*Article*

# Function Computation Under Privacy, Secrecy, Distortion, and Communication Constraints

**Onur Günlü**

Chair of Communications Engineering and Security, University of Siegen, 57076 Siegen, Germany; onur.guenlue@uni-siegen.de

† This paper is an extended version of the papers that will appear in the proceedings of the 2021 International ITG Workshop on Smart Antennas in [1] and the 2021 Asilomar Conference on Signals, Systems, and Computers in [2].

**Abstract:** The problem of reliable function computation is extended by imposing privacy, secrecy, and storage constraints on a remote source whose noisy measurements are observed by multiple parties. The main additions to the classic function computation problem include 1) privacy leakage to an eavesdropper is measured with respect to the remote source rather than the transmitting terminals' observed sequences; 2) the information leakage to a fusion center with respect to the remote source is considered as a new privacy leakage metric; 3) the function computed is allowed 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; 4) two transmitting node observations are used to compute a function. Inner and outer bounds on the rate regions are derived for lossless and lossy single-function computation with two transmitting nodes, which recover previous results in the literature. For special cases that include invertible and partially-invertible functions, and degraded measurement channels, exact lossless and lossy rate regions are characterized, and one exact region is evaluated for an example scenario.

**Keywords:** Information theoretic privacy; secure function computation; remote source; distributed computation.

## 1. Introduction

We consider function computation scenarios in a network with multiple nodes involved. Each node observes a random sequence and all observed random sequences are modeled to be correlated. Recent advancements in network function virtualization [3] and distributed machine learning applications [4] make function computation in a wireless network via software defined networking an important practical problem that should be tackled to improve the performance of future communication systems. In a classic function computation scenario, the nodes exchange messages through authenticated, noiseless, and public communication links, which results in undesired information leakage about the function computed [5–7]. Furthermore, it is possible to reduce the amount of public communications [8,9] by using distributed lossless or lossy source coding methods; see [10–14] for several extensions. The former method uses Slepian-Wolf (SW) coding [15] constructions and the latter allows the function computed to be a distorted version of the target function and applies Wyner-Ziv (WZ) coding [16] methods that result in further reductions compared to the former. A decrease in public communication is important also to limit the information about the computed function leaked to an eavesdropper in the same network, i.e., *secrecy leakage*. In addition to the public messages, an eavesdropper has generally access to a random sequence correlated with other sequences; see [17–19] for various secure function computation extensions.

An important addition to the secure function computation model is a *privacy* constraint that measures the amount of information about the observed sequence leaked to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

an eavesdropper [20]. Providing privacy is necessary to ensure confidentiality of a private sequence that can be reused for future function computations [21,22]. An extension of the results in [20] are given in [23], where two privacy constraints are considered on a remote source whose different noisy measurements are observed by multiple nodes in the same network. The extension in [23] is different from the previous secure and private function computation models due to the posit that there exists a remote source that is the main reason for the correlation between the random sequences observed by the nodes in the same network. It is illustrated via practical examples that considering a remote source hinders unexpected decrease in reliability and unnoticed secrecy leakage [22]. Similarly, such a remote source model is proposed, e.g., in [24] for biometric secrecy and in [25,26] for user or device authentication problems. It is shown in [23] that with such a remote source model two different privacy leakage rate values should be limited, unlike a single constraint considered in [20].

We consider a private remote source whose three noisy versions are used for secure single-function computation. Suppose two nodes transmit public indices to a fusion center to compute one function. In [23], for each function computation one node sends a public index to a fusion center. In [20], cases with two transmitting nodes for function computation are considered for a visible source model, whose results are improved in this work for a remote source model with an additional privacy leakage constraint. Furthermore, we also consider function computation scenarios where the function computed is allowed to be a distorted version of the target function, which is relevant for various recent function computation applications.

### *1.1. Models for Function Inputs and Outputs*

We consider noisy remote source output measurements that are independent and identically distributed (i.i.d.) according to a fixed probability distribution and that are inputs of a target function. This model is reasonable if, e.g., one uses transform-coding algorithms from [27–30] to extract almost i.i.d. symbols, as applied in the biometric security, physical unclonable function, and image and video coding literature. Furthermore, the set of target functions we study are applied per-letter, i.e., the same function is applied to each input symbol; see Section 2 below. These functions are realistic and are used in various recent applications, such as distributed and federated learning applications where the same loss function is applied to each data example [31].

### *1.2. Summary of Contributions*

We extend the lossless and lossy rate region analysis of the single-function computation model with one transmitting node in [23] to consider two transmitting nodes with joint secrecy and privacy constraints, as well as a distortion constraint on the computed function. A summary of the main contributions is as follows.

- The lossless 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 [32]. An outer bound for the same rate region is also provided by using standard properties of Shannon entropy. Inner and outer bounds are shown to not match in general due to different Markov chains imposed.
- The proposed inner and outer bounds are extended for the lossy single-function computation model with two transmitting nodes by considering a distortion metric. Furthermore, effects of considering a distortion constraint, rather than a reliability constraint, on the function computation are discussed.
- For both partially-invertible functions, which define a set that is a proper superset of the set of invertible functions, and invertible functions, we characterize the exact lossless and lossy rate regions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

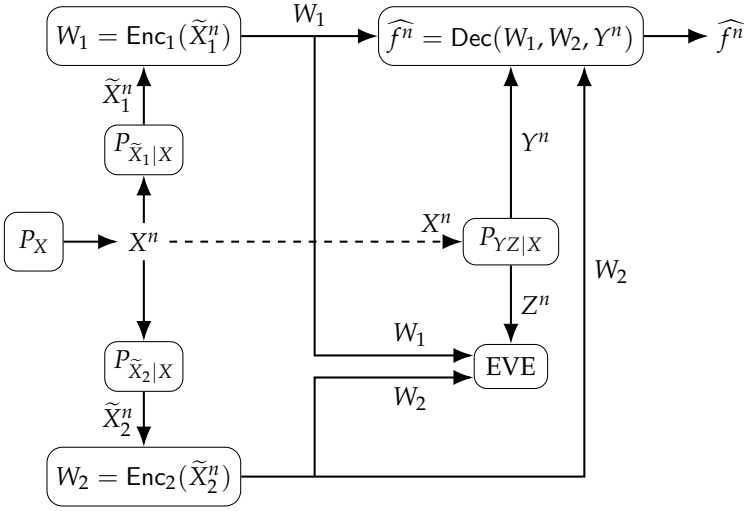


**Figure 1.** Single-function computation problem with two transmitting nodes under secrecy, privacy, and storage (or communication) constraints.

- The exact rate regions for invertible functions are further simplified when the eavesdropper's measurement channel is physically-degraded with respect to the fusion center's channel or vice versa, which result in different bounds on the rates.
- We evaluate the exact rate region for a physically-degraded case with multiplicative Bernoulli noise components.

### *1.3. Organization*

This paper is organized as follows. In Section 2, we introduce the lossless and lossy single-function computation problems with two transmitting nodes under secrecy, privacy, storage, and reliability or distortion constraints. In Section 3, we present the inner and outer bounds for the rate regions of the introduced problems and discuss that the bounds differ because of different Markov chains imposed. In Section 4, we characterize the exact lossless and lossy rate regions for invertible functions, partially-invertible functions, and two different degraded measurement channels, and the rate region for an example case is evaluated. In Section 5, we offer proofs of the inner and outer bounds for the lossless single-function computations with two transmitting nodes. In Section 6, we conclude the paper.

### *1.4. Notation*

Upper case letters represent random variables and lower case letters their realizations. A superscript denotes a sequence of variables, e.g., $X^n = X_1, X_2, \ldots, X_i, \ldots, X_n$, 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, set sizes are written as $|\mathcal{X}|$. Given any $a \in \mathbb{R}$, define $[a]^- = \min\{a, 0\}$. $H_b(c) = -c\log_2 c - (1-c)\log_2(1-c)$ is the binary entropy function for any $c \in [0,1]$.

