Interactive Communication for Data Exchange

Himanshu Tyagi[*] Pramod Viswanath[†] Shun Watanabe[‡]

Abstract—**Two parties observing correlated data seek to exchange their data using interactive communication. How many bits must they communicate? We derive a lower bound on the minimum number of bits that is based on relating the data exchange problem to the secret key agreement problem. Furthermore, we propose an interactive protocol for data exchange which increases the communication size in steps until the task is done and matches the performance of our lower bound. Our single-shot analysis applies to all discrete random variables and yields upper and lower bound of a similar form. In fact, the bounds are asymptotically tight and lead to a characterization of the optimal rate of communication needed for data exchange for a general sequence such as mixture of IID random variables as well as the optimal second-order asymptotic term in the length of communication needed for data exchange for the IID random variables, when the probability of error is fixed. This gives a precise characterization of the asymptotic reduction in the length of optimal communication due to interaction; in particular, two-sided Slepian-Wolf compression is strictly suboptimal.**

I. Introduction

Random correlated data (X, Y) is distributed between two parties with the first observing X and the second Y. What is the optimal communication protocol for the two parties to exchange their data? We allow (randomized) interactive communication protocols and a nonzero probability of error. This basic problem was introduced by El Gamal and Orlitsky in [17] where they presented bounds on the average number of bits of communication needed by deterministic protocols for data exchange without error[1]. When interaction is not allowed, a simple solution is to apply Slepian-Wolf compression [22] for each of the two one-sided data transfer problems. The resulting protocol was shown to be of optimal rate, even in comparison with interactive protocols, when the underlying observations are *independent and identically distributed* (IID) by Csiszár and Narayan in [6]. They considered a multiterminal version of this problem, namely the problem of attaining *omniscience*, and established a lower bound on the rate of communication to show that interaction does not help in improving the asymptotic rate of communication if the probability of error vanishes to 0. However, interaction is known to be beneficial in one-sided data transfer ($cf.$ [18], [26], [7]). Can interaction help to reduce the communication needed for data exchange, and if so, what is the minimum length of interactive communication needed for data exchange?

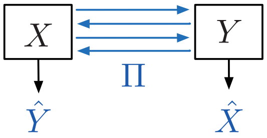

Fig. 1: The data exchange problem.

Our contributions. To address this central problem in information theory, illustrated in Figure 1, we draw on tools from cryptography, communication complexity, and the information spectrum method. Our main contribution is a new approach for proving *converse bounds* for problems with interactive communication that relates efficient communication to secret key agreement, and uses tools from cryptography such as the leftover hash lemma ($cf.$ [20]) and the recently established conditional independence testing bound for the length of a secret key [24]. Furthermore, we propose an *interactive protocol for data exchange* which matches the performance of our lower bound. In fact, our proposed protocol is a recasting in an information spectrum framework of similar protocols that appeared in [7], [26], [4], [11] in different contexts. Our modified analysis allows us to carefully choose the parameters of the protocol and in turn shows that it matches performance of our lower bound. As a consequence of the resulting single-shot bounds, we obtain a characterization of the optimal rate of communication needed for data exchange for a general sequence (X_n, Y_n) such as a mixture of IID random variables as well as the optimal second-order asymptotic term in the length of communication needed for data exchange for the IID random variables (X^n, Y^n), first instance of such a result in source coding with interactive communication[2]. This in turn leads to a precise characterization of the gain in asymptotic length of communication due to interaction.

Organization of the paper. A formal description of the data exchange problem is given in the next section. Section III gives a summary of all our results. Section IV provides a formal description of our protocol and the corresponding upper bound for the communication complexity of data exchange, and Section V contains a formal statement of our converse bound. We omit proofs due to lack of space and provide only an outline of our proof idea.

