Randomized Distributed Function Computation with Semantic Communications: Applications to Privacy

Onur Günlü
Information Theory and Security Laboratory (ITSL),
Linköping University, Linköping, Sweden,
onur.gunlu@liu.se

***Abstract*—Randomized distributed function computation refers to remote function computation where transmitters send data to receivers which compute function outputs that are randomized functions of the inputs. We study the applications of semantic communications in randomized distributed function computation to illustrate significant reductions in the communication load, with a particular focus on privacy. The semantic communication framework leverages generalized remote source coding methods, where the remote source is a randomized version of the observed data. Since satisfying security and privacy constraints generally require a randomization step, semantic communication methods can be applied to such function computation problems, where the goal is to remotely simulate a sequence at the receiver such that the transmitter and receiver sequences follow a target probability distribution. Our performance metrics guarantee (local differential) privacy for each input sequence, used in two different distributed function computation problems, which is possible by using strong coordination methods.**

This work provides lower bounds on Wyner's common information (WCI), which is one of the two corner points of the coordination-randomness rate region characterizing the ultimate limits of randomized distributed function computation. The WCI corresponds to the case when there is no common randomness shared by the transmitter and receiver. Moreover, numerical methods are proposed to compute the other corner point for continuous-valued random variables, for which an unlimited amount of common randomness is available. Results for two problems of practical interest illustrate that leveraging common randomness can decrease the communication load as compared to the WCI corner point significantly. We also illustrate that semantic communication gains over lossless compression methods are achieved also without common randomness, motivating further research on limited common randomness scenarios.

***Index Terms*—Strong coordination, efficient simulation of noisy channels over digital channels, semantic communications for randomized distributed function computation.**

I. Introduction

The problem of reliably transmitting a desired meaning, rather than bit sequences, has led to the development of the *semantic communication* framework [1], [2]. Semantic communications refer to transmission of a signal under a measure of quality that depends on the semantics [3]. Physical layer operations in the current communication systems are not adaptive to the content of the delivered information. Contrarily, semantic communications aim to only transmit the relevant content. For instance, transmission of objects and their relative positions in an image, rather than the pixel values, is an example for semantic communication [3]. Since the semantics of the data available at the transmitter can be seen as a function of the data, the semantic communication problem can be modeled as a remote (or hidden) source coding problem, where the conveyed information is not the data itself but a function of it, while the function output is not available to the transmitter [4], [5, pp. 118], [6, pp. 78]. Such a problem description is a generalization of the lossy source coding problem, as the fidelity measure can be chosen based on the semantics.

We consider a generalization of the remote source coding problem, where a randomized function of the available data is used for distributed function computation. The function computed at the receiver takes randomized functions of the data available at the transmitters as an argument. Thus, this general framework can be considered as a *randomized distributed function computation framework with semantic communications*. Example cases where randomization is necessary include problems with security and privacy constraints, as randomization is usually necessary to provide security and privacy guarantees [7], [8]. We consider randomized distributed function computation problems under a privacy constraint, where a transmitter aims to remotely simulate a sequence at the receiver such that the transmitter and receiver sequences follow a target correlation, i.e., they are distributed according to a given joint probability distribution that provides privacy. Remotely simulating sequences with a minimum amount of communications between the nodes is considered in the literature as a *coordination problem* [9]–[12], alternative names including distributed channel simulation or synthesis. Applications of signal processing methods to achieve a weaker form of coordination, called empirical coordination, have recently received increased attention due to their applications in machine learning problems, see, e.g., [13]–[16]. We, however, consider strong coordination measures [9] to provide coordination guarantees for each instance (rather than for the average behavior over multiple instances) by imposing a joint typicality constraint, which is relevant particularly for security and privacy applications to avoid enabling new attacks.

This work was supported in part by the ZENITH Research and Leadership Career Development Fund, Chalmers Transport Area of Advance, and the ELLIIT funding.

Unlike deterministic function computation problems, common randomness shared by the transmitter and receiver is shown to significantly reduce the communication load required for randomized function computation problems [17]. For a point-to-point channel model, there are two corner points of the coordination-randomness rate region. If there is no common randomness, the minimum communication rate corresponds to the value of Wyner's common information (WCI) $C(\widetilde{X};Y) = \inf_{U:\widetilde{X}-U-Y} I(\widetilde{X},Y;U)$ between the channel input $\widetilde{X}$ and output Y, where $\widetilde{X}-U-Y$ forms a Markov chain [18]. If there is enough common randomness, the minimum communication rate can be reduced to the mutual information $I(\widetilde{X};Y)$. In this work, we compare the WCI and the mutual information $I(\widetilde{X};Y)$ for two problems of practical interest that require randomized distributed function computation. To this end, we impose symmetry conditions to establish lowers bound on the WCI and to compute the mutual information $I(\widetilde{X};Y)$ numerically. For the first problem, we consider a local differential privacy (LDP) constraint imposed to ensure that the randomized (i.e., privatized) output only leaks a limited amount of individual information. For the second problem, we consider a random response mechanism, which can be considered to combine multiple randomized bit responses.

