Secure Lossy Function Computation with Multiple Private Remote Source Observations

Onur Günlü[1], Matthieu Bloch[2], and Rafael F. Schaefer[1]

[1]Chair of Communications Engineering and Security, University of Siegen,
{onur.guenlue, rafael.schaefer}@uni-siegen.de
[2]School of Electrical and Computer Engineering, Georgia Institute of Technology,
matthieu.bloch@ece.gatech.edu

***Abstract*—We consider that multiple noisy observations of a remote source are used by different nodes in the same network to compute a function of the noisy observations under joint secrecy, joint privacy, and individual storage constraints, as well as a distortion constraint on the function computed. Suppose that an eavesdropper has access to one of the noisy observations in addition to the public messages exchanged between legitimate nodes. This model extends previous models by 1) considering a remote source as the source of dependency between the correlated random variables observed at different nodes; 2) allowing the function computed to be a distorted version of the target function, which allows to reduce the storage rate as compared to a reliable function computation scenario in addition to reducing secrecy and privacy leakages; 3) introducing a privacy metric that measures the information leakage about the remote source to the fusion center in addition to the classic privacy metric that measures the leakage to an eavesdropper; 4) considering two transmitting nodes to compute a function rather than one node. Single-letter inner and outer bounds are provided for the considered lossy function computation problem, and simplified bounds are established for two special cases, in which either the computed function is partially invertible or the function is invertible and the measurement channel of the eavesdropper is physically degraded with respect to the measurement channel of the fusion center.**

I. Introduction

Function computation in a network is considered, in which dependent random variables are observed by all nodes. Such a model fits well to recent applications that are based on, e.g., distributed machine learning [1], [2] and network function virtualization [3]. The aim of the nodes is to compute a function of their observed random variables by exchanging public messages over authenticated and noiseless communication links such that a fusion center that observes the public messages can compute the function output. Messages exchanged over public communication links cause information leakage to an eavesdropper about the output of the function computed, which results in *secrecy leakage* [4]–[6]. We consider the general case where an eavesdropper observes both public messages and random variables that are correlated with other random variables, as in [7]–[11]. Due to storage constraints on the communication links, it is also necessary to minimize the amount of messages exchanged over the public communication links [12]. The amount of *public storage* can be reduced by using, e.g., distributed lossless or lossy source coding methods [13]–[17], the latter of which allows the function computed to be a distorted version of the target function and applies Wyner-Ziv (WZ) coding [18] methods that result in further reductions compared to the former. Furthermore, the dependency between the random variables observed by different nodes is posited in [19] to stem from a remote source whose noisy measurements are observed by different nodes in the same network. Such a remote source model has been used for secret key agreement [20]–[22], device authentication [23], and other information-theoretic problems [24, p. 118], [25, p. 78]. For function computation scenarios, we also consider such a remote source. Thus, every function computation in the same network results in information leakage to an eavesdropper about the remote source, which is called *privacy leakage* [26], since the remote source is common to all random sequences observed by each node. Moreover, another privacy leakage metric that measures the information leakage about the remote source to the fusion center is also considered in [19], [27].

We consider function computation scenarios with one secrecy, two privacy, two storage, and one distortion constraints to obtain inner and outer bounds for the rate regions that correspond to the optimal trade-off between all constraints considered. In our models, two transmitting nodes send public messages to a fusion center in order for the fusion center to compute a distorted version of a target function by using both the public message and its noisy observations of the remote source. Our function computation model is a strict extension of previous models considered since 1) in [26] a visible source model is used, which cannot explain how the dependency between different random variables observed by different nodes is established; 2) in [19] only one transmitting node was considered for function computation; 3) in [27] lossless function computation was considered, which imposes the stringent reliability constraint that the function computed should be equal to the target function and this might require a larger amount of public storage as compared to a lossy function computation model considered in this work.

The main contributions of this work are summarized as follows. The lossy single-function computation model with two transmitting nodes is considered and an inner bound for the rate region that characterizes the optimal trade-off between secrecy, privacy, storage, and distortion constraints is established by using the output statistics of random binning (OSRB) method [28]. An outer bound for the same rate region is also provided. Furthermore, effects of considering a distortion constraint, rather than a reliability constraint, on the function computation are discussed. For partially invertible functions, which define a set that is a proper superset of the set of invertible functions, we establish simplified lossy rate region bounds. We also provide simplified bounds for invertible functions when the eavesdropper's measurement channel is physically degraded with respect to the fusion center's channel.

