(S^n, \bar{s}^n) \leq D, (S^n, z^n) \in \mathcal{W}_\delta^n | z^n] \\ & \leq \mathbb{P}[d(S^n, \bar{s}^n) \leq D, \frac{\|z^n\|^2}{nN_Z} \in 1 \pm \delta, \frac{\|z^n\|^2}{nN_Z} + \frac{\|S^n - z^n\|^2}{nN_U} \in 2 \pm \delta | z^n] \end{aligned} \quad (170)$$

$$\leq \mathbb{P}[d(S^n, \bar{s}^n) \leq D, \frac{\|S^n - z^n\|^2}{nN_U} \in 1 \pm 2\delta | z^n]. \quad (171)$$

Denote  $R = \frac{1}{n} \|S^n - z^n\|^2$ , then

$$\begin{aligned} & \mathbb{P}[d(S^n, \bar{s}^n) \leq D, (S^n, z^n) \in \mathcal{W}_\delta^n | z^n] \\ & \leq \int_{N_U(1-2\delta)}^{N_U(1+2\delta)} f_{R|Z^n}(r | z^n) \mathbb{P}[d(S^n, \bar{s}^n) \leq D | R = r, Z^n = z^n] dr. \end{aligned} \quad (172)$$

Given  $Z^n = z^n$ ,  $S^n - z^n \sim \mathcal{N}(\mathbf{0}, N_U \mathbf{I})$ , and on the other hand, Gaussian distribution is isotropic (or invariant to rotation). Hence under condition of  $Z^n = z^n$  and  $R = r$ ,  $S^n$  is uniformly distributed over the sphere with center  $z^n$  and radius  $\sqrt{nr}$ .

$$\mathbb{P}[d(S^n, \bar{s}^n) \leq D | R = r, Z^n = z^n] \leq \frac{\Omega(\theta)}{\Omega(\pi)}, \quad (173)$$

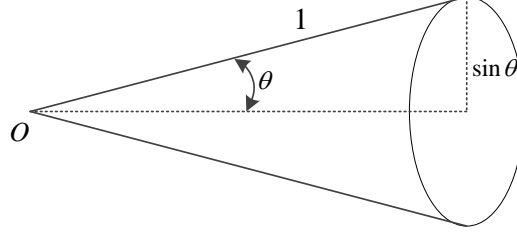


Fig. 6. Cap cut out by the cone on the unit sphere.

where

$$\theta = \arcsin \sqrt{\frac{D}{r}}, \quad (174)$$

and  $\Omega(\theta)$  be solid angle in  $n$  space of a cone with half-angle  $\theta$ , i.e., the area of a spherical cap on a unit sphere (see Fig. 6). To approximate  $\frac{\Omega(\theta)}{\Omega(\pi)}$ , we need the following lemma.

**Lemma 21.** [33] *Let  $\Omega(\theta)$  be solid angle in  $n$  space of a cone with half-angle  $\theta$ , then it holds that*

$$\frac{\Omega(\theta)}{\Omega(\pi)} = \frac{\sin^{n-1} \theta}{\sqrt{2\pi n} \cos \theta} \left( 1 + O\left(\frac{1}{n}\right) \right). \quad (175)$$

Lemma 21 implies

$$\frac{\Omega(\theta)}{\Omega(\pi)} = 2^{n(\log \sin \theta - \frac{1}{n} \log(\sqrt{2\pi n} \sin \theta \cos \theta) + \frac{1}{n} \log(1 + O(\frac{1}{n})))} \quad (176)$$

$$= 2^{n(\log \sin \theta + o(1))}. \quad (177)$$

Combine (173), (174) and (177), then we have for any  $r \in N_U(1 \pm 2\delta)$ ,

$$\mathbb{P}[d(S^n, \bar{s}^n) \leq D | R = r, Z^n = z^n] \leq 2^{n(\log \sqrt{\frac{D}{r}} + o(1))} \quad (178)$$

$$\leq 2^{n(\log \sqrt{\frac{D}{N_U(1-2\delta)}} + o(1))} \quad (179)$$

$$= 2^{-n(\frac{1}{2} \log \frac{N_U}{D} - o(1))}, \quad (180)$$

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ . Combining (172) and (180) gives us  $\mathbb{P}[d(S^n, \bar{s}^n) \leq D, (S^n, z^n) \in \mathcal{W}_\delta^n | z^n] \leq 2^{-n(\frac{1}{2} \log \frac{N_U}{D} - o(1))}$ . This completes the proof of Lemma 19.