In Section II, we introduce the randomized distributed function computation problem that provides function computation guarantees for each compute instance, as well as the metric for achieving LDP. In Section III, we establish a lower bound on the WCI and a numerical method to compute the mutual information $I(\widetilde{X};Y)$ for continuous-valued random variables when a LDP constraint is imposed. In Section IV, we consider a random response mechanism and establish a lower bound on the WCI for a set of discrete symmetric channels.

II. Preliminaries

A. Randomized Distributed Function Computation

The randomized distributed function computation framework [9], with the strong coordination measures, aims to use channels to coordinate sequences of multiple nodes in a network by transmitting the minimum amount of information over channels, which does not require reliable communication of local sequences to other nodes. Coordination is established by exhibiting a joint behavior in local sequences summarized by a target joint probability distribution, which depends on the goals as well as computation and communication resources of each node.

Consider the two-user randomized distributed function computation problem depicted in Fig. 1. The transmitter observes a sequence $\widetilde{X}^n = \widetilde{x}^n$ and jointly encodes $\widetilde{x}^n$ and common randomness $C \in \{1,2,\ldots,2^{nR_0}\}$ shared between the transmitter and receiver, the latter of which can be obtained by using physical unclonable functions [19]. The output of the encoding operation is an index $S \in \{1,2,\ldots,2^{nR}\}$, which is transmitted through a noiseless link to the receiver which wants to synthesize a sequence $Y^n = y^n$ using S and C such that $(\widetilde{x}^n, y^n) \sim Q^n_{\widetilde{X}Y}$. We next define the coordination-randomness region for the two-user randomized distributed function computation problem depicted in Fig. 1.

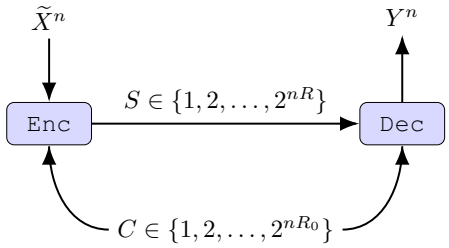

Fig. 1: A randomized distributed function computation model with two nodes and common randomness C that might be available to them. The strong coordination framework imposes that the receiver node outputs a sequence Y^n, given transmitted index S and common randomness C, such that $(\widetilde{x}^n, y^n)$ pair follows a joint probability distribution $Q_{\widetilde{X}Y}$, which represents a cooperation strategy that benefits both nodes and is imposed as in Eq. (1) given below. The rate R of information to be transmitted over channels should be minimized, which broadly corresponds to minimizing the function computation latency and energy consumption.

Definition 1: A coordination-randomness rate pair (R, R_0) is achievable for $Q_{\widetilde{X}Y}$ if, for any $\epsilon_n > 0$, there exist $n \geq 1$ and an encoder-decoder pair such that

$$\lim_{n\to\infty} \left\| P_{\widetilde{X}^n Y^n} - \prod_{i=1}^{n} Q_{\widetilde{X}Y} \right\|_{\text{TV}} = 0 \tag{1}$$

where $P_{\widetilde{X}^n Y^n}$ is the probability distribution induced by the encoder-decoder pair and $\|\cdot\|_{\text{TV}}$ is the total variation distance.

The coordination-randomness region $\mathcal{R}_{\text{RDFC}}$ is the closure of the set of all achievable tuples for the coordination problem depicted in Fig. 1. $\Diamond$

We next provide the coordination-randomness region $\mathcal{R}_{\text{RDFC}}$.

Theorem 1 ([9, Theorem II.1]): For the coordination problem depicted in Fig. 1, the coordination-randomness region $\mathcal{R}_{\text{RDFC}}$ is the union over all probability distributions $P_{\widetilde{X}UY} = P_U P_{\widetilde{X}|U} P_{Y|U}$ of the coordination-randomness rate pairs (R, R_0) satisfying

$$R \geq I(\widetilde{X};U), \tag{2}$$

$$R + R_0 \geq I(\widetilde{X},Y;U), \tag{3}$$

$$\sum_{u\in\mathcal{U}} P_{\widetilde{X}UY} = Q_{\widetilde{X}Y}. \tag{4}$$

The coordination-randomness region given in Theorem 1 remains valid for continuous-valued random variables considered in this work, since there is a discretization procedure [20, Remark 3.8] to generalize the achievability proof to well-behaved continuous-alphabet random variables, such as the (truncated) Gaussian distributions considered below, see also [9, Section VII].

There are two corner points of the coordination-randomness region given in Theorem 1.