In Section II, the lossy single-function computation model with two transmitting nodes and under secrecy, privacy, storage, and distortion constraints is introduced. In Section III, we first provide inner and outer bounds for the lossy single-function computation problem introduced, and then simplified lossy rate region bounds for two special cases are given. In Section IV, we conclude the paper.

A. Notation

Upper case letters represent random variables and lower case letters their realizations. A superscript denotes a sequence of variables, e.g., $X^n = X_1, X_2, \dots, X_i, \dots, X_n$, and a subscript i denotes the position of a variable in a sequence. A random variable X has probability distribution P_X. Calligraphic letters such as $\mathcal{X}$ denote sets and set sizes are written as $|\mathcal{X}|$. $[1\!:\!J]$ denotes the set $\{1,2,\dots,J\}$ for an integer $J \geq 1$, and $X \sim \text{Unif}[1:J]$ is a random variable that is uniformly distributed over the set $[1:J]$.

II. System Model

We consider an independent and identically distributed (i.i.d.) remote source $X^n \sim P_X^n$ that is measured by two transmitting nodes, a fusion center, and an eavesdropper (Eve) through noisy memoryless channels $P_{\widetilde{X}_1|X}$, $P_{\widetilde{X}_2|X}$, and $P_{YZ|X}$, respectively. Thus, we denote the observations of the transmitting nodes as $\widetilde{X}_1^n$ and $\widetilde{X}_2^n$, and the observations of the fusion center and eavesdropper as Y^n and Z^n, respectively. Suppose the source and measurement alphabets are finite sets. Transmitting nodes' encoders $\mathsf{Enc}_1(\cdot)$ and $\mathsf{Enc}_2(\cdot)$ send public indices W_1 and W_2, respectively, to the fusion center over authenticated, one-way, and noiseless communication links. Observing W_1, W_2, and Y^n, the fusion center decoder $\mathsf{Dec}(\cdot)$ estimates a distorted version of the target function $f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)$ that is such that

$$f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n) = \{f(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}, Y_i)\}_{i=1}^n. \tag{1}$$

The considered function computation model with two transmitting nodes is illustrated in Fig. 1 on the next page. We next define achievable rate tuples and the secrecy, privacy, storage, and distortion constraints imposed. We remark that the secrecy leakage constraint imposed does not depend on the properties of the function $f(\cdot,\cdot,\cdot)$ computed, as in [19], [26], [27].

Definition 1. A *lossy* tuple $(R_{\mathrm{s}}, R_{\mathrm{w},1}, R_{\mathrm{w},2}, R_{\ell,\mathrm{Dec}}, R_{\ell,\mathrm{Eve}}, D)$ is *achievable* if, for any $\delta > 0$, there exist $n \geq 1$, two encoders, and one decoder such that

$$\frac{1}{n} I(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n; W_1, W_2 | Z^n) \leq R_{\mathrm{s}} + \delta \;\text{(secrecy)} \tag{2}$$

$$\frac{1}{n} \log\big|\mathcal{W}_1\big| \leq R_{\mathrm{w},1} + \delta \qquad \text{(storage 1)} \tag{3}$$

$$\frac{1}{n} \log\big|\mathcal{W}_2\big| \leq R_{\mathrm{w},2} + \delta \qquad \text{(storage 2)} \tag{4}$$

$$\frac{1}{n} I(X^n; W_1, W_2 | Y^n) \leq R_{\ell,\mathrm{Dec}} + \delta \qquad \text{(privacyDec)} \tag{5}$$

$$\frac{1}{n} I(X^n; W_1, W_2 | Z^n) \leq R_{\ell,\mathrm{Eve}} + \delta \qquad \text{(privacyEve)} \tag{6}$$

$$\mathbb{E}\Big[d(f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n), \widehat{f^n})\Big] \leq D + \delta \qquad \text{(distortion)} \tag{7}$$

where

$$d(f^n, \widehat{f^n}) = \frac{1}{n} \sum_{i=1}^n d(f_i, \widehat{f_i}) \tag{8}$$

is a per-letter distortion metric. The *lossy* region $\mathcal{R}_{\mathrm{D}}$ is the closure of the set of all achievable lossy tuples.