## 2. System Model

We consider the single-function computation model with two transmitting nodes illustrated in Figure 1. Noisy measurements $\widetilde{X}_1^n$ and $\widetilde{X}_2^n$ of an i.i.d. remote source $X^n \sim P_X^n$ through memoryless channels $P_{\widetilde{X}_1|X}$ and $P_{\widetilde{X}_2|X}$, respectively, are observed by two legitimate nodes in a network. Similarly, other noisy measurements $Y^n$ and $Z^n$ of the same remote source are observed by the fusion center and eavesdropper (Eve), respectively, through another memoryless channel $P_{YZ|X}$. Encoders $\mathsf{Enc}_1(\cdot)$ and $\mathsf{Enc}_2(\cdot)$ of the legitimate nodes send indices $W_1$ and $W_2$, respectively, to the fusion center over public communication links with storage rate constraints. The fusion center decoder

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

Dec$(\cdot)$ then uses its observed noisy sequence $Y^n$ and the public indices $W_1$ and $W_2$ to estimate a function $f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)$ 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 source and measurement alphabets are finite sets.

A natural secrecy leakage constraint is to minimize the information leakage about the function output $f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)$ to eavesdropper. However, its analysis depends on the specific function $f(\cdot,\cdot,\cdot)$ computed, so we impose below another secrecy leakage constraint that does not depend on the function used and that provides an upper bound for secrecy leakage for all functions, as considered in [20,23]. Furthermore, we impose two privacy leakage constraints to minimize the information leakage about $X^n$ to the fusion center and eavesdropper because the same remote source would be measured if another function would be computed in the same network (see also [21] for motivations to consider privacy leakage with respect to a remote source) as well as public storage constraints that minimize the rate of storage for transmitting nodes.

We next define lossless and lossy single-function computation rate regions.

### *2.1. Lossless Single-Function Computation*

Consider the single-function computation model illustrated in Figure 1. The corresponding lossless rate region is defined as follows.

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

$$\Pr\left[f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n) \neq \widehat{f^n}\right] \leq \delta \qquad \text{(reliability)} \tag{2}$$

$$\frac{1}{n} I(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n; W_1, W_2 | Z^n) \leq R_s + \delta \qquad \text{(secrecy)} \tag{3}$$

$$\frac{1}{n} \log |\mathcal{W}_1| \leq R_{w,1} + \delta \qquad \text{(storage 1)} \tag{4}$$

$$\frac{1}{n} \log |\mathcal{W}_2| \leq R_{w,2} + \delta \qquad \text{(storage 2)} \tag{5}$$

$$\frac{1}{n} I(X^n; W_1, W_2 | Y^n) \leq R_{\ell,\text{Dec}} + \delta \qquad \text{(privacyDec)} \tag{6}$$

$$\frac{1}{n} I(X^n; W_1, W_2 | Z^n) \leq R_{\ell,\text{Eve}} + \delta \qquad \text{(privacyEve).} \tag{7}$$

The *lossless* region $\mathcal{R}$ is the closure of the set of all achievable lossless tuples. ◊

### *2.2. Lossy Single-Function Computation*

The corresponding lossy rate region for the single-function computation model illustrated in Figure 1 is defined as follows.

**Definition 2.** A *lossy* tuple $(R_s, R_{w,1}, R_{w,2}, R_{\ell,\text{Dec}}, R_{\ell,\text{Eve}}, D)$ is *achievable* if, for any $\delta > 0$, there exist $n \geq 1$, two encoders, and one decoder such that (3)-(7) and

$$\mathbb{E}\left[d(f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n), \widehat{f^n})\right] \leq D + \delta \qquad \text{(distortion)} \tag{8}$$

where

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$d(f^n, \widehat{f^n}) = \frac{1}{n}\sum_{i=1}^{n} d(f_i, \widehat{f_i}) \tag{9}$$

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

## 3. Inner and Outer Bounds

### 3.1. Lossless Single-Function Computation

We first extend the notion of *admissibility* defined in [8] for a single auxiliary random variable to two auxiliary random variables, used in the inner and outer bounds given below for lossless function computation; see also [20, Theorem 3].

**Definition 3.** A pair of (vector) random variables $(U_1, U_2)$ is admissible for a function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ if we have

$$H(f(\widetilde{X}_1, \widetilde{X}_2, Y)|\, U_1, U_2, Y) = 0 \tag{10}$$

and

$$U_1 - \widetilde{X}_1 - (\widetilde{X}_2, Y) \tag{11}$$
$$U_2 - \widetilde{X}_2 - (\widetilde{X}_1, Y) \tag{12}$$

form Markov chains. ◊

We next provide inner and outer bounds for the lossless region $\mathcal{R}$; see Section 5 for a proof sketch.

**Theorem 1.** (Inner Bound): *An achievable lossless 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* $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve})$ *such that* $(U_1, U_2)$ *pair is admissible for the function* $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ *and*

$$R_s \geq \Big[I(U_1, U_2; Z|V_1, V_2, Q) - I(U_1, U_2; Y|V_1, V_2, Q)\Big]^{-} + I(U_1, U_2; \widetilde{X}_1, \widetilde{X}_2|Z) \tag{13}$$
$$R_{w,1} \geq I(V_1; \widetilde{X}_1|V_2, Y) + I(U_1; \widetilde{X}_1|V_1, U_2, Y) \tag{14}$$
$$R_{w,2} \geq I(V_2; \widetilde{X}_2|V_1, Y) + I(U_2; \widetilde{X}_2|U_1, V_2, Y) \tag{15}$$
$$\begin{aligned} R_{w,1} + R_{w,2} \geq\ & I(U_2; \widetilde{X}_2|U_1, V_2, Y) + I(U_1; \widetilde{X}_1|V_1, V_2, Y) \\ & + I(V_2; \widetilde{X}_2|V_1, Y) + I(V_1; \widetilde{X}_1|Y) \end{aligned} \tag{16}$$
$$R_{\ell,Dec} \geq I(U_1, U_2; X|Y) \tag{17}$$
$$R_{\ell,Eve} \geq \Big[I(U_1, U_2; Z|V_1, V_2, Q) - I(U_1, U_2; Y|V_1, V_2, Q)\Big]^{-} + I(U_1, U_2; X|Z) \tag{18}$$

*where we have*

$$P_{QV_1V_2U_1U_2\widetilde{X}_1\widetilde{X}_2XYZ} = P_{Q|V_1V_2}P_{V_1|U_1}P_{U_1|\widetilde{X}_1}P_{\widetilde{X}_1|X}P_{V_2|U_2}P_{U_2|\widetilde{X}_2}P_{\widetilde{X}_2|X}P_XP_{YZ|X}. \tag{19}$$

(Outer Bound): *An outer bound for the lossless region* $\mathcal{R}$ *is the union of the rate tuples in (13), (16)-(18), and*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$R_{w,1} \geq I(V_1; \widetilde{X}_1|V_2, Y) + I(U_1; \widetilde{X}_1|V_1, U_2, Y) - I(V_1; V_2|\widetilde{X}_1, Y) - I(U_1; U_2|\widetilde{X}_1, Y, V_1) \quad (20)$$

$$R_{w,2} \geq I(V_2; \widetilde{X}_2|V_1, Y) + I(U_2; \widetilde{X}_2|U_1, V_2, Y) - I(V_2; V_1|\widetilde{X}_2, Y) - I(U_2; U_1|\widetilde{X}_2, Y, V_2) \quad (21)$$

*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* $P_{\widetilde{X}_2|U_2}$ *such that* $(U_1, U_2)$ *pair is admissible for the function* $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ *and*

$$(Q, V_1) - U_1 - \widetilde{X}_1 - X - (\widetilde{X}_2, Y, Z) \quad (22)$$

$$(Q, V_2) - U_2 - \widetilde{X}_2 - X - (\widetilde{X}_1, Y, Z) \quad (23)$$

*form Markov chains. One can limit the cardinalities to* $|\mathcal{Q}| \leq 2$, $|\mathcal{V}_1| \leq |\widetilde{X}_1| + 6$, $|\mathcal{V}_2| \leq |\widetilde{X}_2| + 6$, $|\mathcal{U}_1| \leq (|\widetilde{X}_1| + 6)^2$, *and* $|\mathcal{U}_2| \leq (|\widetilde{X}_2| + 6)^2$.