First, assume that $R_0 = 0$, i.e., there is no common randomness shared between the transmitter and receiver. Then,

by (2) and (3), we obtain the corner point

$$R \geq C(\widetilde{X};Y) = \inf_{U:\ \widetilde{X}-U-Y} I(\widetilde{X},Y;U) \tag{5}$$

where $C(\widetilde{X};Y)$ is the WCI between the input $\widetilde{X}$ and Y [18]. We remark that in common multi-user information theory results, auxiliary random variables V form Markov chains of the form $V-\widetilde{X}-Y$, for which there are multiple methods to optimize their probability distributions. However, in (5), the auxiliary random variable U forms a Markov chain of the form $\widetilde{X}-U-Y$, which makes it difficult to optimize $P_{\widetilde{X}UY}$ [21]. We illustrate below lower bounds on the WCI for symmetric distributions that are of practical relevance.

Second, assume that there is sufficient common randomness C such that (2) and (3) result in the corner point

$$R \geq I(\widetilde{X};Y) \tag{6}$$

which follows since $I(\widetilde{X};U) \geq I(\widetilde{X};Y)$ due to the data processing inequality. The corner point in (6) can be achieved by $R_0 \geq H(Y|X)$ although a smaller common randomness rate is shown in [22] to be sufficient to achieve it for some cases. The bound in (6) coincides with the reverse Shannon theorem [10] which uses noiseless channels to synthesize a noisy one with the help of an unlimited amount of common randomness shared between the transmitter and receiver. This operation is the reverse of the operation in the channel coding theorem [23].

We remark the following relations [18]

$$I(\widetilde{X};Y) \leq C(\widetilde{X};Y) \leq \min\{H(\widetilde{X}),\ H(Y)\} \tag{7}$$

which clearly illustrate that randomized distributed function computation can bring significant gains in terms of the amount of communication (even without any common randomness) over both (i) transmitting the observed sequence $\widetilde{X}^n$ with rate $H(\widetilde{X})$ to the receiver losslessly and then computing the randomized function output Y^n at the receiver, and (ii) computing the randomized function output Y^n at the transmitter and then transmitting it with rate $H(Y)$ to the receiver losslessly.

B. Local Differential Privacy

We consider LDP to guarantee individual user privacy, which is a randomized distributed function computation problem since privacy is provided by computing a randomized function of the input.

Now, we define (ϵ,δ)–LDP, also called approximate LDP as $\delta \neq 0$. Define $\mathcal{M}(\cdot)$ as a randomized mechanism that maps each $\widetilde{x} \in \widetilde{\mathcal{X}}$ to a probability distribution $\mathcal{M}(\cdot|\widetilde{x}) \in \mathcal{P}(\mathcal{Y})$.

Definition 2 ([24]–[26]): A mechanism $\mathcal{M}: \widetilde{\mathcal{X}} \to \mathcal{P}(\mathcal{Y})$ is (ϵ,δ)-LDP for $\delta \in [0,1]$ and $\epsilon \geq 0$ if we have

$$\sup_{\widetilde{x},\widetilde{x}' \in \widetilde{\mathcal{X}}} \sup_{Y \subseteq \mathcal{Y}} \big(\mathcal{M}(Y|\widetilde{x}) - e^{\epsilon}\mathcal{M}(Y|\widetilde{x}')\big) \leq \delta. \tag{8}$$

$\Diamond$

In the next section, we consider a Gaussian LDP mechanism $\mathcal{M}(\cdot)$ that adds an independent Gaussian noise to satisfy an (ϵ,δ)-LDP constraint, which is formulated as a randomized distributed function computation problem. The Gaussian mechanism, as a low-complexity DP method, can be used, for instance, in private empirical risk minimization methods that are based on, e.g., stochastic gradient descent or Adam optimizer [27].

III. Coordination for Local Differential Privacy with Additive Gaussian Noise

Consider an independent and identically distributed (i.i.d.) Gaussian random variable $X^n \sim \mathcal{N}^n(0,\sigma_x^2)$ that is clipped to the range $[-C,C]$, where $C>0$. Define, for $a \in R$,

$$\beta = \frac{C}{\sigma_X}, \qquad \text{erf}(a) = \frac{2}{\sqrt{\pi}}\int_0^a e^{-t^2}dt, \tag{9}$$

$$\phi(a) = \frac{1}{\sqrt{2\pi}}e^{-a^2/2}, \qquad \gamma(a) = \frac{a\phi(a)}{\text{erf}\big(\frac{a}{\sqrt{2}}\big)}. \tag{10}$$

The clipped output $\widetilde{X}^n$ is an i.i.d. truncated Gaussian random variable with zero mean and variance

$$\sigma_{\widetilde{X}}^2 = \sigma_X^2\Big(1-2\gamma(\beta)\Big). \tag{11}$$

Observing $\widetilde{X}^n$, the transmitter aims to achieve (ϵ,δ)-LDP by considering the Gaussian LDP mechanism, i.e., the sequence Y^n observed at the receiver is such that

$$\{(Y^n-\widetilde{X}^n)|\widetilde{X}^n\} := \{\widetilde{Z}^n|\widetilde{X}^n\} = \widetilde{Z}^n \sim \mathcal{N}^n(0,\sigma_{\widetilde{Z}}^2). \tag{12}$$