III. Inner and Outer Bounds

Given any $a \in \mathbb{R}$, define $[a]^- = \min\{a, 0\}$. We next provide inner and outer bounds for the lossy region $\mathcal{R}_{\mathrm{D}}$; see below for a proof sketch.

Theorem 1. (Outer Bound): *An outer bound for the lossy region $\mathcal{R}_D$ is the union over all P_Q, $P_{V_1|Q}$, $P_{V_2|Q}$, $P_{U_1|V_1}$, $P_{U_2|V_2}$, $P_{\widetilde{X}_1|U_1}$, and $P_{\widetilde{X}_2|U_2}$ of the set of rate tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve}, D)$ such that*

$$\begin{aligned} R_s \geq \Big[I(U_1,U_2;Z|V_1,V_2,Q) - I(U_1,U_2;Y|V_1,V_2,Q)\Big]^- \\ + I(U_1,U_2;\widetilde{X}_1,\widetilde{X}_2|Z) \end{aligned} \tag{9}$$

$$\begin{aligned} R_{w,1} \geq I(V_1;\widetilde{X}_1|V_2,Y) + I(U_1;\widetilde{X}_1|V_1,U_2,Y) \\ - I(V_1;V_2|\widetilde{X}_1,Y) - I(U_1;U_2|\widetilde{X}_1,Y,V_1) \end{aligned} \tag{10}$$

$$\begin{aligned} R_{w,2} \geq I(V_2;\widetilde{X}_2|V_1,Y) + I(U_2;\widetilde{X}_2|U_1,V_2,Y) \\ - I(V_2;V_1|\widetilde{X}_2,Y) - I(U_2;U_1|\widetilde{X}_2,Y,V_2) \end{aligned} \tag{11}$$

$$\begin{aligned} R_{w,1} + R_{w,2} \geq I(U_2;\widetilde{X}_2|U_1,V_2,Y) + I(U_1;\widetilde{X}_1|V_1,V_2,Y) \\ + I(V_2;\widetilde{X}_2|V_1,Y) + I(V_1;\widetilde{X}_1|Y) \end{aligned} \tag{12}$$

$$R_{\ell,Dec} \geq I(U_1,U_2;X|Y) \tag{13}$$

$$\begin{aligned} R_{\ell,Eve} \geq \Big[I(U_1,U_2;Z|V_1,V_2,Q) - I(U_1,U_2;Y|V_1,V_2,Q)\Big]^- \\ + I(U_1,U_2;X|Z) \end{aligned} \tag{14}$$

$$D \geq \mathbb{E}[d(f(\widetilde{X}_1,\widetilde{X}_2,Y), g(U_1,U_2,Y))] \tag{15}$$

for some function $g(\cdot,\cdot,\cdot)$ and where

$$(Q,V_1) - U_1 - \widetilde{X}_1 - X - (\widetilde{X}_2,Y,Z) \tag{16}$$

$$(Q,V_2) - U_2 - \widetilde{X}_2 - X - (\widetilde{X}_1,Y,Z) \tag{17}$$

form Markov chains. One can limit the cardinalities to $|\mathcal{Q}| \leq 2$, $|\mathcal{V}_1| \leq |\widetilde{X}_1|+7$, $|\mathcal{V}_2| \leq |\widetilde{X}_2|+7$, $|\mathcal{U}_1| \leq (|\widetilde{X}_1|+7)^2$, and $|\mathcal{U}_2| \leq (|\widetilde{X}_2|+7)^2$.