We remark that if the joint probability distribution in (19) is imposed on the outer bound, (20) and (21) recover (14) and (15), respectively, because then

$$(V_1, U_1) - \widetilde{X}_1 - (Y, U_2, V_2) \quad (24)$$

$$(V_2, U_2) - \widetilde{X}_2 - (Y, U_1, V_1) \quad (25)$$

form Markov chains for (19). However, the outer bound that satisfies (22) and (23) defines a rate region that is in general larger than the rate region defined by the inner bound that satisfies (19). Thus, inner and outer bounds generally differ. The results in Theorem 1 recovers previous results including [20, Theorem 3] and, naturally, also other results that are recovered by these previous results such as the SW coding region.

### *3.2. Lossy Single-Function Computation*

We next provide inner and outer bounds for the lossy region $\mathcal{R}_D$; see below for a proof sketch.

**Theorem 2.** (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 (13)-(18) and*

$$D \geq \mathbb{E}[d(f(\widetilde{X}_1, \widetilde{X}_2, Y), g(U_1, U_2, Y))] \quad (26)$$

*for some function* $g(\cdot, \cdot, \cdot)$ *and where* $P_{QV_1V_2U_1U_2\widetilde{X}_1\widetilde{X}_2XYZ}$ *is equal to (19).*

(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* $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve}, D)$ *in (13), (16)-(18), (20), (21), and (26) such that (22) and (23) 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* $|\mathcal{U}_2| \leq (|\widetilde{X}_2| + 7)^2$.

**Proof Sketch.** The achievability proof of the lossy function computation problem follows from the achievability proof of its lossless version given in Section 5.1 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 $g(U_1, U_2, Y)$ that satisfies

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$g^n(U_1^n, U_2^n, Y^n) = \{g(U_{1,i}, U_{2,i}, Y_i)\}_{i=1}^n \tag{27}$$

$$\mathbb{E}[d(f^n(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n), g^n(U_1^n, U_2^n, Y^n))] \leq D + \epsilon_n \tag{28}$$

where $\epsilon_n > 0$ such that $\epsilon_n \to 0$ when $n \to \infty$. 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 [33, pp. 26], constraint in (8) is satisfied.

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 function computation problem given in Section 5.2. However, the proof for the lossless function computation problem requires the auxiliary random variables to be admissible as defined in Definition 3, unlike the lossy function computation problem. Thus, the outer bound proof for Theorem 2 follows by replacing the admissibility step (96) in the outer bound proof for the lossless function computation problem with the step

$$\begin{aligned}
&n(D+\delta_n)\\
&\overset{(a)}{\geq} \mathbb{E}\Big[\sum_{i=1}^n d\Big(f_i(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}, Y_i), \widehat{f}_i(W_1, W_2, Y^n)\Big)\Big]\\
&\overset{(b)}{\geq} \mathbb{E}\Big[\sum_{i=1}^n d\Big(f_i(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}, Y_i), g_i(W_1, W_2, Y^n, X^{i-1}, Z^{i-1})\Big)\Big]\\
&\overset{(c)}{=} \mathbb{E}\Big[\sum_{i=1}^n d\Big(f_i(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}, Y_i), g_i(W_1, W_2, Y_i^n, X^{i-1}, Z^{i-1})\Big)\Big]\\
&\overset{(d)}{=} \mathbb{E}\Big[\sum_{i=1}^n d\Big(f(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}, Y_i), g(U_{1,i}, U_{2,i}, Y_i)\Big)\Big]
\end{aligned} \tag{29}$$

where $(a)$ follows by (8) and (9), $(b)$ 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 given in (100), and $(d)$ follows from the definitions of $U_{1,i}$ and $U_{2,i}$ given in (91) and (92), respectively. Furthermore, the proof of the cardinality bounds for the lossy case follows from the proof for the lossless case 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

$$(V_1, V_2) - (U_1, U_2, Y) - g(U_1, U_2, Y). \tag{30}$$

□

Entirely similar to Theorem 1, the inner and outer bounds given in Theorem 2 do not match in general because of different Markov chains imposed.

**Remark 1.** Since all secrecy and privacy rate terms given in the outer bounds in Theorems 1 and 2, i.e., lower bounds in (13), (17), and (18), are generally strictly positive, strong secrecy or strong privacy constraints cannot be satisfied in general for the lossless and lossy single-function computation problems.

We next provide the exact rate regions, i.e., rate regions for which inner and outer bounds match, for various sets of computed functions $f(\cdot,\cdot,\cdot)$ and measurement channels $P_{YZ|X}$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

## 4. Rate Regions for Special Sets of Computed Functions and Measurement Channels

The terms that characterize the exact rate regions of the lossless and lossy function computation problems for various sets of functions and channels are the same, except 1) removal of the admissibility requirement; 2) addition of a distortion constraint; 3) increase in the cardinality bounds on the auxiliary random variables for the lossy case as compared to the lossless case. Thus, we provide the exact rate regions only for the lossless case. However, we remark that the optimal auxiliary random variables for lossless and lossy cases might differ. Therefore, the corresponding lossless and lossy rate regions might look different for the same joint probability distribution $P_{\widetilde{X}_1\widetilde{X}_2XYZ}$.

### 4.1. Partially-Invertible Functions

We now 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 [11,34]

$$H(\widetilde{X}_1|f(\widetilde{X}_1, \widetilde{X}_2, Y), Y) = 0. \tag{31}$$

For such functions, it is straightforward to show that we have the following exact rate region for the lossless function computation problem with two transmitting nodes. The proof of Lemma 1 follows from Theorem 1 by assigning $U_1 = \widetilde{X}_1$ and constant $V_1$, and then by applying the Markov chain (23) to (13). Furthermore, by symmetry the exact lossless rate region 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 $U_2 = \widetilde{X}_2$ and constant $V_2$, and then applying (22) to ( 13).

**Lemma 1.** *The lossless rate region $\mathcal{R}$ when $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is a partially-invertible function with respect to $\widetilde{X}_1$ is the set of all tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve})$ such that $U_2$ is admissible for the function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ and*

$$R_s \geq \Big[I(\widetilde{X}_1, U_2; Z|V_2, Q) - I(\widetilde{X}_1, U_2; Y|V_2, Q)\Big]^- + H(\widetilde{X}_1|U_2, Z) + I(U_2; \widetilde{X}_2|Z) \tag{32}$$

$$R_{w,1} \geq H(\widetilde{X}_1|U_2, Y) \tag{33}$$

$$R_{w,2} \geq I(V_2; \widetilde{X}_2|Y) + I(U_2; \widetilde{X}_2|\widetilde{X}_1, V_2, Y) \tag{34}$$

$$R_{w,1} + R_{w,2} \geq I(U_2; \widetilde{X}_2|\widetilde{X}_1, V_2, Y) + H(\widetilde{X}_1|V_2, Y) + I(V_2; \widetilde{X}_2|Y) \tag{35}$$

$$R_{\ell,Dec} \geq I(\widetilde{X}_1, U_2; X|Y) \tag{36}$$

$$R_{\ell,Eve} \geq \Big[I(\widetilde{X}_1, U_2; Z|V_2, Q) - I(\widetilde{X}_1, U_2; Y|V_2, Q)\Big]^- + I(\widetilde{X}_1, U_2; X|Z) \tag{37}$$

*such that (23) form a Markov chain. One can limit the cardinalities to $|\mathcal{Q}| \leq 2$, $|\mathcal{V}_2| \leq |\widetilde{\mathcal{X}}_2| + 6$, and $|\mathcal{U}_2| \leq (|\widetilde{\mathcal{X}}_2| + 6)^2$.*