If we have

$$\sigma_{\widetilde{Z}}^2 = \frac{8C^2}{\epsilon^2}\log\big(\frac{1.25}{\delta}\big), \tag{13}$$

then the (ϵ,δ)-LDP constraint is satisfied for all $0 \leq \epsilon \leq 1$ [28], [29], which follows since the ℓ_2-sensitivity of the clipped input $\widetilde{X}$ is $2C$. We also have

$$\sigma_Y^2 = \sigma_{\widetilde{X}}^2 + \sigma_{\widetilde{Z}}^2 \tag{14}$$

which follows since $\widetilde{X}^n$ and $\widetilde{Z}^n$ are independent.

We next consider the two corner points of the coordination-randomness rate region, given in (5) and (6), for the (ϵ,δ)-LDP problem defined above. We remark that both corner points, i.e., the WCI $C(\widetilde{X};Y)$ and the mutual information $I(\widetilde{X};Y)$, are invariant to mean values of $\widetilde{X}$ and Y. Define, for $a \in R$,

$$\{a\}^+ = \max\{a,0\}, \qquad \Phi(a) = 0.5\Big(1+\text{erf}\big(\frac{a}{\sqrt{2}}\big)\Big), \tag{15}$$

$$\bar{\beta} = \frac{C}{\sigma_{\widetilde{X}}^2}, \qquad m(a) = \frac{a}{\sigma_Y}. \tag{16}$$

Now, we provide a lower bound on the WCI and a numerical calculation method for the mutual information $I(\widetilde{X};Y)$.

Theorem 2: Consider a randomized distributed function computation problem with a clipped Gaussian input in the range $[-C,C]$ under an (ϵ,δ)-LDP constraint satisfied by using a Gaussian LDP mechanism. We have the following lower bound on the WCI $C(\widetilde{X};Y)$:

$$C(\widetilde{X};Y) \geq \left\{0.5\log\left(1+\frac{2\sigma_{\widetilde{X}}}{\sigma_Y-\sigma_{\widetilde{X}}}\right) + \log\left(\frac{\text{erf}\big(\frac{\beta}{\sqrt{2}}\big)}{\sqrt{1-2\gamma(\beta)}}\right) - \gamma(\beta)\right\}^+. \tag{17}$$

Moreover, the mutual information $I(\widetilde{X};Y)$ can be computed numerically as

$$I(\widetilde{X};Y) = -\mathbb{E}_{p_Y}[\log(p_Y(Y))] - 0.5\log(2\pi e\sigma_{\widetilde{Z}}^2) \quad (18)$$

where we have the probability density function (pdf), for $y \in (-\infty,\infty)$,

$$p_Y(y) = \frac{\phi(m(y))\left[\Phi\left(\frac{\bar{\beta}\sigma_Y^2 - \sigma_{\widetilde{X}}y}{\sigma_{\widetilde{Z}}\sigma_Y}\right) - \Phi\left(-\frac{\bar{\beta}\sigma_Y^2 + \sigma_{\widetilde{X}}y}{\sigma_{\widetilde{Z}}\sigma_Y}\right)\right]}{\text{erf}\left(\frac{\bar{\beta}}{\sqrt{2}}\right)\sigma_Y}. \quad (19)$$

Proof Sketch: The lower bound on the WCI follows from [30, Theorem 1] that establishes the following lower bound

$$C(\widetilde{X};Y) \geq \left\{C(\widetilde{X}_g;Y_g) + h(\widetilde{X},Y) - h(\widetilde{X}_g,Y_g)\right\}^+ \quad (20)$$

where $(\widetilde{X}_g,Y_g) \sim \mathcal{N}(0,K_{\widetilde{X}Y})$ given that the joint pdf $p_{\widetilde{X},Y}$ satisfies the cross-covariance matrix constraint $K_{\widetilde{X}Y}$. Since $(\widetilde{X}_g,Y_g)$ are jointly Gaussian, we have [18]

$$C(\widetilde{X}_g;Y_g) = \frac{1}{2}\log\left(\frac{1+|\rho_{\widetilde{X}Y}|}{1-|\rho_{\widetilde{X}Y}|}\right) \quad (21)$$

where $\rho_{\widetilde{X}Y} = \dfrac{\sigma_{\widetilde{X}}}{\sigma_Y} = \dfrac{\sigma_{\widetilde{X}}}{\sqrt{\sigma_{\widetilde{X}}^2+\sigma_{\widetilde{Z}}^2}}$ is the correlation coefficient between $\widetilde{X}$ and Y. Moreover, we have