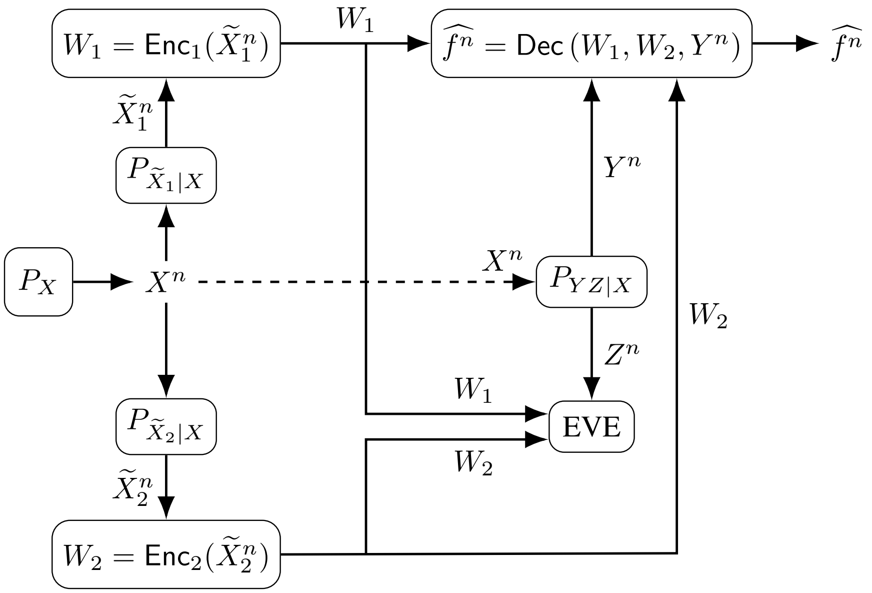

Fig. 1. Lossy single-function computation model that uses two transmitting nodes, where the computed function is allowed to be a distorted version of the target function.

(Inner Bound): *An achievable lossy region is the union over all P_Q, $P_{V_1|Q}$, $P_{V_2|Q}$, $P_{U_1|V_1}$, $P_{U_2|V_2}$, $P_{\widetilde{X}_1|U_1}$, and $P_{\widetilde{X}_2|U_2}$ of the rate tuples in (9), (12)-(15), and*

$$R_{w,1} \geq I(V_1;\widetilde{X}_1|V_2,Y) + I(U_1;\widetilde{X}_1|V_1,U_2,Y) \qquad (18)$$

$$R_{w,2} \geq I(V_2;\widetilde{X}_2|V_1,Y) + I(U_2;\widetilde{X}_2|U_1,V_2,Y) \qquad (19)$$

and where we have

$$\begin{aligned} &P_{QV_1V_2U_1U_2\widetilde{X}_1\widetilde{X}_2XYZ} \\ &= P_{Q|V_1V_2}P_{V_1|U_1}P_{U_1|\widetilde{X}_1}P_{\widetilde{X}_1|X}P_{V_2|U_2}P_{U_2|\widetilde{X}_2}P_{\widetilde{X}_2|X}P_XP_{YZ|X}. \end{aligned} \qquad (20)$$

Proof Sketch: The proof of the outer bound applies the standard properties of the Shannon entropy and follows mainly from the outer bound proof for the lossless version of the function computation problem depicted in Fig. 1, for which the distortion constraint (7) is replaced with a reliability constraint; see [27, Section IV] for the inner and outer bound proofs for the lossless function computation problem. However, the proof for the lossless function computation problem requires the auxiliary random variables to be admissible as defined in [12], unlike the lossy function computation problem. Thus, the outer bound proof for Theorem 1 follows by replacing the admissibility step in the outer bound proof for the lossless function computation problem with the steps

$$\begin{aligned} &n(D+\delta_n) \\ &\overset{(a)}{\geq} \mathbb{E}\Big[\sum_{i=1}^{n} d\left(f_i(\widetilde{X}_{1,i},\widetilde{X}_{2,i},Y_i), \widehat{f}_i(W_1,W_2,Y^n)\right)\Big] \\ &\overset{(b)}{\geq} \mathbb{E}\Big[\sum_{i=1}^{n} d\left(f_i(\widetilde{X}_{1,i},\widetilde{X}_{2,i},Y_i), g_i(W_1,W_2,Y^n,X^{i-1},Z^{i-1})\right)\Big] \\ &\overset{(c)}{=} \mathbb{E}\Big[\sum_{i=1}^{n} d\left(f_i(\widetilde{X}_{1,i},\widetilde{X}_{2,i},Y_i), g_i(W_1,W_2,Y_i^n,X^{i-1},Z^{i-1})\right)\Big] \\ &\overset{(d)}{=} \mathbb{E}\Big[\sum_{i=1}^{n} d\left(f(\widetilde{X}_{1,i},\widetilde{X}_{2,i},Y_i), g(U_{1,i},U_{2,i},Y_i)\right)\Big] \end{aligned} \qquad (21)$$