II. Problem Formulation

Let the first and the second party, respectively, observe discrete random variables X and Y taking values in finite

[*]Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore 560012, India. Email: htyagi@ece.iisc.ernet.in

[†]Coordinated Science Laboratory and Dept. of ECE, University of Illinois, Urbana-Champaign, IL 61801, USA. Email: pramodv@illinois.edu

[‡]Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan. Email: shunwata@cc.tuat.ac.jp

[1] They also illustrated the advantage of using randomized protocols when error is allowed

[2] In a different context, recently [2] showed that the second-order asymptotic term in the size of good channel codes can be improved using feedback.

sets $\mathcal{X}$ and $\mathcal{Y}$. The two parties wish to know each other's observation using interactive communication over a noiseless (error-free) channel. Specifically, the parties communicate using a communication protocol Π which consists of a sequence of random variables $\Pi_1, \Pi_2, \ldots, \Pi_r$ and a bounded stopping time[3] $T \leq r$. For each odd i, the first party sends Π_i, which is a (stochastic) function of X and the past communication $\Pi^{i-1} = (\Pi_1, \ldots, \Pi_{i-1})$. Similarly for each even i, the second party sends Π_i, which is a function of Y and the past communication. The protocol stops at round T where the event $\{T = i\}$ must be recognizable by both the parties, *i.e.*, T is a stopping time with respect to the observations of both parties. Thus, the overall transmission is given by $\Pi_1, \ldots, \Pi_T$.

Without loss of generality, we assume that each transmission Π_i takes value in $\{0,1\}^*$, the set of binary strings of varying length. The *length of a protocol* Π is the maximum accumulated number of bits transmitted in any realization of the protocol.

Definition 1. For $0 \leq \varepsilon < 1$, a protocol Π attains *ε-data exchange* (ε-DE) if there exist (stochastic) functions $\hat{Y}$ and $\hat{X}$ of (X, Π) and (Y, Π), respectively, such that

$$\mathrm{P}(\hat{X} = X,\ \hat{Y} = Y) \geq 1 - \varepsilon.$$

The *minimum communication for ε-DE $L_\varepsilon(X,Y)$* is the infimum of lengths of protocols that attain ε-DE, *i.e.*, $L_\varepsilon(X,Y)$ is the minimum number of bits that must be communicated by the two parties in order to exchange their observed data with probability of error less than ε.

Protocols with 2 rounds of communication Π_1 and Π_2, which are functions of only X and Y, respectively, are termed *simple protocols*. For a comparison, we denote by $L^{\mathrm{s}}_\varepsilon(X,Y)$ the minimum communication for ε-DE by a simple protocol.

III. Summary of Results

To describe our results, denote by $h(X) = -\log \mathrm{P}_X(X)$ and $h(X|Y) = -\log \mathrm{P}_{X|Y}(X|Y)$, respectively, the *entropy density* of X and the *conditional entropy density* of X given Y. Also, pivotal in our results is a quantity we call the *sum conditional entropy density* of X and Y defined as

$$h(X \triangle Y) := h(X|Y) + h(Y|X).$$

An interactive data exchange protocol. Our data exchange protocol is based on an interactive version of the Slepian-Wolf protocol where the length of the communication is increased in steps until the second party decodes the data of the first. Similar protocols have been proposed earlier for distributed data compression in [7], [26], for protocol simulation in [4], and for secret key agreement in [12], [11].

In order to send X to an observer of Y, a single-shot version of the Slepian-Wolf protocol was proposed in [16] (see, also, [8, Lemma 7.2.1]). Roughly speaking, this protocol simply hashes X to as many bits as the right most point in the spectrum[4] of $\mathrm{P}_{X|Y}$. The main shortcoming of this protocol for our purpose is that it sends the same number of bits for every realization of (X,Y). However, we would like to use as few bits as possible for sending X to party 2 so that the remaining bits can be used for sending Y to party 1. Note that once X is recovered by party 2 correctly, it can send Y to Party 1 without error using, say, Shannon-Fano-Elias coding (eg. see [5, Section 5]); the length of this second communication is $\lceil h(Y|X) \rceil$ bits. Our protocol accomplishes the first part above using roughly $h(X|Y)$ bits of communication.

Specifically, in order to send X to Y we use a *spectrum slicing technique* introduced in [8] (see, also, [12], [11]). We divide the domain $[\lambda_{\min}, \lambda_{\max}]$ of spectrum of $\mathrm{P}_{X|Y}$ into N slices size Δ each; see Figure 2 for an illustration.

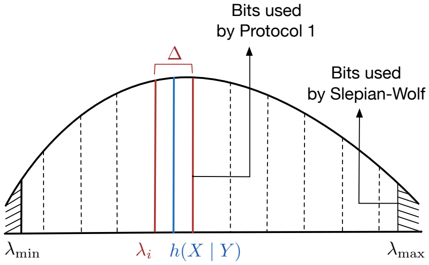

Fig. 2: Spectrum slicing in Protocol 1.