$$\begin{aligned} h(\widetilde{X}_g,Y_g) &= \log(2\pi e\sigma_{\widetilde{X}}\sigma_{\widetilde{Z}}) \\ &= \log(2\pi e\sigma_X\sqrt{1-2\gamma(\beta)}\sigma_{\widetilde{Z}}) \end{aligned} \quad (22)$$

which follows by (11). Similarly, we obtain

$$\begin{aligned} h(\widetilde{X},Y) &= h(\widetilde{X}) + h(\widetilde{Z}) \\ &= \log\left(2\pi e\sigma_X\text{erf}\left(\frac{\beta}{\sqrt{2}}\right)\sigma_{\widetilde{Z}}\right) - \gamma(\beta) \end{aligned} \quad (23)$$

which follows from the properties of the truncated Gaussian random variable $\widetilde{X}$.

Now, consider the mutual information

$$I(\widetilde{X};Y) = h(Y) - h(\widetilde{Z}) = h(Y) - \log(\sqrt{2\pi e}\sigma_{\widetilde{Z}}) \quad (24)$$

which can be calculated numerically from the pdf p_Y given in (19), which is addressed next. We have $Y^n = \widetilde{X}^n + \widetilde{Z}^n$, where $\widetilde{X}^n$ is i.i.d. truncated Gaussian, $\widetilde{Z}^n$ is i.i.d. Gaussian, and $\widetilde{X}^n$ and $\widetilde{Z}^n$ are independent. Thus, we can obtain the pdf of a sum of independent Gaussian and truncated Gaussian random variables by using [31, Lemma 3.1]. ∎

We remark that the lower bound on the WCI $C(\widetilde{X};Y)$, given in (17), is tight if $(\widetilde{X},Y)$ are jointly Gaussian random variables. Moreover, this lower bound can be tighter than the bound given in (7), i.e., the error between the WCI $C(\widetilde{X};Y)$ and (17) can be smaller than the error between the WCI $C(\widetilde{X};Y)$ and the mutual information $I(\widetilde{X};Y)$. Example parameter ranges, obtained through random search, for which (17) is tighter than (7) are given in Table I, where the mutual information $I(\widetilde{X};Y)$ is computed numerically by using (18).

The results in Table I illustrate the significant gains, up to 98.5 times in the last row, from the available common randomness in reducing the amount of information to be transmitted to compute a randomized function output, corresponding to satisfying an (ϵ,δ)-LDP constraint. Moreover, based on the results in Table I, smaller ϵ (i.e., better LDP) and σ_X (i.e., smaller input variance) values result in larger communication load gains from using common randomness for distributed randomized function computation, i.e., the lower bound on the ratio $C(\widetilde{X};Y)/I(\widetilde{X};Y)$ increases.

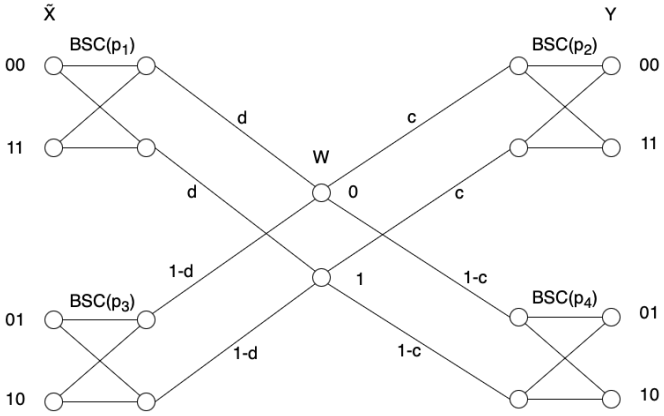

Fig. 2: An example random response setting, where the joint probability distribution $Q_{\widetilde{X}Y}$ can be represented as a combination of mixtures of binary symmetric channels (BSCs). For simplicity, suppose W is uniformly distributed.

IV. Coordination for Symmetric Random Response

We next consider a random response setting, which can be used, e.g., as a LDP mechanism [32], with discrete inputs and outputs, where the input is $\widetilde{X}^n$ and the random response that is simulated at the receiver is Y^n such that they are jointly distributed according to a given probability mass function (pmf) $Q_{\widetilde{X}Y}^n$ that can be represented as in Fig. 2, where we consider 2-bit outputs for simplicity. Suppose the binary random variable W in the middle is uniformly distributed. We remark that this joint pmf is symmetric, as the channels $P_{\widetilde{X}|W}$ and $P_{Y|W}$ can be written as mixtures of binary symmetric channels (BSCs). For the random response setting depicted in Fig. 2, we define $\widetilde{X} = (\widetilde{X}_1,\widetilde{X}_2)$ and $Y = (Y_1,Y_2)$, so we have

$$\begin{bmatrix}\widetilde{X}_1\\ \widetilde{X}_2\\ Y_1\\ Y_2\end{bmatrix} = W\begin{bmatrix}1\\1\\1\\1\end{bmatrix} \oplus \begin{bmatrix}B_1\\B_2\\B_3\\B_4\end{bmatrix} \quad (25)$$