## APPENDIX H

### PROOF OF LEMMA 20

Observe that

$$\begin{aligned} & \mathbb{P}[d(S^n, \bar{s}^n) \leq D, S^n \in \mathcal{W}_\delta^n | s'^n] \\ &= \mathbb{P}[d(S^n, \bar{s}^n) \leq D, \frac{\|S^n\|^2}{nN_S} \in 1 \pm \delta | s'^n] \end{aligned} \quad (181)$$

$$= \mathbb{P}_\Psi[d(\Psi^T s'^n, \bar{s}^n) \leq D | s'^n] 1 \left\{ \frac{\|s'^n\|^2}{nN_S} \in 1 \pm \delta \right\}. \quad (182)$$

Since  $\Psi$  is uniformly distributed on orthogonal matrices set (so is  $\Psi^T$  as stated in Lemma (4)), Lemma 5 implies that for any  $s'^n$ ,  $\Psi^T s'^n$  is uniformly distributed over the sphere with center at the origin  $O$  and radius  $\|s'^n\|$ . Hence

$$\mathbb{P}_\Psi[d(\Psi^T s'^n, \bar{s}^n) \leq D | s'^n] \leq \frac{\Omega(\theta)}{\Omega(\pi)}, \quad (183)$$

where as described in the previous section,  $\Omega(\theta)$  denotes solid angle in  $n$  space of a cone with half-angle  $\theta$ , and

$$\theta = \arcsin \sqrt{\frac{D}{\frac{1}{n} \|s'^n\|^2}}. \quad (184)$$

From Lemma 21, we have for any  $s'^n$  such that  $\frac{\|s'^n\|^2}{nN_S} \in 1 \pm \delta$ ,

$$\mathbb{P}_\Psi[d(\Psi^T s'^n, \bar{s}^n) \leq D | s'^n] \leq 2^{n \left( \log \sqrt{\frac{D}{\frac{1}{n} \|s'^n\|^2}} + o(1) \right)} \quad (185)$$

$$\leq 2^{n \left( \log \sqrt{\frac{D}{N_S(1-\delta)}} + o(1) \right)} \quad (186)$$

$$\leq 2^{-n \left( \frac{1}{2} \log \frac{N_S}{D} + o(1) \right)}, \quad (187)$$

where  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ . Combining (182) and (187) gives us

$$\mathbb{P}[d(S^n, \bar{s}^n) \leq D, S^n \in \mathcal{W}_\delta^n | s'^n] \leq 2^{-n \left( \frac{1}{2} \log \frac{N_S}{D} + o(1) \right)}. \quad (188)$$

This completes the proof of Lemma 20.

## APPENDIX I

### PROOF OF LEMMA 6

By choosing  $\check{S}$  to be independent of  $Z$ , we have  $R_{S|Z}(D_0) = \min_{P_{\check{S}|SZ}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S} | Z) \leq \min_{P_{\check{S}|S}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}) = R_S(D_0) = \frac{1}{2} \log^+ \left( \frac{\lambda}{D_0} \right)$ . Then we only need to prove  $R_{S|Z}(D_0) \leq R_{S|Z}^{(\text{UB})}(D_0)$ .

First consider the case of  $\frac{\lambda N_0}{P' + N_0} \leq D_0 \leq \lambda$ . Assume

$$Q = \begin{cases} 1, & \text{with probability } p; \\ 0, & \text{with probability } 1 - p, \end{cases} \quad (189)$$

independent of  $(S, Z)$ , denotes a timesharing random variable, and also assume

$$\check{S}_Q = \begin{cases} \beta_0 \Psi_K Z, & \text{if } Q = 1; \\ 0, & \text{if } Q = 0, \end{cases} \quad (190)$$

where  $\beta_0 = \frac{\sqrt{\lambda P'}}{P' + N_0}$  and  $\Psi_K$  is defined in (36). Then

$$\mathbb{E}d(S, \check{S}_Q) = \mathbb{E}_Q \mathbb{E} [d(S, \check{S}_Q) | Q] \quad (191)$$