The protocol begins with the left most slice and party 1 sends $\lambda_{\min} + \Delta$ hash bits to party 2. If party 2 can find a unique x that is compatible with the received hash bits, it sends back an ACK and the protocol stops. Else, party 2 sends back a NACK and the protocol now moves to the next round, in which Party 1 sends additional Δ hash bits. The parties keep on moving to the next slice until either party 2 sends an ACK or all slices are covered. It is easy to show that this protocol is reliable and uses not more than $h(X|Y) + \Delta + N$ bits of communication for each realization of (X,Y). As mentioned above, once party 2 gets X, it sends back Y using $h(Y|X) + 1$ bits, thereby resulting in an overall communication of $h(X \triangle Y) + 1$ bits. In our applications, we shall choose N and Δ to be of negligible order in comparison with the tail bounds on $h(X \triangle Y)$. Thus, we have the following upper bound on $L_\varepsilon(X,Y)$. (The statement here is rough; see Theorem 2 below for a precise version.)

Result 1 (Rough statement of the single-shot upper bound). *For every* $0 < \varepsilon < 1$,

$$L_\varepsilon(X,Y) \lesssim \inf\{\gamma : \mathbb{P}(h(X \triangle Y) > \gamma) \leq \varepsilon\}.$$

A converse bound. Our next result, which is perhaps the main contribution of this paper, is a lower bound on $L_\varepsilon(X,Y)$. This bound is derived by connecting the data exchange problem to the two-party secret key agreement problem. For an illustration of our approach in the case of IID random variables X^n and Y^n, note that the optimal rate of a secret key that can be generated is given by $I(X \wedge Y)$, the mutual information between X and Y [15], [1]. Also, using a privacy amplification

[3] The bounded assumption is cosmetic and is made only for concreteness. Our results remain valid even when this assumption is dropped.

[4] Spectrum of a distribution P_X refers to the histogram of $-\log \mathrm{P}_X$.

argument (*cf.* [3], [20]), it can be shown that a data exchange protocol using nR bits can yield roughly $n(H(XY) - R)$ bits of secret key. Therefore, $I(X \wedge Y)$ exceeds $H(XY) - R$, which further gives $R \geq H(X|Y) + H(Y|X)$. This connection between secret key agreement and data exchange was noted first in [6] where it was used for designing an optimal rate secret key agreement protocol. Our converse proof is, in effect, a single-shot version of this argument.

Specifically, the "excess" randomness generated when the parties observing X and Y share a communication Π can be extracted as a secret key independent of Π using the *leftover hash lemma* [14], [21]. Thus, denoting by $S_\varepsilon(X, Y)$ the maximum length of secret key and by H the length of the common randomness (*cf.* [1]) generated by the two parties during the protocol, we get $H - L_\varepsilon(X, Y) \leq S_\varepsilon(X, Y)$.

Next, we apply the recently established conditional independence testing upper bound for $S_\varepsilon(X, Y)$ [24], [25], which follows by reducing a binary hypothesis testing problem to secret key agreement. However, the resulting lower bound on $L_\varepsilon(X, Y)$ is good only when the spectrum of P_{XY} is concentrated. Heuristically, this slack in the lower bound arises since we are lower bounding the worst-case communication complexity of the protocol for data exchange – the resulting lower bound need not apply for every (X, Y) but only for a few realizations of (X, Y) with probability greater than ε. To remedy this shortcoming, we once again take recourse to spectrum slicing and show that there exists a slice of the spectrum of P_{XY} where the protocol requires sufficiently large number of bits; Figure 3 illustrates this approach. The resulting lower bound on $L_\varepsilon(X, Y)$ is stated below roughly, and a precise statement is given in Theorem 3.

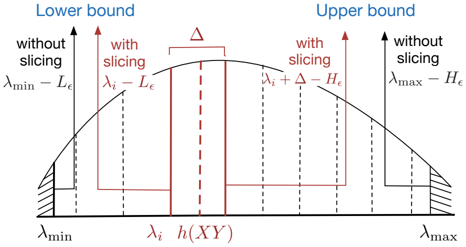

Fig. 3: Bounds on secret key length leading to the converse. Here $L_\varepsilon := L_\varepsilon(X, Y)$ and H_ε denotes the ε-tail of $h(X \triangle Y)$.