where $\oplus$ represents modulo-2 summation, and B_1-B_4 are mutually dependent binary random variables and they are jointly independent of W. Observe from (25) that since W is uniformly distributed, we have

$$P_{\widetilde{X}_1\widetilde{X}_2Y_1Y_2}(\widetilde{x}_1,\widetilde{x}_2,y_1,y_2) = P_{\widetilde{X}_1\widetilde{X}_2Y_1Y_2}(\bar{\widetilde{x}}_1,\bar{\widetilde{x}}_2,\bar{y}_1,\bar{y}_2) \quad (26)$$

where $\bar{y}=1-y$ is the one's complement of y. Define

$$q_{Y_1Y_2} = P_{Y_1Y_2|W}(y_1,y_2|0), \quad (27)$$

σ_X	ϵ	δ	WCI bound (17)	$I(\widetilde{X};Y)$ (18)	(17) / (18)
0.4938	0.8918	0.0097	0.0324	0.0019	17.05
0.4451	0.7917	0.0058	0.0307	0.0012	25.58
0.6112	0.9256	0.0016	0.0029	0.0019	1.53
0.2064	0.6399	0.0023	0.0186	0.0003	62.00
0.3839	0.3637	0.0061	0.0127	0.0002	63.50
0.4947	0.9884	0.0019	0.0298	0.0018	16.56
0.3280	0.4663	0.0032	0.0197	0.0002	98.50

TABLE I: WCI lower bound computations from (17), numerical computations of $I(\widetilde{X};Y)$ from (18), and their ratios for the (ϵ,δ)-LDP constraints imposed.

then the subchannel probabilities are

$$c = q_{00} + q_{11}, \qquad 1-c = q_{01} + q_{10} \tag{28}$$

and the crossover probabilities of the BSCs are $p_2 = q_{11}/c$ and $p_4 = q_{10}/(1-c)$. The subchannel probabilities and the crossover probabilities for the channel $P_{\widetilde{X}|W}$ can be defined similarly. Moreover, extensions to joint pmfs that can be decomposed into more than 2 BSCs are possible by using similar steps as in [33]. Given a joint pmf $Q_{\widetilde{X}Y}$ for a random response mechanism, one can check whether the symmetry condition in (26) is satisfied, which simplifies the computation of a lower bound on the WCI $C(\widetilde{X};Y)$, as described below.

We first state a lower bound on the WCI $C(\widetilde{X};Y)$, which is a summary of the results in Theorems 4 and 5 in [34]. Define

$$\alpha = \sqrt{\frac{kx-1}{k-1}}, \qquad x_k^* = \frac{k^2-3k+3}{k(k-1)}, \tag{29}$$

$$f_1(x) = -\frac{2}{k}\big[(1+(k-1)\alpha)\log(1+(k-1)\alpha) + (k-1)(1-\alpha)\log(1-\alpha)\big] + 2\log k, \tag{30}$$

$$f_2(x) = 2\log k - 2(k-1)\frac{\log(k-1)}{k-2}\left(x - \frac{1}{k}\right). \tag{31}$$

Consider a $k \times k$ joint probability distribution matrix $\mathbf{Q}_{\widetilde{X}Y}$ with elements $Q_{i,j}$, representing the pmf $Q_{\widetilde{X}Y}$. For instance, we have $k = 4$ for the probability distribution depicted in Fig. 2. We remark that it suffices to consider square matrices $\mathbf{Q}_{\widetilde{X}Y}$, since adding or removing rows or columns of zeros does not affect the WCI calculations [34].

Denote the trace operation as $\mathrm{tr}(\cdot)$, and define the maxitrace operation as

$$\mathrm{maxtr}\,\mathbf{Q} = \max_{\pi\in\mathcal{S}_k} \sum_{i=1}^{k} Q_{i,\pi(i)} \tag{32}$$

where $\mathcal{S}_k$ is the set of all permutations of the indices $\{1,2,\ldots,k\}$.

Theorem 3: Let $k > 2$ and $1/k \le x \le 1$, and

- if $x_k^* \le x \le 1$, consider
$$f(x) := f_1(x) \tag{33}$$
- and, otherwise, if $1/n \le x \le x_k^*$, consider
$$f(x) := f_2(x). \tag{34}$$

Given a probability distribution matrix $\mathbf{Q}_{\widetilde{X}Y}$, we have the following lower bound on the WCI $C(\widetilde{X};Y)$

$$C(\widetilde{X};Y) \ge H(\widetilde{X},Y) - f(\mathrm{maxtr}(\mathbf{Q}_{\widetilde{X}Y})). \tag{35}$$

Now, we compute the lower bound given in Theorem 3 for joint probability distributions $Q_{\widetilde{X}Y}$ that can be represented as in Fig. 2 and (25) to compare it with the corresponding mutual information $I(\widetilde{X};Y)$. We illustrate example parameter ranges, obtained through random search, for which (35) is tighter than (7) in Table II.