where (a) follows by (7) and (8), (b) follows since there exists a function $g_i(\cdot,\cdot,\cdot)$ that achieves a distortion that is not greater than the distortion achieved by $\widehat{f}_i(W_1,W_2,Y^n)$, where the distortion is measured with respect to $f_i(\widetilde{X}_{1,i},\widetilde{X}_{2,i},Y_i)$, since $g_i(\cdot,\cdot,\cdot)$ has additional inputs, (c) follows from the Markov chain

$$Y^{i-1} - (X^{i-1},Z^{i-1},W_1,W_2,Y_i,Y_{i+1}^n) - f_i \qquad (22)$$

and (d) follows from the definitions of

$$U_{1,i} \triangleq (W_1,X^{i-1},Y_{i+1}^n,Z^{i-1}) \qquad (23)$$

$$U_{2,i} \triangleq (W_2,X^{i-1},Y_{i+1}^n,Z^{i-1}). \qquad (24)$$

We next introduce a uniformly distributed time-sharing random variable $Q \sim \text{Unif}[1:n]$ that is independent of other random variables. By defining $X = X_Q$, $\widetilde{X}_1 = \widetilde{X}_{1,Q}$, $\widetilde{X}_2 = \widetilde{X}_{2,Q}$, $Y = Y_Q$, $Z = Z_Q$, $V_1 = V_{1,Q}$, $V_2 = V_{2,Q}$, $U_1 = (U_{1,Q},Q)$, $U_2 = (U_{2,Q},Q)$, and $f = f_Q$ such that (16) and (17) form Markov chains, the proof of the outer bound follows. Furthermore, the proof of the cardinality bounds follows from [24, Lemma 15.4] since we preserve the same probability and conditional entropy values as being preserved for the lossless function computation problem with the addition of preserving the value of $g(U_1,U_2,Y) = g(U_1,U_2,V_1,V_2,Y)$, following from the Markov chain

$$(V_1,V_2) - (U_1,U_2,Y) - g(U_1,U_2,Y). \qquad (25)$$

For the proof of the inner bound, we use the OSRB method that assigns two random bin indices to each auxiliary sequence $U_1^n=u_1^n$, $U_2^n=u_2^n$, $V_1^n=v_1^n$, and $V_2^n=v_2^n$ separately. The first set of random bin indices represents the public choices of two encoders and one decoder, whereas the second set of random bin indices the public messages sent to the fusion center from encoders. The fusion center that observes all public random bin indices as well as Y^n applies the following successive decoding order:

1) using Y^n and public bin indices, the fusion center estimates V_1^n as $\widehat{V}_1^n$;
2) using $(Y^n,\widehat{V}_1^n)$ and public bin indices, V_2^n is estimated as $\widehat{V}_2^n$;
3) using $(Y^n,\widehat{V}_1^n,\widehat{V}_2^n)$ and public bin indices, U_1^n is estimated as $\widehat{U}_1^n$;
4) using $(Y^n,\widehat{V}_1^n,\widehat{V}_2^n,\widehat{U}_1^n)$ and public bin indices, U_2^n is estimated as $\widehat{U}_2^n$.

Furthermore, by swapping the indices 1 and 2 in the decoding order above the other corner point in the achievable rate region is obtained, so it suffices to analyze the given decoding order. We impose conditions on the rates of the first and second sets of random bin indices to ensure reliable estimations at the fusion center [28, Lemma 1] as well as to ensure that the choices of encoders and decoders are independent of the random sequences observed [28, Theorem 1]. Using the OSRB method consecutively, various different recoverability cases that indicate whether it is possible obtain single-letter terms are analyzed. All cases are bounded by the same mutual information terms. We remark that the achievability proof of the lossy function computation problem follows from the achievability proof of its lossless version by replacing the admissibility constraint with the constraint that $P_{U_1|\widetilde{X}_1}$, $P_{V_1|U_1}$, $P_{U_2|\widetilde{X}_2}$, and $P_{V_2|U_2}$ are chosen such that there exists a function $g(U_1,U_2,Y)$ such that