Result 2 (Rough statement of the single-shot lower bound). *For every* $0 < \varepsilon < 1$,

$$L_\varepsilon(X, Y) \gtrsim \inf\{\gamma : \mathbb{P}\left(h(X \triangle Y) > \gamma\right) \leq \varepsilon\}.$$

Note that the upper and lower bound on $L_\varepsilon(X, Y)$ appear to be same in the two results above only because we have ignore some error terms. Nevertheless, the form above captures the spirit of our bounds on $L_\varepsilon(X, Y)$. In fact, the displayed term dominates asymptotically and leads to tight bounds in the asympotitic regime.

Asymptotic optimality. The single-shot bounds stated above are asymptotically tight up to the first order term for any sequence of random variables (X_n, Y_n), and up to the second order term for a sequence of IID random variables (X^n, Y^n).

Specifically, consider a general source sequence $(\mathbf{X}, \mathbf{Y}) = \{(X_n, Y_n)\}_{n=1}^\infty$. We are interested in characterizing the minimum asymptotic rate of communication for asymptotically error-free data exchange, and seek its comparison with the minimum rate possible using simple protocols.

Definition 2. The minimum rate of communication for data exchange R^* is defined as

$$R^*(\mathbf{X}, \mathbf{Y}) = \inf_{\varepsilon_n} \limsup_{n \to \infty} \frac{1}{n} L_{\varepsilon_n}(X_n, Y_n),$$

where the infimum is over all $\varepsilon_n \to 0$ as $n \to \infty$. The corresponding minimum rate for simple protocols is denoted by R_s^*.

Denote by $\overline{H}(\mathbf{X} \triangle \mathbf{Y})$, $\overline{H}(\mathbf{X}|\mathbf{Y})$, and $\overline{H}(\mathbf{Y}|\mathbf{X})$, respectively, the $\limsup$ in probability of random variables $h(X_n \triangle Y_n)$, $h(X_n|Y_n)$, and $h(Y_n|X_n)$. The quantity $\overline{H}(\mathbf{X}|\mathbf{Y})$ is standard in information spectrum method [9], [8] and corresponds to the asymptotically minimum rate of communication needed to send X_n to an observer of Y_n [16] (see, also, [8, Lemma 7.2.1]). Thus, a simple communication protocol of rate $\overline{H}(\mathbf{X}|\mathbf{Y}) + \overline{H}(\mathbf{Y}|\mathbf{X})$ can be used to accomplish data exchange. In fact, a standard converse argument can be used to show the optimality of this rate. Therefore, when we restrict ourselves to simple protocols, the asymptotically minimum rate of communication needed is

$$R_s^*(\mathbf{X}, \mathbf{Y}) = \overline{H}(\mathbf{X}|\mathbf{Y}) + \overline{H}(\mathbf{Y}|\mathbf{X}).$$

As an illustration, consider the case when (X_n, Y_n) are generated by a mixture of two n-fold IID distributions $\mathrm{P}^{(1)}_{X^nY^n}$ and $\mathrm{P}^{(2)}_{X^nY^n}$. For this case, the right-side above equals

$$\max\{H(X^{(1)} \mid Y^{(1)}), H(X^{(2)} \mid Y^{(2)})\} + \max\{H(Y^{(1)} \mid X^{(1)}), H(Y^{(2)} \mid X^{(2)})\}.$$

Can we improve this rate by using interactive communication? Using our single-shot bounds for $L_\varepsilon(X, Y)$, we answer this question in the affirmative.

Result 3 (Min rate of communication for data exchange). *For a sequence of sources* $(\mathbf{X}, \mathbf{Y}) = \{(X_n, Y_n)\}_{n=1}^\infty$,

$$R^*(\mathbf{X}, \mathbf{Y}) = \overline{H}(\mathbf{X} \triangle \mathbf{Y}).$$

For the mixture of IID example above,

$$\overline{H}(\mathbf{X} \triangle \mathbf{Y}) = \max\{H(X^{(1)} \mid Y^{(1)}) + H(Y^{(1)} \mid X^{(1)}), H(X^{(2)} \mid Y^{(2)}) + H(Y^{(2)} \mid X^{(2)})\},$$

and therefore, simple protocols are strictly suboptimal in general. Note that while the standard information spectrum techniques suffice to prove the converse when we restrict to simple protocols, their extension to interactive protocols is unclear and our single-shot converse above is needed.