$$= p \frac{\lambda N_0}{P' + N_0} + (1 - p) \lambda. \quad (192)$$

Therefore, to satisfy distortion constraint  $\mathbb{E}d(S, \check{S}_Q) \leq D_0$ , it is sufficient to set  $p = \frac{(\lambda - D_0)(P' + N_0)}{\lambda P'}$ . Substituting  $\check{S}_Q$  into  $R_{S|Z}(D_0)$ , we have

$$R_{S|Z}(D_0) \leq I(S; \check{S}_Q|Z) \quad (193)$$

$$\leq I(S; \check{S}_Q Q|Z) \quad (194)$$

$$= I(S; \check{S}_Q|QZ) \quad (195)$$

$$= pI(S; \check{S}_Q|Z, Q = 1) + (1 - p)I(S; \check{S}_Q|Z, Q = 0) \quad (196)$$

$$= pI(S; \check{S}_Q|Z, Q = 1) \quad (197)$$

$$\leq pI(S; K|Z, Q = 1) \quad (198)$$

$$\leq pH(K) \quad (199)$$

$$= p \quad (200)$$

$$= \frac{(\lambda - D_0)(P' + N_0)}{\lambda P'}, \quad (201)$$

where (195) follows from  $Q$  is independent of  $(S, Z)$ .

Next consider the case of  $0 \leq D_0 \leq \frac{\lambda N_0}{P' + N_0}$ . Observe that

$$R_{S|Z}(D_0) = \min_{P_{\check{S}|SZ}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}|Z) \quad (202)$$

$$= \min_{P_{\check{S}|SZK}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}|Z) \quad (203)$$

$$\leq \min_{P_{\check{S}|SZK}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}K|Z) \quad (204)$$

$$= I(S; K|Z) + \min_{P_{\check{S}|SZK}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}|ZK) \quad (205)$$

$$\leq H(K) + \min_{P_{\check{S}|SZK}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}|ZK), \quad (206)$$

where (203) follows since, on one hand, by setting  $P_{\check{S}|SZK}$  in (203) as  $P_{\check{S}|SZ}$  we have (202)  $\geq$  (203); on the other hand, given  $P_{SZK}$ , both the constraint and the optimization objective only depend on  $P_{\check{S}|SZ} = \sum_k P_{\check{S}|SZK} P_{K|SZ}$ , hence it suffices to optimize (203) over  $P_{\check{S}|SZ}$ .

The first term of (206) satisfies

$$H(K) = 1. \quad (207)$$

By (38) and (39), we have  $S$  and  $\Psi_K Z$  are jointly Gaussian, i.e.,

$$S = \beta_0 \Psi_K Z + V'_0 \quad (208)$$

where  $\beta_0 = \frac{\sqrt{\lambda P'}}{P' + N_0}$ ,  $\Psi_K$  is defined in (36), and  $V'_0 \sim \mathcal{N}\left(0, \frac{\lambda N_0}{P' + N_0}\right)$  is independent of  $\Psi_K Z$ . Hence we can also write

$$\check{S}^* = \beta_0 \Psi_K Z + V''_0 \quad (209)$$

and

$$S = \check{S}^* + \Delta V_0'', \quad (210)$$

where  $V_0'' \sim \mathcal{N}\left(0, \frac{\lambda N_0}{P' + N_0} - D_0\right)$  and  $\Delta V_0'' \sim \mathcal{N}(0, D_0)$  are independent of each other and also independent of  $\Psi_K Z$ . Therefore, we can bound the second term in (206) as

$$\min_{P_{\check{S}|SZK}: \mathbb{E}d(S, \check{S}) \leq D_0} I(S; \check{S}|ZK) \leq I(S; \check{S}^*|ZK) \quad (211)$$

$$= h(S|ZK) - h(S|ZK\check{S}^*) \quad (212)$$

$$= h(S|K) + h(Z|SK) - h(Z|K) - h(S - \check{S}^*|ZK\check{S}^*) \quad (213)$$

$$\leq \frac{1}{2} \log 2\pi e \lambda + \frac{1}{2} \log 2\pi e N_0 - \frac{1}{2} \log 2\pi e (P' + N_0) - h(S - \check{S}^*) \quad (214)$$

$$\leq \frac{1}{2} \log \left( \frac{\lambda N_0}{D_0 (P' + N_0)} \right), \quad (215)$$

where (211) follows since  $P_{\check{S}^*|SZK}$  satisfies the constraint  $\mathbb{E}d(S, \check{S}^*) \leq D_0$ , and (214) follows since  $\Psi_K Z \rightarrow \check{S}^* \rightarrow S$  forms a Markov chain and  $S - \check{S}^*$  is independent of  $\check{S}^*$ .