### 4.2. Invertible Functions

Suppose now we impose the condition that the function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is *invertible*, i.e., we have [11,34]

$$H(\widetilde{X}_1, \widetilde{X}_2|f(\widetilde{X}_1, \widetilde{X}_2, Y), Y) = 0. \tag{38}$$

We provide in Lemma 2 below the exact rate region for the lossless function computation problem with two transmitting nodes when the function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is invertible. The proof of Lemma 2 follows from Theorem 1 by assigning $U_1 = \widetilde{X}_1$, $U_2 = \widetilde{X}_2$, and constant $V_1$ and $V_2$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

**Lemma 2.** *The lossless rate region $\mathcal{R}$ when $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is an invertible function is the set of all tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve})$ satisfying*

$$R_s \geq \left[I(\widetilde{X}_1, \widetilde{X}_2; Z|Q) - I(\widetilde{X}_1, \widetilde{X}_2; Y|Q)\right]^- + H(\widetilde{X}_1, \widetilde{X}_2|Z) \tag{39}$$
$$R_{w,1} \geq H(\widetilde{X}_1|\widetilde{X}_2, Y) \tag{40}$$
$$R_{w,2} \geq H(\widetilde{X}_2|\widetilde{X}_1, Y) \tag{41}$$
$$R_{w,1} + R_{w,2} \geq H(\widetilde{X}_1, \widetilde{X}_2|Y) \tag{42}$$
$$R_{\ell,Dec} \geq I(\widetilde{X}_1, \widetilde{X}_2; X|Y) \tag{43}$$
$$R_{\ell,Eve} \geq \left[I(\widetilde{X}_1, \widetilde{X}_2; Z|Q) - I(\widetilde{X}_1, \widetilde{X}_2; Y|Q)\right]^- + I(\widetilde{X}_1, \widetilde{X}_2; X|Z) \tag{44}$$

*where $Q - (\widetilde{X}_1, \widetilde{X}_2) - X - (Y, Z)$ form a Markov chain. One can limit the cardinality to $|\mathcal{Q}| \leq 2$.*

### *4.3. Invertible Functions and Two Different Degraded Channels*

The lossless rate region given in Lemma 2 can be further simplified by imposing conditions on the measurement channel $P_{YZ|X}$ in addition to the function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ being invertible. We next characterize the lossless rate regions for two different physically-degraded channels.

#### 4.3.1. Eve's Channel is Physically-Degraded

Suppose the measurement channel $P_{YZ|X}$ is physically-degraded such that

$$P_{YZ|X} = P_{Y|X} P_{Z|Y}. \tag{45}$$

For invertible functions and physically-degraded measurement channels $P_{YZ|X}$ as defined in (45), we provide the exact lossless rate region in Lemma 3. The proof of Lemma 3 follows from Lemma 2 and by using the following Markov chain for this case

$$(\widetilde{X}_1, \widetilde{X}_2) - X - Y - Z \tag{46}$$

which follows by (45).

**Lemma 3.** *The lossless rate region $\mathcal{R}$ when $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is an invertible function and $P_{YZ|X}$ is as given in (45) is the set of all tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve})$ satisfying (40)-(43) and*

$$R_s \geq H(\widetilde{X}_1, \widetilde{X}_2|Y) \tag{47}$$
$$R_{\ell,Eve} \geq I(\widetilde{X}_1, \widetilde{X}_2; X|Y). \tag{48}$$

#### 4.3.2. Fusion Center's Channel is Physically-Degraded

Suppose the measurement channel $P_{YZ|X}$ is physically-degraded such that

$$P_{YZ|X} = P_{Z|X} P_{Y|Z}. \tag{49}$$

For invertible functions and physically-degraded measurement channels $P_{YZ|X}$ as defined in (49), we provide the exact lossless rate region in Lemma 4. The proof of Lemma 4 follows from Lemma 2 and by using the following Markov chain for this case

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$(\widetilde{X}_1, \widetilde{X}_2) - X - Z - Y \tag{50}$$

which follows by (49).

**Lemma 4.** *The lossless rate region $\mathcal{R}$ when $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is an invertible function and $P_{YZ|X}$ is as given in (49) is the set of all tuples $(R_s, R_{w,1}, R_{w,2}, R_{\ell,Dec}, R_{\ell,Eve})$ satisfying (40)-(43) and*

$$R_s \geq H(\widetilde{X}_1, \widetilde{X}_2|Z) \tag{51}$$
$$R_{\ell,Eve} \geq I(\widetilde{X}_1, \widetilde{X}_2; X|Z). \tag{52}$$

**Remark 2.** The rate regions given in Lemmas 2-4 can be plotted by computing the terms that characterize the regions since $P_{\widetilde{X}_1\widetilde{X}_2XYZ}$ is fixed for function computation problems considered. However, the rate region given in Lemma 1, similar to the inner bounds given in Theorems 1 and 2, might not be easy to characterize due to the requirement to optimize the auxiliary random variables whose cardinalities are bounded by large terms. Thus, evaluating the rate region for a function computation problem with two transmitting terminals is generally significantly more difficult than characterization of the rate region for function computation with one transmitting terminal; see [23] for an information bottleneck example for the latter problem.

We next evaluate the lossless rate region $\mathcal{R}$ by using Lemma 4 for specific measurement channels when $f(\widetilde{X}_1, \widetilde{X}_2, Y)$ is an invertible function.

### *4.4. Lossless Rate Region Example*

Suppose measurement channels in Figure 1 have binary input and output alphabets with multiplicative Bernoulli noise components, i.e., we have $\mathcal{X} = \widetilde{\mathcal{X}}_1 = \widetilde{\mathcal{X}}_2 = \mathcal{Z} = \mathcal{Y} = \mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}_Z = \mathcal{S}_Y = \{0, 1\}$ and

$$\widetilde{X}_1 = S_1 \cdot X, \qquad \widetilde{X}_2 = S_2 \cdot X, \qquad Z = S_Z \cdot X, \qquad Y = S_Y \cdot X \tag{53}$$

where $S_1$, $S_2$, $X$, and $(S_Z, S_Y)$ are mutually independent, and we have $P_X(1) = 0.5$, $P_{S_1}(1) = \beta_1$, $P_{S_2}(1) = \beta_2$, $P_{S_ZS_Y}(0,0) = (1-q)$, $P_{S_ZS_Y}(1,1) = q\alpha$, and $P_{S_ZS_Y}(1,0) = q(1-\alpha)$ for fixed $\beta_1, \beta_2, q, \alpha \in [0, 1]$, so (49) is satisfied; see also [35, Section IV-A]. Using Lemma 4 for the given probability distributions, we evaluate the lossless rate region $\mathcal{R}$ for an invertible function computation scenario with two transmitting nodes, in which, e.g., $\beta_1 = 0.2$, $\beta_2 = 0.11$, $\alpha = 0.3$, and $q = 0.25$ and obtain the lossless rate region that is characterized by

$$R_{\mathrm{s}} \geq 0.7579 \text{ bits/symbol}, \qquad R_{\mathrm{w},1} \geq 0.4626 \text{ bits/symbol}, \tag{54}$$
$$R_{\mathrm{w},2} \geq 0.3021 \text{ bits/symbol}, \qquad R_{\mathrm{w},1} + R_{\mathrm{w},2} \geq 0.7686 \text{ bits/symbol}, \tag{55}$$
$$R_{\ell,\mathrm{Dec}} \geq 0.1577 \text{ bits/symbol}, \qquad R_{\ell,\mathrm{Eve}} \geq 0.1469 \text{ bits/symbol} \tag{56}$$

where the sum-storage rate constraint is active since the sum of the bounds on $R_{\mathrm{w},1}$ and $R_{\mathrm{w},2}$ is smaller than the bound on $(R_{\mathrm{w},1} + R_{\mathrm{w},2})$.