Turning now to the case of IID random variables, *i.e.* when $X_n = X^n = (X_1, ..., X_n)$ and $Y_n = Y^n = (Y_1, ..., Y_n)$ are

n-IID repetitions of random variables (X, Y). For brevity, denote by $R^*(X, Y)$ the corresponding minimum rate of communication for data exchange, and by $H(X\triangle Y)$ and V, respectively, the mean and the variance of $h(X\triangle Y)$. Earlier, Csiszár and Narayan [6] showed that $R^*(X, Y) = H(X\triangle Y)$. We are interested in a finer asymptotic analysis than this first order characterization.

In particular, we are interested in characterizing the asymptotic behavior of $L_\varepsilon(X^n, Y^n)$ up to to the second-order term, for every fixed ε in (0,1). We need the following notation:

$$R^*_\varepsilon(X, Y) = \lim_{n\to\infty} \frac{1}{n} L_\varepsilon(X^n, Y^n), \quad 0 < \varepsilon < 1.$$

Note that $R^*(X, Y) = \sup_{\varepsilon\in(0,1)} R^*_\varepsilon(X, Y)$. Our next result shows that $R^*_\varepsilon(X, Y)$ does not depend on ε and constitutes a *strong converse* for the result in [6].

Result 4 (**Strong converse**). *For every* $0 < \varepsilon < 1$,

$$R^*_\varepsilon(X, Y) = H(X\triangle Y).$$

In fact, this result follows from a general result characterizing the second-order asymptotic term[5].

Result 5 (**Second-order asymptotic behavior**). *For every* $0 < \varepsilon < 1$,

$$L_\varepsilon\left(X^n, Y^n\right) = nH(X\triangle Y) + \sqrt{nV}Q^{-1}(\varepsilon) + o(\sqrt{n}),$$

where $Q(a)$ *is the tail probability of the standard Gaussian distribution.*

IV. A General Achievability Scheme

We begin with an interactive scheme for sending X to an observer of Y, which hashes (bins) X into a few values as in the scheme of [16], but unlike that scheme, increases the hash-size gradually, starting with $\lambda_1 = \lambda_{\min}$ and increasing the size Δ-bits at a time until either X is recovered or $\lambda_{\max}$ bits have been sent. After each transmission, Party 2 sends either an ACK-NACK feedback signal; the protocol stops when an ACK symbol is received.

As mentioned in the introduction, we rely on spectrum slicing. Our protocol focuses on the "essential spectrum" of $h(X|Y)$, *i.e.*, those values of (X, Y) for which $h(X|Y) \in (\lambda_{\min}, \lambda_{\max})$. For $\lambda_{\min}, \lambda_{\max}, \Delta > 0$ with $\lambda_{\max} > \lambda_{\min}$, let

$$N = \frac{\lambda_{\max} - \lambda_{\min}}{\Delta}, \tag{1}$$

and $\lambda_i = \lambda_{\min} + (i-1)\Delta, \quad 1 \le i \le N$. Further, let

$$\mathcal{T}_0 = \Big\{h(X|Y) \ge \lambda_{\max} \text{ or } h(Y|X) < \lambda_{\min}\Big\}, \tag{2}$$

and for $1 \le i \le N$, let $\mathcal{T}_i$ denote the ith slice of the spectrum given by $\mathcal{T}_i = \Big\{(x, y) : \lambda_i \le h_{\mathrm{P}_{X|Y}}(x|y) < \lambda_i + \Delta\Big\}$. Note that $\mathcal{T}_0$ corresponds to the complement of "typical event." Finally, let $\mathcal{H}_l(\mathcal{X})$ denote the set of all mappings $h : \mathcal{X} \to \{0, 1\}^l$.

Our protocol for transmitting X to an observer of Y is described in Protocol 1. The lemma below bounds the probability of error for Protocol 1 when $(x, y) \in \mathcal{T}_i$, $1 \le i \le N$.