Combining (206), (207) and (215) gives us  $R_{S|Z}(D_0) \leq R_{S|Z}^{(\text{UB})}(D_0)$ . This completes the proof of Lemma 6.

## APPENDIX J

### PROOF OF THEOREM 11

Following similar steps to the proof for the scalar Gaussian case, it is easy to prove  $\mathcal{R}^{(i)}$  is achievable by the permutation based scheme. However, for the orthogonal-transform based scheme, the proof for the scalar Gaussian case cannot be applied to the vector Gaussian case directly, and some details need to be treated specially. Next we give a proof for this case.

Following similar steps to the proof of Theorem 6, it can be shown that for vector Gaussian case, the distortion constraints and power constraint are satisfied for the tuples given in Theorem 11. Next we prove the secrecy constraint is also satisfied.

Define events

$$\mathcal{A}_j \triangleq \{(S_j^n, Z_j'^n) \in \mathcal{W}_\delta^n(S_j, Z_j)\}, \quad (216)$$

$$\mathcal{A} \triangleq \prod_{j \in [m]} \mathcal{A}_j, \quad (217)$$

for  $\delta > 0$ . Similar to Lemma 18, it can be shown that for any  $\delta > 0$ ,  $\mathbb{P}[\mathcal{A}_j^c] \rightarrow 0$ , as  $n \rightarrow \infty$ .

The derivation up to (160) still holds for vector Gaussian case. Hence

$$\begin{aligned} & \mathbb{P}_{\mathcal{C}Z^n} \left[ \max_{R_n \text{ Hcodes}} \mathbb{P}[d(\mathbf{S}^n, \check{\mathbf{S}}^n) \leq D_0 | \mathcal{C}Z^n] > \tau_n \right] \\ & \leq \mathbb{P}_{\mathcal{C}Z^n} \left[ \max_{\check{\mathbf{s}}^n \in \mathbb{R}^{mn}} \sum_{k=1}^{2^{nR_K}} \mathbb{P}[d(\mathbf{S}^n, [\check{\mathbf{s}}]^n) \leq D'_0, \mathcal{A} | \Psi_{j,k}, j \in [m], \mathbf{Z}^n] > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n, \end{aligned} \quad (218)$$

where  $D'_0 \triangleq (\sqrt{D_0} + m\Delta)^2$ .

Observe that  $\mathbb{P}[d(\mathbf{S}^n, [\tilde{\mathbf{s}}]^n) \leq D_0, \mathcal{A}|\Psi_{j,k}, j \in [m], \mathbf{Z}^n] = \int_{\sum_{j=1}^m d(s_j^n, [\tilde{s}_j]^n) \leq D'_0, \mathcal{A}} \prod_{j=1}^m f(s_j^n | \Psi_{j,k}, z_j^n) ds_j^n$ . One may expect to exchange  $\int$  with  $\prod$ , in order to write the expression as  $\prod_{j=1}^m \mathbb{P}[d(S_j^n, [\tilde{s}_j]^n) \leq d_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n]$  for some  $d_j, j \in [m]$  such that  $\sum_{j=1}^m d_j \leq D'_0$ . However, obviously this is not feasible. To address this problem, we need to discretize the source, and then eliminate the  $\int$  operation since after discretization it becomes a  $\sum$  operation with the number of summands polynomial in  $n$ .

Discretize  $S$  by  $[S] = \Delta \cdot \text{Round}(\frac{S}{\Delta})$ . Then we have

$$\begin{aligned} & \mathbb{P}[d(\mathbf{S}^n, [\tilde{\mathbf{s}}]^n) \leq D'_0, \mathcal{A}|\Psi_{j,k}, j \in [m], \mathbf{Z}^n] \\ & \leq \mathbb{P}[d([\mathbf{S}]^n, [\tilde{\mathbf{s}}]^n) \leq D''_0, \mathcal{A}|\Psi_{j,k}, j \in [m], \mathbf{Z}^n], \end{aligned} \quad (219)$$