The results in Table II illustrate the significant gains, up to 8.36 times in the last row, from the available common randomness in reducing the amount of information to be transmitted to compute a randomized function output. Moreover, based on the results in Table II, smaller mutual information $I(\widetilde{X};Y)$ values result in larger communication load gains from using common randomness for distributed randomized function computation, i.e., the lower bound on the ratio $C(\widetilde{X};Y)/I(\widetilde{X};Y)$ increases. The WCI lower bound is up to 116.55 times smaller than the lossless compression rates $H(\widetilde{X})$ and $H(Y)$, illustrating that randomized distributed function computation can reduce the communication load significantly already without any common randomness.

The results in Sections III and IV illustrate significant gains over classical function computation methods by using semantic communication methods for randomized distributed function computation. Our results motivate further research on the minimum achievable communication load for a fixed amount of common randomness, which is important as common randomness is a true commodity, e.g., for devices with limited storage.

REFERENCES

[1] D. Gündüz, Z. Qin, I. E. Aguerri, H. S. Dhillon, Z. Yang, A. Yener, K. K. Wong, and C.-B. Chae, "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 41, no. 1, pp. 5–41, Nov. 2022.

[2] Q. Lan, D. Wen, Z. Zhang, Q. Zeng, X. Chen, P. Popovski, and K. Huang, "What is semantic communication? A view on conveying

$(p_1, p_2, p_3, p_4, c, d)$	WCI Bound (35)	$I(\widetilde{X};Y)$	(35) / $I(\widetilde{X};Y)$	$H(\widetilde{X})$	$H(Y)$	$\frac{\min\{H(\widetilde{X}),H(Y)\}}{(35)}$
(0.05, 0.45, 0.5, 0.25, 0.45, 0.4)	0.0977	0.0238	4.10	1.3662	1.3813	13.984
(0.05, 0.1, 0.2, 0.4, 0.45, 0.3)	0.0998	0.0822	1.21	1.3040	1.3813	13.069
(0.15, 0.25, 0, 0.05, 0.45, 0.4)	0.3276	0.2570	1.27	1.3662	1.3813	4.170
(0.25, 0.4, 0.1, 0.25, 0.5, 0.25)	0.0967	0.0403	2.40	1.2555	1.3863	12.980
(0.45, 0.4, 0.3, 0.1, 0.45, 0.3)	0.1315	0.0216	6.09	1.3040	1.3813	9.916
(0.3, 0.5, 0.3, 0, 0.5, 0.4)	0.1364	0.0411	3.32	1.3662	1.3863	10.010
(0.1, 0.05, 0.1, 0, 0.1, 0.05)	0.5325	0.3601	1.48	0.8917	1.0182	1.675
(0.25, 0.35, 0.25, 0.25, 0.3, 0.15)	0.1087	0.0255	4.26	1.1159	1.3040	10.261
(0.5, 0, 0.45, 0.3, 0.4, 0.45)	0.0117	0.0014	8.36	1.3813	1.3662	116.550

TABLE II: WCI lower bound computations from (35), exact computations of $I(\widetilde{X};Y)$ from $Q_{\widetilde{X}Y}$, their ratios, lossless source coding rates $H(\widetilde{X})$ and $H(Y)$ to transmit $\widetilde{X}^n$ and Y^n, respectively, and the ratios $\min\{H(\widetilde{X}), H(Y)\}/(35)$.

meaning in the era of machine intelligence," *Journal of Communications and Information Networks (JCIN)*, vol. 6, no. 4, pp. 336–371, Dec. 2021.

[3] D. Gündüz, F. Chiariotti, K. Huang, A. E. Kalør, S. Kobus, and P. Popovski, "Timely and massive communication in 6G: Pragmatics, learning, and inference," *IEEE BITS the Information Theory Magazine*, vol. 3, no. 1, pp. 27–40, Oct. 2023.

[4] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Transactions on Information Theory (T-IT)*, vol. 8, no. 5, pp. 293–304, Sep. 1962.

[5] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

[6] T. Berger, "Rate-distortion theory," *Wiley Encyclopedia of Telecommunications*, 2003.

[7] M. Bloch, O. Günlü, A. Yener, F. Oggier, H. V. Poor, L. Sankar, and R. F. Schaefer, "An overview of information-theoretic security and privacy: Metrics, limits and applications," *IEEE Journal on Selected Areas in Information Theory (JSAIT)*, vol. 2, no. 1, pp. 5–22, Mar. 2021.

[8] D. Data, G. R. Kurri, J. Ravi, and V. M. Prabhakaran, "Interactive secure function computation," *IEEE Transactions on Information Theory (T-IT)*, vol. 66, no. 9, pp. 5492–5521, Mar. 2020.