Protocol 1: Interactive Slepian-Wolf compression

```
Input: Observations $X$ and $Y$, uniform public
       randomness $V$, and parameters $l, \Delta$
Output: Estimate $\hat{X}$ of $X$ at party 2
Both parties use $V$ to select $h_1$ uniformly from $\mathcal{H}_l(\mathcal{X})$
Party 1 sends $\Pi_1 = h_1(X)$
if Party 2 finds a unique $x \in \mathcal{T}_1$ with hash value
$h_1(x) = \Pi_1$ then
    set $\hat{X} = x$
    send back $\Pi_2$ = ACK
else
    send back $\Pi_2$ = NACK
while $2 \le i \le N$ and party 2 did not send an ACK do
    Both parties use $V$ to select $h_i$ uniformly from
    $\mathcal{H}_\Delta(\mathcal{X})$, independent of $h_1, ..., h_{i-1}$
    Party 1 sends $\Pi_{2i-1} = h_i(X)$
    if Party 2 finds a unique $x \in \mathcal{T}_i$ with hash value
    $h_j(x) = \Pi_{2j-1}, \forall 1 \le j \le i$ then
        set $\hat{X} = x$
        send back $\Pi_{2i}$ = ACK
    else
        if More than one such $x$ found then
            protocol declares an error
        else
            send back $\Pi_{2i}$ = NACK
    Reset $i \to i + 1$
if No $\hat{X}$ found at party 2 then
    Protocol declares an error
```

Theorem 1 (**Interactive Slepian-Wolf**). *Protocol 1 with* $l = \lambda_{\min} + \Delta + \eta$ *sends at most* $(h\left(X|Y\right) + \Delta + N + \eta)$ *bits when the observations are* $(X, Y) \notin \mathcal{T}_0$ *and has probability of error less than*

$$\mathbb{P}\left(\hat{X} \ne X\right) \le \mathrm{P}_{XY}\left(\mathcal{T}_0\right) + N2^{-\eta}.$$

Note that when $\mathcal{T}_0$ is chosen to be of small probability, Protocol 1 sends essentially the same number of bits in the worst-case as the Slepian-Wolf protocol.

Returning to the data exchange problem, our protocol for data exchange builds upon Protocol 1 and uses it to first transmit X to the second party (observing Y). Once Party 2 has recovered X correctly, it sends Y to Party 1 without error using, say, Shannon-Fano-Elias coding (eg. see [5, Section 5]); the length of this second communication is $\lceil h(Y|X) \rceil$ bits. When the accumulated number of bits communicated in the protocol exceeds a prescribed length $l_{\max}$, the parties abort the protocol and declare an error.[6] Using Theorem 1,

[5] Following the pioneering work of Strassen [23], study of these second-order terms in coding theorems has been revived recently by Hayashi [10], [13] and Polyanskiy, Poor, and Verdú [19].

[6] Alternatively, we can use the (noninteractive) Slepian-Wolf coding by setting the size of hash as $l_{\max} - (h(X|Y) + \Delta + N + \eta)$.

the probability of error of the combined protocol is bounded above as follows.

Theorem 2 (**Interactive data exchange protocol**). *Given $\lambda_{\min}, \lambda_{\max}, \Delta, \eta > 0$ and for N in (1), there exists a protocol for data exchange of length $l_{\max}$ such that*

$$\begin{aligned}&\mathbb{P}\left(X \neq \hat{X} \textit{ or } Y \neq \hat{Y}\right)\\ &\quad\le \mathbb{P}\left(h(X\triangle Y) + \Delta + N + \eta + 1 > l_{\max}\right)\\ &\qquad + \mathrm{P}_{XY}\left(\mathcal{T}_0\right) + N2^{-\eta}.\end{aligned}$$

Thus, we attain ε-DE using a protocol of length $l_{\max} = \lambda_\varepsilon + \Delta + N$, where λ_ε is the ε-tail of $h(X\triangle Y)$.

V. Converse Bound

Our converse bound, while heuristically simple, is technically involved. Unfortunately, due to lack of space we can only provide a sketch here.

Our converse proof, too, relies on spectrum slicing to find the part of the spectrum of P_{XY} where the protocol communicates large number of bits. As in the achievability part, we shall focus on the "essential spectrum" of $h(XY)$.

Given $\lambda_{\max}$, $\lambda_{\min}$, and $\Delta > 0$, let N be as in (1) and the set $\mathcal{T}_0$ be as in (2), with $h_{\mathrm{P}_{X|Y}}(x|y)$ replaced by $h_{\mathrm{P}_{XY}}(xy)$ in those definitions.

Theorem 3. *For $0 \le \varepsilon < 1$, $0 < \eta < 1 - \varepsilon$, and parameters Δ, N as above, the following lower bound on $L_\varepsilon(X,Y)$ holds for every $\gamma > 0$:*

$$\begin{aligned}L_\varepsilon(X,Y) \ge \gamma + 3\log\left(\mathrm{P}_\gamma - \varepsilon - \mathrm{P}_{XY}\left(\mathcal{T}_0\right) - \frac{1}{N}\right)_+&\\ + \log(1-2\eta) - \Delta - 6\log N - 4\log\frac{1}{\eta} - 1&,\end{aligned}$$

where $\mathrm{P}_\gamma := \mathrm{P}_{XY}\left(h(X\triangle Y) > \gamma\right)$.