$$g^n(U_1^n,U_2^n,Y^n)=\{g(U_{1,i},U_{2,i},Y_i)\}_{i=1}^n \tag{26}$$

$$\mathbb{E}[d(f^n(\widetilde{X}_1^n,\widetilde{X}_2^n,Y^n),g^n(U_1^n,U_2^n,Y^n))]\leq D+\epsilon_n \tag{27}$$

where $\epsilon_n>0$ such that $\epsilon_n\rightarrow 0$ when $n\rightarrow\infty$. Since all $(\widetilde{x}_1^n,\widetilde{x}_2^n,y^n,u_1^n,u_2^n)$ tuples are in the jointly typical set with high probability, by the typical average lemma [29, pp. 26], constraint in (7) is satisfied. Furthermore, a time-sharing random variable Q such that $P_{V_1V_2Q}=P_QP_{V_1|Q}P_{V_2|Q}$ is used to convexify the rate region. ■

Remark 1. *Since all terms given in the outer bound in Theorem 1, i.e., lower bounds in (9)-(15), are generally strictly positive, strong secrecy or strong privacy constraints cannot be satisfied in general for the lossy function computation problem depicted in Fig. 1.*

One can show that the terms in (10) and (11) recover the terms in (18) and (19), respectively, if the joint probability distribution given in (20) is imposed on the outer bound since the negative terms in (10) and (11) are constant for (20) because of the Markov chains

$$(V_1,U_1)-\widetilde{X}_1-(Y,U_2,V_2) \tag{28}$$

$$(V_2,U_2)-\widetilde{X}_2-(Y,U_1,V_1). \tag{29}$$

However, the rate region that is defined by the outer bound that satisfies (16) and (17) is in general larger than the one that is defined by the inner bound that satisfies (20), so the outer and inner bounds do not match in general.

We next impose the condition that the function $f(\widetilde{X}_1,\widetilde{X}_2,Y)$ is partially invertible with respect to $\widetilde{X}_1$, i.e., we have [13], [30]

$$H(\widetilde{X}_1|f(\widetilde{X}_1,\widetilde{X}_2,Y),Y)=0. \tag{30}$$

For such functions, it is straightforward to show that we have the following achievable rate region for the lossy function computation problem with two transmitting nodes. We remark that the proof of Corollary 1 follows from the inner bound in Theorem 1 by assigning $U_1=\widetilde{X}_1$ and the corresponding outer bound can be similarly obtained from Theorem 1. Furthermore, by symmetry the lossy rate region bounds for a function $f(\widetilde{X}_1,\widetilde{X}_2,Y)$ that is partially invertible with respect to $\widetilde{X}_2$ can be obtained by assigning $U_2=\widetilde{X}_2$.

Corollary 1. *The lossy region $\mathcal{R}_D$ when $f(\widetilde{X}_1,\widetilde{X}_2,Y)$ is a partially invertible function with respect to $\widetilde{X}_1$ includes the set of all tuples $(R_s,R_{w,1},R_{w,2},R_{\ell,Dec},R_{\ell,Eve},D)$ satisfying*

$$R_s\geq\Big[I(\widetilde{X}_1,U_2;Z|V_1,V_2,Q)-I(\widetilde{X}_1,U_2;Y|V_1,V_2,Q)\Big]^- +H(\widetilde{X}_1|Z)+I(U_2;\widetilde{X}_2|\widetilde{X}_1,Z) \tag{31}$$

$$R_{w,1}\geq H(\widetilde{X}_1|V_2,Y)-I(\widetilde{X}_1;U_2|V_1,V_2,Y)$$

$$R_{w,2}\geq I(V_2;\widetilde{X}_2|V_1,Y)+I(U_2;\widetilde{X}_2|\widetilde{X}_1,V_2,Y) \tag{32}$$

$$R_{w,1}+R_{w,2}\geq I(U_2;\widetilde{X}_2|\widetilde{X}_1,V_2,Y)+H(\widetilde{X}_1|V_1,V_2,Y) +I(V_2;\widetilde{X}_2|V_1,Y)+I(V_1;\widetilde{X}_1|Y) \tag{33}$$