## 5. Proof of Theorem 1

### *5.1. Inner Bound*

**Proof Sketch.** The OSRB method [32] is used for the proof of achievability by applying the steps given in [36, Section 1.6]. Let

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$(V_1^n, V_2^n, U_1^n, U_2^n, \widetilde{X}_1^n, \widetilde{X}_2^n, X^n, Y^n, Z^n) \tag{57}$$

be i.i.d. according to $P_{V_1V_2U_1U_2\widetilde{X}_1\widetilde{X}_2XYZ}$ that can be obtained from (19) with fixed $P_{U_1|\widetilde{X}_1}$, $P_{V_1|U_1}$, $P_{U_2|\widetilde{X}_2}$, and $P_{V_2|U_2}$ such that the pair $(U_1, U_2)$ is admissible for a function $f(\widetilde{X}_1, \widetilde{X}_2, Y)$, so $(U_1^n, U_2^n)$ is also admissible since random variables in (57) are i.i.d.

To each $v_1^n$ assign two random bin indices $(F_{\mathrm{v}_1}, W_{\mathrm{v}_1})$ such that $F_{\mathrm{v}_1} \in [1 : 2^{n\widetilde{R}_{\mathrm{v}_1}}]$ and $W_{\mathrm{v}_1} \in [1 : 2^{nR_{\mathrm{v}_1}}]$. Furthermore, to each $u_1^n$ assign two random indices $(F_{\mathrm{u}_1}, W_{\mathrm{u}_1})$ such that $F_{\mathrm{u}_1} \in [1 : 2^{n\widetilde{R}_{\mathrm{u}_1}}]$ and $W_{\mathrm{u}_1} \in [1 : 2^{nR_{\mathrm{u}_1}}]$. Similarly, random indices $(F_{\mathrm{v}_2}, W_{\mathrm{v}_2})$ and $(F_{\mathrm{u}_2}, W_{\mathrm{u}_2})$ are assigned to each $v_2^n$ and $u_2^n$, respectively. The indices $F_1 = (F_{\mathrm{v}_1}, F_{\mathrm{u}_1})$, and $F_2 = (F_{\mathrm{v}_2}, F_{\mathrm{u}_2})$ represent the public choice of two encoders and one decoder, whereas $W_1 = (W_{\mathrm{v}_1}, W_{\mathrm{u}_1})$ and $W_2 = (W_{\mathrm{v}_2}, W_{\mathrm{u}_2})$ are the public messages sent by the encoders $\mathsf{Enc}_1(\cdot)$ and $\mathsf{Enc}_2(\cdot)$, respectively, to the fusion center.

We consider the following decoding order:

1. observing $(Y^n, F_{\mathrm{v}_1}, W_{\mathrm{v}_1})$, the decoder $\mathsf{Dec}(\cdot)$ estimates $V_1^n$ as $\widehat{V}_1^n$;
2. observing $(Y^n, \widehat{V}_1^n, F_{\mathrm{v}_2}, W_{\mathrm{v}_2})$, the decoder estimates $V_2^n$ as $\widehat{V}_2^n$;
3. observing $(Y^n, \widehat{V}_1^n, \widehat{V}_2^n, F_{\mathrm{u}_1}, W_{\mathrm{u}_1})$, the decoder estimates $U_1^n$ as $\widehat{U}_1^n$;
4. observing $(Y^n, \widehat{V}_1^n, \widehat{V}_2^n, \widehat{U}_1^n, F_{\mathrm{u}_2}, W_{\mathrm{u}_2})$, the decoder estimates $U_2^n$ as $\widehat{U}_2^n$.

By swapping indices 1 and 2 in the decoding order another corner point in the achievable rate region is obtained, so we analyze the given decoding order but also provide the results for the other corner point.

Consider Step 1 in the decoding order given above. Using a SW [15] decoder, one can reliably estimate $V_1^n$ from $(Y^n, F_{\mathrm{v}_1}, W_{\mathrm{v}_1})$ such that the expected value of the error probability taken over the random bin assignments vanishes when $n \to \infty$, if we have [32, Lemma 1]

$$\widetilde{R}_{\mathrm{v}_1} + R_{\mathrm{v}_1} > H(V_1|Y). \tag{58}$$

Similarly, Step 2, 3, and 4 estimations are reliable if we have

$$\widetilde{R}_{\mathrm{v}_2} + R_{\mathrm{v}_2} > H(V_2|V_1, Y) \tag{59}$$

$$\widetilde{R}_{\mathrm{u}_1} + R_{\mathrm{u}_1} > H(U_1|V_1, V_2, Y) \tag{60}$$

$$\widetilde{R}_{\mathrm{u}_2} + R_{\mathrm{u}_2} > H(U_2|V_1, V_2, U_1, Y) \overset{(a)}{=} H(U_2|V_2, U_1, Y) \tag{61}$$

where $(a)$ follows from the Markov chain $V_1 - U_1 - (U_2, V_2, Y)$. Therefore, (2) is satisfied if (58)-(61) are satisfied.

The public index $F_{\mathrm{v}_1}$ is almost independent of $\widetilde{X}_1^n$, so it is almost independent of $(\widetilde{X}_1^n, \widetilde{X}_2^n, X^n, Y^n, Z^n)$, if we have [32, Theorem 1]

$$\widetilde{R}_{\mathrm{v}_1} < H(V_1|\widetilde{X}_1) \tag{62}$$

because then the expected value, which is taken over the random bin assignments, of the variational distance between the joint probability distributions $\mathrm{Unif}[1{:}2^{n\widetilde{R}_{\mathrm{v}_1}}] \cdot P_{\widetilde{X}_1^n}$ and $P_{F_{\mathrm{v}_1}\widetilde{X}_1^n}$ vanishes when $n \to \infty$. Furthermore, the public index $F_{\mathrm{u}_1}$ is almost independent of $(V_1^n, \widetilde{X}_1^n)$, so it is almost independent of $(V_1^n, \widetilde{X}_1^n, \widetilde{X}_2^n, X^n, Y^n, Z^n)$, if we have

$$\widetilde{R}_{\mathrm{u}_1} < H(U_1|V_1, \widetilde{X}_1). \tag{63}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

Similarly, $F_{\mathrm{v}_2}$ is almost independent of $\widetilde{X}_2^n$ if we have

$$\widetilde{R}_{\mathrm{v}_2} < H(V_2|\widetilde{X}_2) \tag{64}$$

and $F_{\mathrm{u}_2}$ is almost independent of $(V_2^n, \widetilde{X}_2^n)$ if we have

$$\widetilde{R}_{\mathrm{u}_2} < H(U_2|V_2, \widetilde{X}_2). \tag{65}$$

To satisfy (58)-(65), for any $\epsilon > 0$ we fix

$$\begin{aligned}
\widetilde{R}_{\mathrm{v}_1} &= H(V_1|\widetilde{X}_1) - \epsilon && (66)\\
R_{\mathrm{v}_1} &= I(V_1;\widetilde{X}_1) - I(V_1;Y) + 2\epsilon && (67)\\
\widetilde{R}_{\mathrm{v}_2} &= H(V_2|\widetilde{X}_2) - \epsilon && (68)\\
R_{\mathrm{v}_2} &= I(V_2;\widetilde{X}_2) - I(V_2;V_1,Y) + 2\epsilon && (69)\\
\widetilde{R}_{\mathrm{u}_1} &= H(U_1|V_1,\widetilde{X}_1) - \epsilon && (70)\\
R_{\mathrm{u}_1} &= I(U_1;\widetilde{X}_1|V_1) - I(U_1;V_2,Y|V_1) + 2\epsilon && (71)\\
\widetilde{R}_{\mathrm{u}_2} &= H(U_2|V_2,\widetilde{X}_2) - \epsilon && (72)\\
R_{\mathrm{u}_2} &= I(U_2;\widetilde{X}_2|V_2) - I(U_2;U_1,Y|V_2) + 2\epsilon. && (73)
\end{aligned}$$