where  $D''_0 \triangleq (\sqrt{D'_0} + m\Delta)^2$ . In addition, observe that

$$\begin{aligned} & \mathbb{P}[d([\mathbf{S}]^n, [\tilde{\mathbf{s}}]^n) \leq D'_0, \mathcal{A}|\Psi_{j,k}, j \in [m], \mathbf{Z}^n] \\ & = \mathbb{P}\left[\sum_{j=1}^m d([S_j]^n, [\tilde{s}_j]^n) \leq D'_0, \mathcal{A}|\Psi_{j,k}, j \in [m], \mathbf{Z}^n\right] \end{aligned} \quad (220)$$

$$= \sum_{\mathbf{d} \in \mathcal{D}^{mn}} \prod_{j=1}^m \mathbb{P}[d([S_j]^n, [\tilde{s}_j]^n) = d_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n] \quad (221)$$

$$\leq \sum_{\mathbf{d} \in \mathcal{D}^{mn}} \prod_{j=1}^m \mathbb{P}[d(S_j^n, [\tilde{s}_j]^n) \leq d'_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n], \quad (222)$$

where  $d'_j \triangleq (\sqrt{d_j} + \Delta)^2$ , and

$$\begin{aligned} \mathcal{D}^{mn} & \triangleq \{d([\mathbf{s}]^n, [\tilde{\mathbf{s}}]^n) : [\mathbf{s}]^n, [\tilde{\mathbf{s}}]^n \in \mathcal{B}^{mn}, d([\mathbf{s}]^n, [\tilde{\mathbf{s}}]^n) \leq D'_0\} \\ \mathcal{B}^{mn} & \triangleq \{[\tilde{\mathbf{s}}]^n \in \mathcal{N}^{mn} : \|\tilde{s}_j\|^n \leq \sqrt{n\Gamma_j}, 1 \leq j \leq m\} \end{aligned}$$

with  $\sqrt{\Gamma_j} \triangleq \sqrt{\lambda_j(1+\delta)} + \Delta + \sqrt{mD'_0}$ . (221) follows from that  $\mathcal{A}_j$  implies  $\|[S_j]^n\| \leq \sqrt{n\lambda_j(1+\delta)} + \Delta$ , hence it is sufficient to only consider the case of  $\|[\tilde{s}_j]^n\| \leq \sqrt{n\Gamma_j}$ . (222) is obtained by using triangle inequality again.

Combining (218), (219) and (222) gives us

$$\begin{aligned} & \mathbb{P}_{\mathcal{CZ}^n} \left[ \max_{R_n \text{Hcodes}} \mathbb{P}[d(\mathbf{S}^n, \tilde{\mathbf{S}}^n) \leq D_0 | \mathcal{CZ}^n] > \tau_n \right] \\ & \leq |\mathcal{B}^{mn}| \max_{[\tilde{\mathbf{s}}]^n \in \mathcal{B}^{mn}} \mathbb{P}_{\mathcal{CZ}^n} \left[ \sum_{k=1}^{2^{nR_K}} \xi_{k, \mathbf{z}^n}([\tilde{\mathbf{s}}]^n) > \tau'_n 2^{-n(R_n - R_K)} \right] + \epsilon'_n, \end{aligned} \quad (223)$$

where

$$\xi_{k, \mathbf{z}^n}([\tilde{\mathbf{s}}]^n) \triangleq \sum_{\mathbf{d} \in \mathcal{D}^{mn}} \prod_{j=1}^m \mathbb{P}[d(S_j^n, [\tilde{s}_j]^n) \leq d'_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n]. \quad (224)$$

Given  $[\tilde{\mathbf{s}}]^n$  and  $\mathbf{z}^n$ ,  $\xi_{k, \mathbf{z}^n}([\tilde{\mathbf{s}}]^n), k \in [2^{nR_K}]$  are i.i.d. with mean

$$\mathbb{E}_{\mathcal{C}} \xi_{k, \mathbf{z}^n}([\tilde{\mathbf{s}}]^n) = \mathbb{E}_{\mathcal{C}} \sum_{\mathbf{d} \in \mathcal{D}^{mn}} \prod_{j=1}^m \mathbb{P}[d(S_j^n, [\tilde{s}_j]^n) \leq d'_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n] \quad (225)$$

$$= \sum_{\mathbf{d} \in \mathcal{D}^{mn}} \prod_{j=1}^m \mathbb{E}_{\Psi_{j,k}} \mathbb{P}[d(S_j^n, [\tilde{s}_j]^n) \leq d'_j, \mathcal{A}_j | \Psi_{j,k}, Z_j^n]. \quad (226)$$