$$R_{\ell,Dec}\geq I(\widetilde{X}_1,U_2;X|Y) \tag{34}$$

$$R_{\ell,Eve}\geq\Big[I(\widetilde{X}_1,U_2;Z|V_1,V_2,Q)-I(\widetilde{X}_1,U_2;Y|V_1,V_2,Q)\Big]^- +I(\widetilde{X}_1,U_2;X|Z) \tag{35}$$

$$D\geq\mathbb{E}[d(f(\widetilde{X}_1,\widetilde{X}_2,Y),\ell(\widetilde{X}_1,U_2,Y))] \tag{36}$$

for some function $\ell(\cdot,\cdot,\cdot)$ such that (20) follows with $U_1=\widetilde{X}_1$.

Similar to partially invertible functions, we can establish the simplified lossy rate region bounds for invertible functions, i.e., we have

$$H(\widetilde{X}_1,\widetilde{X}_2|f(\widetilde{X}_1,\widetilde{X}_2,Y),Y)=0. \tag{37}$$

Furthermore, we also impose the condition that the measurement channel $P_{YZ|X}$ is physically degraded such that

$$P_{YZ|X}=P_{Y|X}P_{Z|Y}. \tag{38}$$

For invertible functions and physically degraded measurement channels $P_{YZ|X}$, as defined in (38), we provide a simplified achievable lossy rate region in Corollary 2. The proof of

Corollary 2 follows from Theorem 1 by assigning $U_1 = \widetilde{X}_1$, $U_2 = \widetilde{X}_2$, constant V_1 and V_2, and by using the following Markov chain for this case

$$(\widetilde{X}_1, \widetilde{X}_2) - X - Y - Z \tag{39}$$

which follows by (38). Note that choosing V_1 and V_2 constant further simplifies the achievable rate region, which might provide a suboptimal result.

Corollary 2. *When $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is an invertible function and $P_{YZ|X}$ is as given in (38), the lossy region $\mathcal{R}_D$ includes the set of all tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve}, D)$ satisfying*

$$R_s \geq H(\widetilde{X}_1, \widetilde{X}_2|Y) \tag{40}$$

$$R_{w,1} \geq H(\widetilde{X}_1|\widetilde{X}_2, Y) \tag{41}$$

$$R_{w,2} \geq H(\widetilde{X}_2|\widetilde{X}_1, Y) \tag{42}$$

$$R_{w,1} + R_{w,2} \geq H(\widetilde{X}_1, \widetilde{X}_2|Y) \tag{43}$$

$$R_{\ell,Dec} \geq I(\widetilde{X}_1, \widetilde{X}_2; X|Y) \tag{44}$$

$$R_{\ell,Eve} \geq I(\widetilde{X}_1, \widetilde{X}_2; X|Y) \tag{45}$$

$$D \geq \mathbb{E}[d(f(\widetilde{X}_1, \widetilde{X}_2, Y), k(\widetilde{X}_1, \widetilde{X}_2, Y))] \tag{46}$$

for some function $k(\cdot, \cdot, \cdot)$.

IV. Conclusion

We considered the function computation problem, where three nodes observe correlated random variables and aim to compute a target function of their observations at the fusion center node. We modeled the source of the correlation between these nodes by positing that all three random variables are noisy observations of a remote random source. Furthermore, we imposed one secrecy, two privacy, two storage, and one distortion constraints on this function computation problem to define a lossy rate region by considering an eavesdropper with a correlated random variable and by allowing the function computed to be a distorted version of the target function. We proposed inner and outer bounds for the lossy rate region. The secrecy leakage and privacy leakage rates that are measured with respect to the eavesdropper were shown to be different due to the remote source considered, unlike in the literature. Furthermore, we established simplified lossy rate region bounds for functions that are partially invertible with respect to one of the transmitting node observations and also for invertible functions when the measurement channel of the fusion center and eavesdropper is physically degraded.

In future work, we will propose inner and outer bounds for the lossy multi-function computation problem with multiple transmitting nodes.

Acknowledgment

This work has been supported in part by the German Research Foundation (DFG) under the Grant SCHA 1944/9 - 1 and in part by the US National Science Foundation (NSF) under the Grant CCF 1955401.