Thus, a protocol attaining ε-DE must communicate roughly as many bits as ε-tail of $h(X\triangle Y)$.

We close with an outline of our proof. The main idea is to relate data exchange to secret key agreement, which is done in the following two steps:

1) Given a protocol Π for ε-DE of length l, use the leftover hash lemma to extract an ε-secret key of length roughly $\lambda_{\min} - l$.
2) The length of the secret key that has been generated is bounded above by $S_\varepsilon(X,Y)$, the maximum possible length of an ε-secret key. Use the conditional independence testing bound in [24], [25] to further upper bound $S_\varepsilon(X,Y)$, thereby obtaining a lower bound for l.

This approach leads to a loss of $\lambda_{\max} - \lambda_{\min}$, the length of the spectrum of P_{XY}. However, since we are lower bounding the worse-case communication complexity, we can divide the spectrum into small slices of length Δ, and show that there is a slice where the communication is high enough by applying the steps above to the conditional distribution given that (X,Y) lie in a given slice. This reduces the loss from $\lambda_{\max} - \lambda_{\min}$ to Δ.

References

[1] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography–part i: Secret sharing," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.

[2] Y. Altuğ and A. B. Wagner, "Feedback can improve the second-order coding performance in discrete memoryless channels," *ISIT*, 2014.

[3] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, "Generalized privacy amplification," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, November 1995.

[4] M. Braverman and A. Rao, "Information equals amortized communication," in *FOCS*, 2011, pp. 748–757.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory. 2nd edition.* John Wiley & Sons Inc., 2006.

[6] I. Csiszár and P. Narayan, "Secrecy capacities for multiple terminals," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3047–3061, December 2004.

[7] M. Feder and N. Shulman, "Source broadcasting with unknown amount of receiver side information," in *ITW*, Oct 2002, pp. 127–130.

[8] T. S. Han, *Information-Spectrum Methods in Information Theory [English Translation].* Springer, 2003.

[9] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[10] M. Hayashi, "Second-order asymptotics in fixed-length source coding and intrinsic randomness," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4619–4637, Oct 2008.

[11] M. Hayashi, H. Tyagi, and S. Watanabe, "Secret key agreement: General capacity and second-order asymptotics," *arXiv:1411.0735*, 2014.

[12] ——, "Secret key agreement: General capacity and second-order asymptotics," *ISIT*, 2014.

[13] M. Hayashi, "Information spectrum approach to second-order coding rate in channel coding," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4947–4966, Novemeber 2009.

[14] R. Impagliazzo, L. A. Levin, and M. Luby, "Pseudo-random generation from one-way functions," *Proc. Annual Symposium on Theory of Computing*, pp. 12–24, 1989.

[15] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[16] S. Miyake and F. Kanaya, "Coding theorems on correlated general sources," *IIEICE Trans. Fundamental*, vol. E78-A, no. 9, pp. 1063–1070, September 1995.

[17] A. Orlitsky and A. El Gamal, "Communication with secrecy constraints," *STOC*, pp. 217–224, 1984.

[18] A. Orlitsky, "Worst-case interactive communication i: Two messages are almost optimal," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1111–1126, 1990.

[19] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[20] R. Renner, "Security of quantum key distribution," *Ph. D. Dissertation, ETH Zurich*, 2005.

[21] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *ASIACRYPT*, 2005, pp. 199–216.

[22] D. Slepian and J. Wolf, "Noiseless coding of correlated information source," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[23] V. Strassen, "Asymptotische abschätzungen in Shannon's informationstheorie," *Third Prague Conf. Inf. Theory*, pp. 689–723, 1962, English translation: http://www.math.cornell.edu/ pmlut/strassen.pdf.

[24] H. Tyagi and S. Watanabe, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *EUROCRYPT*, 2014, pp. 369–386.

[25] ——, "Converses for secret key agreement and secure computing," *CoRR*, vol. abs/1404.5715, 2014.

[26] E.-H. Yang and D.-K. He, "Interactive encoding and decoding for one way learning: Near lossless recovery with side information at the decoder," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1808–1824, April 2010.