To bound (223), we need to bound  $|\mathcal{B}^{mn}|$  and  $|\mathcal{D}^{mn}|$  first. Similar to (149), it can be shown

$$|\mathcal{B}^{mn}| \leq 2^{\frac{n}{2} \sum_{j=1}^m \log \frac{2\pi e(\sqrt{\Gamma_j} + \sqrt{\Delta^2})^2}{\Delta^2}} + O(\log n). \quad (227)$$

In addition, for  $[s_j]^n, [\tilde{s}_j]^n$  such that  $\|[s_j]^n\| \leq \sqrt{n\Gamma_j}$ ,  $\|[\tilde{s}_j]^n\| \leq \sqrt{n\Gamma_j}$ , using triangle inequality we have

$$d([s_j]^n, [\tilde{s}_j]^n) = \|[s_j]^n - [\tilde{s}_j]^n\|^2 \leq 4n\Gamma_j. \quad (228)$$

Combine it with

$$d([s_j]^n, [\tilde{s}_j]^n) = \frac{1}{n} \sum_{i=1}^n ([s_{i,j}] - [\tilde{s}_{i,j}])^2 = \frac{\Delta^2}{n} \sum_{i=1}^n (l_{i,j} - \tilde{l}_{i,j})^2, \quad (229)$$

where  $l_{i,j} \triangleq \text{Round}(\frac{s_{i,j}}{\Delta})$  and  $\tilde{l}_{i,j} \triangleq \text{Round}(\frac{\tilde{s}_{i,j}}{\Delta})$  are both integers, then we have

$$\sum_{i=1}^n (l_{i,j} - \tilde{l}_{i,j})^2 \leq \frac{4n^2\Gamma_j}{\Delta^2}. \quad (230)$$

In addition,  $\sum_{i=1}^n (l_{i,j} - \tilde{l}_{i,j})^2 \in \mathbb{N} \cup \{0\}$ , hence

$$|\mathcal{D}^{mn}| \leq \prod_{j=1}^m \left( \frac{4n^2\Gamma_j}{\Delta^2} + 1 \right). \quad (231)$$

That is,  $|\mathcal{D}^{mn}|$  is bounded by a polynomial term of  $n$ .

If we can show that the probability in (223) decays doubly exponentially fast with  $n$ , then the proof will be complete. To that end, by using Lemma 19 we have

$$\xi_{k,z^n}([\tilde{\mathbf{s}}]^n) \leq |\mathcal{D}^{mn}| 2^{-n(\sum_{j=1}^m R_{S_j|Z_j}(d'_j) - o(1))} \quad (232)$$

$$= 2^{-n(\sum_{j=1}^m R_{S_j|Z_j}(d'_j) - o(1))} \quad (233)$$

$$\leq 2^{-nm(R_{S|Z}(D_0 + \epsilon_\Delta) - o(1))}, \quad (234)$$

where  $R_{S|Z}(D)$  given in (52) denotes the conditional rate-distortion function for source  $\mathbf{S}$  with side information  $\mathbf{Z}$  at both encoder and decoder,  $o(1)$  is a term that vanishes as  $\delta \rightarrow 0$  and  $n \rightarrow \infty$ , and  $\epsilon_\Delta$  is a term that vanishes as  $\Delta \rightarrow 0$ . (233) follows from that  $|\mathcal{D}^{mn}|$  grows only polynomially fast with  $n$ .

From Lemma 20, we have for any  $\mathbf{z}^n$ ,

$$\mathbb{E}_C \xi_{k,z^n}([\tilde{\mathbf{s}}]^n) \leq 2^{-nm(R_S(D_0 + \epsilon_\Delta) - o(1))}, \quad (235)$$

where  $R_S(D)$  given in (51) denotes the point-to-point rate-distortion function for  $\mathbf{S}$ .

Using these bounds and applying Lemmas 12, we have that the probability in (223) decays doubly exponentially fast with  $n$ . This completes the proof of Theorem 11.

## REFERENCES

- [1] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10-21, Jan. 1949.
- [2] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 4, pp. 558-567, Oct. 1965.