**Public Message (Storage) Rates**: (67) and (71) result in a public message (storage) rate $R_{\mathrm{w}_1}$ of

$$\begin{aligned}
R_{\mathrm{w}_1} &= R_{\mathrm{v}_1} + R_{\mathrm{u}_1}\\
&\overset{(a)}{=} I(V_1;\widetilde{X}_1|Y) + H(U_1|V_1,V_2,Y) - H(U_1|V_1,\widetilde{X}_1) + 4\epsilon\\
&\overset{(b)}{=} I(V_1;\widetilde{X}_1|Y) + I(U_1;\widetilde{X}_1|V_1,V_2,Y) + 4\epsilon
\end{aligned} \tag{74}$$

where $(a)$ follows because $V_1 - \widetilde{X}_1 - Y$ form a Markov chain and $(b)$ follows because $U_1 - (V_1, \widetilde{X}_1) - (V_2, Y)$ form a Markov chain. Furthermore, (69) and (73) result in a storage rate $R_{\mathrm{w}_2}$ of

$$\begin{aligned}
R_{\mathrm{w}_2} &= R_{\mathrm{v}_2} + R_{\mathrm{u}_2}\\
&\overset{(a)}{=} I(V_2;\widetilde{X}_2|V_1,Y) + H(U_2|U_1,V_2,Y) - H(U_2|V_2,\widetilde{X}_2) + 4\epsilon\\
&\overset{(b)}{=} I(V_2;\widetilde{X}_2|V_1,Y) + I(U_2;\widetilde{X}_2|U_1,V_2,Y) + 4\epsilon
\end{aligned} \tag{75}$$

where $(a)$ follows from the Markov chain $V_2 - \widetilde{X}_2 - (V_1, Y)$ and $(b)$ from $U_2 - (V_2, \widetilde{X}_2) - (U_1, Y)$. We remark that if the indices 1 and 2 in the decoding order given above are swapped, the other corner point with

$$\begin{aligned}
R'_{\mathrm{w}_1} &= I(V_1;\widetilde{X}_1|V_2,Y) + I(U_1;\widetilde{X}_1|U_2,V_1,Y) + 4\epsilon && (76)\\
R'_{\mathrm{w}_2} &= I(V_2;\widetilde{X}_2|Y) + I(U_2;\widetilde{X}_2|V_1,V_2,Y) + 4\epsilon && (77)
\end{aligned}$$

is achieved.

**Privacy Leakage to Decoder**: We have

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$
\begin{aligned}
&I(X^n; W_1, W_2, F_1, F_2|Y^n) \\
&= I(X^n; W_1, W_2|F_1, F_2, Y^n) + I(X^n; F_1, F_2|Y^n) \\
&\overset{(a)}{\leq} H(X^n|Y^n) - H(X^n|W_1, W_2, F_1, F_2, V_1^n, V_2^n, U_1^n, U_2^n, Y^n) + 4\epsilon_n \\
&\overset{(b)}{=} H(X^n|Y^n) - H(X^n|U_1^n, U_2^n, Y^n) + 4\epsilon_n \\
&\overset{(c)}{=} nI(U_1, U_2; X|Y) + 4\epsilon_n
\end{aligned} \tag{78}
$$

where
$(a)$ follows for some $\epsilon_n > 0$ with $\epsilon_n \to 0$ when $n \to \infty$ because

$$
\begin{aligned}
&I(X^n; F_1, F_2|Y^n) \\
&= I(X^n; F_{\mathrm{v}_1}|Y^n) + I(X^n; F_{\mathrm{u}_1}|F_{\mathrm{v}_1}, Y^n) + I(X^n; F_{\mathrm{v}_2}|F_{\mathrm{v}_1}, F_{\mathrm{u}_1}, Y^n) \\
&\qquad + I(X^n; F_{\mathrm{u}_2}|F_{\mathrm{v}_1}, F_{\mathrm{u}_1}, F_{\mathrm{v}_2}, Y^n) \\
&\leq 4\epsilon_n
\end{aligned} \tag{79}
$$

since 1) by (62) $F_{\mathrm{v}_1}$ is almost independent of $(X^n, Y^n)$; 2) by (63) $F_{\mathrm{u}_1}$ is almost independent of $(V_1^n, X^n, Y^n)$ and because $V_1^n$ determines $F_{\mathrm{v}_1}$; 3) by (64) $F_{\mathrm{v}_2}$ is almost independent of $(U_1^n, V_1^n, X^n, Y^n)$ and because $(V_1^n, U_1^n)$ determine $(F_{\mathrm{v}_1}, F_{\mathrm{u}_1})$; 4) by (65) $F_{\mathrm{u}_2}$ is almost independent of $(V_2^n, U_1^n, V_1^n, X^n, Y^n)$ and because $(V_1^n, U_1^n, V_2^n)$ determine $(F_{\mathrm{v}_1}, F_{\mathrm{u}_1}, F_{\mathrm{v}_2})$;
$(b)$ follows because $(V_1^n, V_2^n, U_1^n, U_2^n)$ determine $(W_1, W_2, F_1, F_2)$ and from the Markov chains $V_1^n - U_1^n - (X^n, Y^n, U_2^n, V_2^n)$ and $V_2^n - U_2^n - (X^n, Y^n, U_1^n)$;
$(c)$ follows because $(X^n, U_1^n, U_2^n, Y^n)$ are i.i.d.

**Privacy Leakage to Eve**: We have

$$
\begin{aligned}
&I(X^n; W_1, W_2, F_1, F_2|Z^n) \\
&\overset{(a)}{=} H(W_1, W_2, F_1, F_2|Z^n) - H(W_1, W_2, F_1, F_2|X^n) \\
&\overset{(b)}{=} H(W_1, W_2, F_1, F_2|Z^n) - H(W_{u_1}, F_{u_1}, W_{u_2}, F_{u_2}, V_1^n, V_2^n|X^n) \\
&\quad + H(V_1^n|W_1, W_2, F_1, F_2, X^n) + H(V_2^n|V_1^n, W_1, W_2, F_1, F_2, X^n) \\
&\overset{(c)}{\leq} H(W_1, W_2, F_1, F_2|Z^n) - H(W_{u_1}, F_{u_1}, W_{u_2}, F_{u_2}, V_1^n, V_2^n|X^n) + 2n\epsilon_n' \\
&\overset{(d)}{=} H(W_1, W_2, F_1, F_2|Z^n) - H(U_1^n, U_2^n, V_1^n, V_2^n|X^n) \\
&\quad + H(U_1^n|W_{u_1}, F_{u_1}, W_{u_2}, F_{u_2}, V_1^n, V_2^n, X^n) \\
&\quad + H(U_2^n|U_1^n, W_{u_1}, F_{u_1}, W_{u_2}, F_{u_2}, V_1^n, V_2^n, X^n) + 2n\epsilon_n' \\
&\overset{(e)}{\leq} H(W_1, W_2, F_1, F_2|Z^n) - H(U_1^n, U_2^n, V_1^n, V_2^n|X^n) + 4n\epsilon_n' \\
&\overset{(f)}{=} H(W_1, W_2, F_1, F_2|Z^n) - nH(U_1, U_2, V_1, V_2|X) + 4n\epsilon_n'
\end{aligned} \tag{80}
$$

where $(a)$ follows because $(W_1, W_2, F_1, F_2) - X^n - Z^n$ form a Markov chain, $(b)$ follows since $(V_1^n, V_2^n)$ determine $(F_{v_1}, W_{v_1}, F_{v_2}, W_{v_2})$, $(c)$ follows for some $\epsilon_n' > 0$ such that $\epsilon_n' \to 0$ when $n \to \infty$ because $(F_{v_1}, W_{v_1}, X^n)$ can reliably recover $V_1^n$ by (58), and similarly because $(F_{v_2}, W_{v_2}, V_1^n, X^n)$ can reliably recover $V_2^n$ by (59) both due to the Markov chain $(V_1^n, V_2^n) - X^n - Y^n$, $(d)$ follows because $(U_1^n, U_2^n)$ determine $(F_{u_1}, W_{u_1}, F_{u,2}, W_{u_2})$, $(e)$ follows because $(F_{u_1}, W_{u_1}, V_1^n, V_2^n, X^n)$ can reliably recover $U_1^n$ by (60) and the inequality

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$H(U_1|V_1,V_2,Y) \geq H(U_1|V_1,V_2,X) \tag{81}$$

that follows from

$$I(U_1;V_1,V_2,X) - I(U_1;V_1,V_2,Y) \geq I(U_1;V_1,V_2,X) - I(U_1;V_1,V_2,Y,X) = 0 \tag{82}$$

since $U_1 - (V_1,V_2,X) - Y$ form a Markov chain. Furthermore, $(F_{u_2}, W_{u_2}, V_1^n, V_2^n, U_1^n, X^n)$ can reliably recover $U_2^n$ by (61) and the inequality

$$H(U_2|V_1,V_2,U_1,Y) \geq H(U_2|V_1,V_2,U_1,X) \tag{83}$$

that can be proved entirely similarly to (82) by using the Markov chain $U_2 - (V_1,V_2,U_1,X) - Y$, and $(f)$ follows because $(U_1^n, U_2^n, V_1^n, V_2^n, X^n)$ are i.i.d.