References

[1] J. B. Pred, S. B. Kulkarni, and H. V. Poor, "Distributed learning in wireless sensor networks," *IEEE Sign. Process. Mag.*, vol. 23, no. 4, pp. 56–69, July 2006.

[2] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 2031–2063, Thirdquarter 2020.

[3] R. Mijumbi, J. Serrat, J.-L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 236–262, Firstquarter 2016.

[4] H. Tyagi, P. Narayan, and P. Gupta, "When is a function securely computable?" *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6337–6350, Oct. 2011.

[5] A. C. Yao, "How to generate and exchange secrets," in *IEEE Symp. Foundations Comp. Sci.*, Toronto, ON, Canada, Oct. 1986, pp. 162–167.

[6] ——, "Protocols for secure computations," in *IEEE Symp. Foundations Comp. Sci.*, Chicago, IL, Nov. 1982, pp. 160–164.

[7] D. Gunduz, E. Erkip, and H. V. Poor, "Secure lossless compression with side information," in *IEEE Inf. Theory Workshop*, Porto, Portugal, May 2008, pp. 169–173.

[8] H. Tyagi and S. Watanabe, "Converses for secret key agreement and secure computing," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4809–4827, July 2015.

[9] M. Goldenbaum, H. Boche, and H. V. Poor, "On secure computation over the binary modulo-2 adder multiple-access wiretap channel," in *IEEE Inf. Theory Workshop*, Cambridge, U.K., Sep. 2016, pp. 21–25.

[10] V. Prabhakaran and K. Ramchandran, "On secure distributed source coding," in *IEEE Inf. Theory Workshop*, Lake Tahoe, CA, Sep. 2007, pp. 442–447.

[11] H. Tyagi and S. Watanabe, "A bound for multiparty secret key agreement and implications for a problem of secure computing," in *Int. Conf. Theory Appl. Crypt. Techn.*, Copenhagen, Denmark, May 2014, pp. 369–386.

[12] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

[13] M. Sefidgaran and A. Tchamkerten, "Computing a function of correlated sources: A rate region," in *IEEE Int. Symp. Inf. Theory*, St. Petersburg, Russia, July-Aug. 2011, pp. 1856–1860.

[14] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6180–6195, Aug. 2011.

[15] H. Kowshik and P. R. Kumar, "Optimal function computation in directed and undirected graphs," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3407–3418, Feb. 2012.

[16] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Sep. 2007.

[17] S. Kannan and P. Viswanath, "Multi-session function computation and multicasting in undirected graphs," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 702–713, Mar. 2013.

[18] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[19] O. Günlü, M. Bloch, and R. F. Schaefer, "Secure multi-function computation with private remote sources," June 2021, [Online]. Available: arxiv.org/abs/2106.09485.

[20] O. Günlü, R. F. Schaefer, and H. V. Poor, "Biometric and physical identifiers with correlated noise for controllable private authentication," July 2020, [Online]. Available: arxiv.org/abs/2001.00847.

[21] O. Günlü, "Key agreement with physical unclonable functions and biometric identifiers," Ph.D. dissertation, TU Munich, Germany, Nov. 2018, published by Dr.-Hut Verlag in Feb. 2019.

[22] O. Günlü, R. F. Schaefer, and G. Kramer, "Private authentication with physical identifiers through broadcast channel measurements," in *IEEE Inf. Theory Workshop*, Visby, Sweden, Aug. 2019, pp. 1–5.

[23] Y. Wang, S. Rane, S. C. Draper, and P. Ishwar, "A theoretical analysis of authentication, privacy, and reusability across secure biometric systems," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 6, pp. 1825–1840, July 2012.

[24] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge, U.K.: Cambridge University Press, 2011.

[25] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[26] W. Tu and L. Lai, “On function computation with privacy and secrecy constraints,” *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6716–6733, Oct. 2019.

[27] O. Günlü, M. Bloch, and R. F. Schaefer, “Multiple noisy private remote source observations for secure function computation,” in *Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Oct.-Nov. 2021, to appear.

[28] M. H. Yassaee, M. R. Aref, and A. Gohari, “Achievability proof via output statistics of random binning,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6760–6786, Nov. 2014.

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

[30] T. Ericson and J. Körner, “Successive encoding of correlated sources,” *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 390–395, May 1983.