- [3] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1147-1158, May 2003.
- [4] K. H. Lee, and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, no. 12, pp. 1283-1290, Dec. 1976.
- [5] S. Shamai, S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 564-579, Mar. 1998.
- [6] V. Prabhakaran, R. Puri, and K. Ramchandran, "Hybrid digital-analog codes for source-channel broadcast of Gaussian sources over Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4573-4588, Aug. 2011.
- [7] L. Yu, H. Li, and W. Li, "Wireless scalable video coding using a hybrid digital-analog scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 2, pp. 331-345, Feb. 2014.
- [8] L. Yu, H. Li, and W. Li, "Wireless cooperative video coding using a hybrid digital-analog scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 436-450, Mar. 2015.
- [9] L. Yu, H. Li, and W. Li, "Distortion bounds for source broadcast over degraded channel," *IEEE Int. Symp. Inf. Theory (ISIT)*, 2016.
- [10] L. Yu, H. Li, and W. Li, "Distortion bounds for source broadcast problem," Submitted to *IEEE Trans. Inf. Theory*, 2016.
- [11] Z. Reznicek, M. Feder, and R. Zamir, "Distortion bounds for broadcasting with bandwidth expansion," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3778-3788, Aug. 2006.
- [12] C. Tian, S. Diggavi, and S. Shamai, "Approximate characterizations for the Gaussian source broadcasting distortion region," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 124-136, Jan. 2011.
- [13] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 4, pp. 656-715, Oct. 1949.
- [14] C. Schieler, and P. Cuff, "The henchman problem: Measuring secrecy by the minimum distortion in a list," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3436-3450, Jun. 2016.
- [15] L. Yu, H. Li, and W. Li, "Source-channel secrecy for Shannon cipher system," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2596-2622, Apr. 2017.
- [16] H. Yamamoto, "Rate-distortion theory for the Shannon cipher system," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 827-835, May 1997.
- [17] M. P. Wilson, and K. Narayanan, "Transmitting an analog Gaussian source over a Gaussian wiretap channel under SNR mismatch," *IEEE International Conference on Telecommunications*, pp. 44-47, April 2010.
- [18] G. Bagherikaram and K. Plataniotis, "Secure hybrid digital-analog Wyner-Ziv coding," in *IEEE 22nd International Symposium on Personal Indoor and Mobile Radio Communications*, pp. 1161-1166, Sep. 2011.
- [19] G. Bagherikaram and K. Plataniotis, "Secure joint source-channel coding with interference known at the transmitter," *IET Communications*, vol. 6, no. 17, pp. 2796-2808, Jan. 2013.
- [20] S. C. Kak and N. S. Jayant, "On speech encryption using waveform scrambling," *Bell Syst. Tech. J.*, vol. 56, pp. 781-808, May-Jun. 1977.
- [21] A. D. Wyner, "An analog scrambling scheme which does not expand bandwidth, Part I: Discrete time," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 261-274, May 1979.
- [22] R. Ahlswede and G. Dueck, "Good codes can be produced by a few permutations," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, pp. 430-443, May 1982.
- [23] W. Kang and N. Liu, "Compressing encrypted data: Achieving optimality and strong secrecy via permutations," *IEEE Trans. on Inf. Theory*, vol. 62, no. 12, pp. 7153-7163, Dec. 2016.
- [24] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol. 53, pp. 228-236, Mar. 1965.
- [25] D. Slepian, "Group codes for the Gaussian channel," *Bell Syst. Tech. J.*, vol. 47, pp. 575-602, Apr. 1968.
- [26] T. Ericson, "A min-max theorem for antijamming group codes," *IEEE Trans. Inf. Theory*, vol. IT-30, pp. 792-799, Nov. 1984.
- [27] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian MAC," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2714-2752, Jun. 2010.
- [28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991.
- [29] I. Csiszár, and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.
- [30] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.
- [31] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*, Wiley and Sons, New York, NY, 1983.
- [32] I. Sason and S. Verdú, "f-divergence inequalities," *IEEE Trans. on Inf. Theory*, vol. 62, no. 11, pp. 5973-6006, Nov. 2016.
- [33] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611-656, 1959.

- [34] S.-W. Ho and R. W. Yeung, "On information divergence measures and a unified typicality," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 5893–5905, Dec. 2010.