In (80), obtaining single letter bounds on the term $H(W_1,W_2,F_1,F_2|Z^n)$ requires analysis of numerous decodability cases, whereas there are only six different decodability cases analyzed in [23] for secure function computation with a single transmitting node. To simplify our analysis by applying the results in [23], we combine the decoding order Steps 1 and 2 given above such that $(V_1,V_2)$ are treated jointly and, similarly, we combine Steps 3 and 4 such that $(U_1,U_2)$ are treated jointly. Using the combined steps, we can consider the six decodability cases analyzed in [23, Section V-A] by replacing $V^n$ with $(V_1^n,V_2^n)$ and $U^n$ with $(U_1^n,U_2^n)$, respectively, in the proof. Since in (80) the second term $-nH(U_1,U_2,V_1,V_2|X)$ can be obtained by applying the same replacement to the second term in [23, Eq. (54)], we obtain from (80) and these decodability analyses that

$$\begin{aligned} &I(X^n;W_1,W_2,F_1,F_2|Z^n) \\ &\leq n([I(U_1,U_2;Z|V_1,V_2) - I(U_1,U_2;Y|V_1,V_2) + \epsilon]^- \\ &\qquad + I(U_1,U_2;X|Z) + 4\epsilon_n' + \epsilon_n'') \end{aligned} \tag{84}$$

for some $\epsilon_n'' > 0$ such that $\epsilon_n'' \to 0$ when $n \to \infty$.

**Secrecy Leakage (to Eve)**: We obtain

$$\begin{aligned} &I(\widetilde{X}_1^n,\widetilde{X}_2^n,Y^n;W_1,W_2,F_1,F_2|Z^n) \\ &\overset{(a)}{=} H(W_1,W_2,F_1,F_2|Z^n) - H(W_1,W_2,F_1,F_2|\widetilde{X}_1^n,\widetilde{X}_2^n) \\ &\overset{(b)}{=} H(W_1,W_2,F_1,F_2|Z^n) - H(W_{u_1},W_{u_2},F_{u_1},F_{u_2},V_1^n,V_2^n|\widetilde{X}_1^n,\widetilde{X}_2^n) \\ &\quad + H(V_1^n|W_1,W_2,F_1,F_2,\widetilde{X}_1^n,\widetilde{X}_2^n) + H(V_2^n|V_1^n,W_1,W_2,F_1,F_2,\widetilde{X}_1^n,\widetilde{X}_2^n) \\ &\overset{(c)}{\leq} H(W_1,W_2,F_1,F_2|Z^n) - H(W_{u_1},W_{u_2},F_{u_1},F_{u_2},V_1^n,V_2^n|\widetilde{X}_1^n,\widetilde{X}_2^n) + 2n\epsilon_n' \\ &\overset{(d)}{=} H(W_1,W_2,F_1,F_2|Z^n) - H(U_1^n,U_2^n,V_1^n,V_2^n|\widetilde{X}_1^n,\widetilde{X}_2^n) + 2n\epsilon_n' \\ &\quad + H(U_1^n|W_{u_1},W_{u_2},F_{u_1},F_{u_2},V_1^n,V_2^n,\widetilde{X}_1^n,\widetilde{X}_2^n) \\ &\quad + H(U_2^n|U_1^n,W_{u_1},W_{u_2},F_{u_1},F_{u_2},V_1^n,V_2^n,\widetilde{X}_1^n,\widetilde{X}_2^n) \\ &\overset{(e)}{\leq} H(W_1,W_2,F_1,F_2|Z^n) - H(U_1^n,U_2^n,V_1^n,V_2^n|\widetilde{X}_1^n,\widetilde{X}_2^n) + 4n\epsilon_n' \\ &\overset{(f)}{\leq} H(W_1,W_2,F_1,F_2|Z^n) - nH(U_1,U_2,V_1,V_2|\widetilde{X}_1,\widetilde{X}_2) + 4n\epsilon_n' \end{aligned} \tag{85}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

where $(a)$ follows from the Markov chain $(W_1, W_2, F_1, F_2) - (\widetilde{X}_1^n, \widetilde{X}_2^n) - (Y^n, Z^n)$, $(b)$ follows since $(V_1^n, V_2^n)$ determine $(F_{v_1}, W_{v_1}, F_{v_2}, W_{v_2})$, $(c)$ follows because $(F_{v_1}, W_{v_1}, \widetilde{X}_1^n, \widetilde{X}_2^n)$ can reliably recover $V_1^n$ by (58), and similarly because $(F_{v_2}, W_{v_2}, V_1^n, \widetilde{X}_1^n, \widetilde{X}_2^n)$ can reliably recover $V_2^n$ by (59) both due to the Markov chain $(V_1^n, V_2^n) - (\widetilde{X}_1^n, \widetilde{X}_2^n) - Y^n$, $(d)$ follows since $(U_1^n, U_2^n)$ determine $(F_{u_1}, W_{u_1}, F_{u_2}, W_{u_2})$, $(e)$ follows because $(F_{u_1}, W_{u_1}, V_1^n, V_2^n, \widetilde{X}_1^n, \widetilde{X}_2^n)$ can reliably recover $U_1^n$ by (60) and the inequality

$$H(U_1|V_1, V_2, Y) \geq H(U_1|V_1, V_2, \widetilde{X}_1, \widetilde{X}_2^n) \tag{86}$$

that can be proved similarly to (82) due to the Markov chain $U_1 - (V_1, V_2, \widetilde{X}_1, \widetilde{X}_2) - Y$. Furthermore, $(F_{u_2}, W_{u_2}, V_1^n, V_2^n, U_1^n, \widetilde{X}_1^n, \widetilde{X}_2^n)$ can reliably recover $U_2^n$ by (61) and the inequality

$$H(U_2|V_1, V_2, U_1, Y) \geq H(U_2|V_1, V_2, U_1, \widetilde{X}_1, \widetilde{X}_2) \tag{87}$$

that can be proved by using the Markov chain $U_2 - (V_1, V_2, U_1, \widetilde{X}_1, \widetilde{X}_2) - Y$, and $(f)$ follows because $(U_1^n, U_2^n, V_1^n, V_2^n, \widetilde{X}_1^n, \widetilde{X}_2^n)$ are i.i.d.