[9] P. Cuff, "Distributed channel synthesis," *IEEE Transactions on Information Theory (T-IT)*, vol. 59, no. 11, pp. 7071–7096, Aug. 2013.

[10] C. Bennett, P. Shor, J. Smolin, and A. Thapliyal, "Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem," *IEEE Transactions on Information Theory (T-IT)*, vol. 48, no. 10, pp. 2637–2655, Oct. 2002.

[11] G. Kramer and S. A. Savari, "Communicating probability distributions," *IEEE Transactions on Information Theory (T-IT)*, vol. 53, no. 2, pp. 518–525, Jan. 2007.

[12] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," *IEEE Transactions on Information Theory (T-IT)*, vol. 56, no. 1, pp. 438–449, Jan. 2010.

[13] M. Havasi, R. Peharz, and J. M. Hernández-Lobato, "Minimal random code learning: Getting bits back from compressed model parameters," in *International Conference on Learning Representations (ICLR)*, Jan. 2019.

[14] B. Isik, F. Pase, D. Gunduz, S. Koyejo, T. Weissman, and M. Zorzi, "Adaptive compression in federated learning via side information," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Apr. 2024, pp. 487–495.

[15] B. Phan, A. Khisti, and C. Louizos, "Importance matching lemma for lossy compression with side information," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Apr. 2024, pp. 1387–1395.

[16] A. Shah, W.-N. Chen, J. Balle, P. Kairouz, and L. Theis, "Optimal compression of locally differentially private mechanisms," in *International Conference on Artificial Intelligence and Statistics (AISTATS)*, Mar. 2022, pp. 7680–7723.

[17] G. R. Kurri, V. Ramachandran, S. R. B. Pillai, and V. M. Prabhakaran, "Multiple access channel simulation," in *IEEE International Symposium on Information Theory (ISIT)*, July 2021, pp. 2411–2416.

[18] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory (T-IT)*, vol. 21, no. 2, pp. 163–179, Mar. 1975.

[19] O. Günlü and R. F. Schaefer, "An optimality summary: Secret key agreement with physical unclonable functions," *Entropy*, vol. 23, no. 1, p. 16, Dec. 2020.

[20] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge University Press, 2011.

[21] S.-L. Huang, X. Xu, L. Zheng, and G. W. Wornell, "A local characterization for Wyner common information," in *IEEE International Symposium on Information Theory (ISIT)*, June 2020, pp. 2252–2257.

[22] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Transactions on Information Theory (T-IT)*, vol. 56, no. 9, pp. 4181–4206, Aug. 2010.

[23] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[24] S. Asoodeh, M. Aliakbarpour, and F. P. Calmon, "Local differential privacy is equivalent to contraction of an f-divergence," in *IEEE International Symposium on Information Theory (ISIT)*, July 2021, pp. 545–550.

[25] A. Evfimievski, J. Gehrke, and R. Srikant, "Limiting privacy breaches in privacy preserving data mining," in *ACM Symposium on Principles of Database Systems (PODS)*, June 2003, pp. 211–222.

[26] S. P. Kasiviswanathan, H. K. Lee, K. Nissim, S. Raskhodnikova, and A. Smith, "What can we learn privately?" in *IEEE Symposium on Foundations of Computer Science (FOCS)*, Oct. 2008, pp. 531–540.

[27] R. Bassily, A. Smith, and A. Thakurta, "Private empirical risk minimization: Efficient algorithms and tight error bounds," in *IEEE Symposium on Foundations of Computer Science (FOCS)*, Oct. 2014, pp. 464–473.

[28] C. Dwork, A. Roth *et al.*, "The algorithmic foundations of differential privacy," *Foundations and Trends® in Theoretical Computer Science*, vol. 9, no. 3–4, pp. 211–407, 2014.

[29] M. Seif, R. Tandon, and M. Li, "Wireless federated learning with local differential privacy," in *IEEE International Symposium on Information Theory (ISIT)*, June 2020, pp. 2604–2609.

[30] E. Sula and M. Gastpar, "Lower bound on relaxed Wyner's common information," in *IEEE International Symposium on Information Theory (ISIT)*, July 2021, pp. 1510–1515.

[31] E. Drysdale, "A parametric distribution for exact post-selection inference with data carving," *preprint arXiv:2305.12581*, May 2023.

[32] B. Bebensee, "Local differential privacy: A tutorial," *preprint arXiv:1907.11908*, 2019.

[33] O. Günlü and G. Kramer, "Privacy, secrecy, and storage with multiple noisy measurements of identifiers," *IEEE Transactions on Information Forensics and Security (TIFS)*, vol. 13, no. 11, pp. 2872–2883, Nov. 2018.

[34] H. S. Witsenhausen, "Values and bounds for the common information of two discrete random variables," *SIAM Journal on Applied Mathematics (SIAP)*, vol. 31, no. 2, pp. 313–333, Sep. 1976.