We remark that the terms in (85) are entirely similar to the terms in (80). One can show that all steps of the decodability analysis from [23, Section V-A] that is applied to ( 80) can be applied also to (85) by replacing $X$ with $(\widetilde{X}_1, \widetilde{X}_2)$, so we obtain

$$\begin{aligned} &I(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n; W_1, W_2, F_1, F_2|Z^n) \\ &\leq n[I(U_1, U_2; Z|V_1, V_2) - I(U_1, U_2; Y|V_1, V_2) + \epsilon]^- \\ &\quad + nI(U_1, U_2; \widetilde{X}_1, \widetilde{X}_2|Z) + 5n\epsilon_n'. \end{aligned} \tag{88}$$

We consider that the public indices $(F_1, F_2)$ are generated uniformly at random and the encoders generate $(V_1^n, U_1^n)$ and $(V_2^n, U_2^n)$ according to $P_{V_1^n U_1^n V_2^n U_2^n|\widetilde{X}_1^n F_1 \widetilde{X}_2^n F_2}$ obtained from the binning scheme above. This procedure induces a joint probability distribution that is almost equal to $P_{V_1 V_2 U_1 U_2 \widetilde{X}_1 \widetilde{X}_2 XYZ}$ fixed as in (19) [36, Section 1.6]. Since the privacy and secrecy leakage metrics considered above are expectations over all possible realizations $F = f$, applying the selection lemma [37, Lemma 2.2], these results prove the achievability for Theorem 1 by choosing an $\epsilon > 0$ such that $\epsilon \to 0$ when $n \to \infty$. We remark that the achievable region is convexified by using a time-sharing random variable $Q$ such that $P_{QV_1V_2} = P_Q P_{V_1|Q} P_{V_2|Q}$, required because of the $[\cdot]^-$ operation. □

### 5.2. Outer Bound

**Proof Sketch.** Assume that for some $n \geq 1$ and $\delta_n > 0$, there exist two encoders and a decoder such that (2)-(7) are satisfied for some tuple $(R_s, R_{w_1}, R_{w,2}, R_{\ell,\text{Dec}}, R_{\ell,\text{Eve}})$. Let

$$V_{1,i} \triangleq (W_1, Y_{i+1}^n, Z^{i-1}) \tag{89}$$
$$V_{2,i} \triangleq (W_2, Y_{i+1}^n, Z^{i-1}) \tag{90}$$
$$U_{1,i} \triangleq (X^{i-1}, W_1, Y_{i+1}^n, Z^{i-1}) \tag{91}$$
$$U_{2,i} \triangleq (X^{i-1}, W_2, Y_{i+1}^n, Z^{i-1}) \tag{92}$$

that satisfy the Markov chains

$$V_{1,i} - U_{1,i} - \widetilde{X}_{1,i} - X_i - (\widetilde{X}_{2,i}, Y_i, Z_i) \tag{93}$$
$$V_{2,i} - U_{2,i} - \widetilde{X}_{2,i} - X_i - (\widetilde{X}_{1,i}, Y_i, Z_i). \tag{94}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

**Admissibility of $(\mathbf{U_1}, \mathbf{U_2})$**: Define

$$n\epsilon_n = n\delta_n |\widetilde{\mathcal{X}}_1||\widetilde{\mathcal{X}}_2||\mathcal{Y}| + H_b(\delta_n) \tag{95}$$

such that $\epsilon_n \to 0$ if $\delta_n \to 0$. Using Fano's inequality and (2), we obtain

$$\begin{aligned}
n\epsilon_n &\geq H(f^n|\widehat{f^n}) \\
&\overset{(a)}{=} H(f^n|\bar{f}^n) = \sum_{i=1}^{n} H(f_i|\bar{f}_i) \\
&\geq \sum_{i=1}^{n} H(f_i|\bar{f}^n) \overset{(b)}{\geq} \sum_{i=1}^{n} H(f_i|W_1, W_2, Y^n) \\
&\geq \sum_{i=1}^{n} H(f_i|W_1, W_2, Y^n, X^{i-1}, Z^{i-1}) \\
&\overset{(c)}{=} \sum_{i=1}^{n} H(f_i|W_1, W_2, Y_{i+1}^n, X^{i-1}, Z^{i-1}, Y_i) \\
&\overset{(d)}{=} \sum_{i=1}^{n} H(f_i|U_{1,i}, U_{2,i}, Y_i)
\end{aligned} \tag{96}$$

where $(a)$ follows from [38, Lemma 2] that proves that when $n \to \infty$, there exists an i.i.d. random variable $\bar{f}^n$ that satisfies both

$$H(f^n|\widehat{f^n}) = H(f^n|\bar{f}^n) \tag{97}$$

and the Markov chain

$$\widehat{f^n} - \bar{f}^n - (W_1, W_2, Y^n) \tag{98}$$

$(b)$ follows from the data processing inequality because of the Markov chain

$$f^n - (W_1, W_2, Y^n) - \bar{f}^n \tag{99}$$

and permits randomized decoding, $(c)$ follows from the Markov chain

$$Y^{i-1} - (X^{i-1}, Z^{i-1}, W_1, W_2, Y_i, Y_{i+1}^n) - f_i \tag{100}$$

and $(d)$ follows from the definitions of $U_{1,i}$ and $U_{2,i}$.
**Public Message (Storage) Rates**: We obtain

$$\begin{aligned}
n(R_{w_1} + \delta_n) &\overset{(a)}{\geq} \log|\mathcal{W}_1| \\
&\geq H(W_1|Y^n) - H(W_1|\widetilde{X}_1^n, Y^n) \\
&= H(\widetilde{X}_1^n|Y^n) - H(\widetilde{X}_1^n|W_1, Y^n) \\
&= H(\widetilde{X}_1^n|Y^n) - \sum_{i=1}^{n} H(\widetilde{X}_{1,i}|\widetilde{X}_1^{i-1}, W_1, Y^n) \\
&\overset{(b)}{=} H(\widetilde{X}_1^n|Y^n) - \sum_{i=1}^{n} H(\widetilde{X}_{1,i}|\widetilde{X}_1^{i-1}, W_1, Y_{i+1}^n, Y_i)
\end{aligned}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$
\begin{aligned}
&\overset{(c)}{\geq} H(\widetilde{X}_1^n|Y^n) - \sum_{i=1}^{n} H(\widetilde{X}_{1,i}|X^{i-1}, Z^{i-1}, W_1, Y_{i+1}^n, Y_i) \\
&\overset{(d)}{=} nH(\widetilde{X}_1|Y) - \sum_{i=1}^{n} H(\widetilde{X}_{1,i}|U_{1,i}, Y_i) \\
&= \sum_{i=1}^{n} I(U_{1,i}; \widetilde{X}_{1,i}|Y_i) \\
&\overset{(e)}{=} \sum_{i=1}^{n} [I(V_{1,i}; \widetilde{X}_{1,i}|Y_i) + I(U_{1,i}; \widetilde{X}_{1,i}|Y_i, V_{1,i})] \\
&= \sum_{i=1}^{n} \Big[ I(V_{1,i}; \widetilde{X}_{1,i}, V_{2,i}|Y_i) - I(V_{1,i}; V_{2,i}|\widetilde{X}_{1,i}, Y_i) + I(U_{1,i}; \widetilde{X}_{1,i}, U_{2,i}|Y_i, V_{1,i}) \\
&\qquad - I(U_{1,i}; U_{2,i}|\widetilde{X}_{1,i}, Y_i, V_{1,i}) \Big] \\
&\geq \sum_{i=1}^{n} \Big[ I(V_{1,i}; \widetilde{X}_{1,i}|Y_i, V_{2,i}) - I(V_{1,i}; V_{2,i}|\widetilde{X}_{1,i}, Y_i) + I(U_{1,i}; \widetilde{X}_{1,i}|Y_i, V_{1,i}, U_{2,i}) \\
&\qquad - I(U_{1,i}; U_{2,i}|\widetilde{X}_{1,i}, Y_i, V_{1,i}) \Big]
\end{aligned} \tag{101}
$$

where $(a)$ follows by (4), $(b)$ follows from the Markov chain

$$
Y^{i-1} - (\widetilde{X}_1^{i-1}, W_1, Y_{i+1}^n, Y_i) - \widetilde{X}_{1,i} \tag{102}
$$

$(c)$ follows from the data processing inequality applied to the Markov chain

$$
(X^{i-1}, Z^{i-1}) - (\widetilde{X}_1^{i-1}, W_1, Y_{i+1}^n, Y_i) - \widetilde{X}_{1,i} \tag{103}
$$

$(d)$ follows from the definition of $U_{1,i}$, and $(e)$ follows by (93). Similarly, one can show by symmetry that we have

$$
\begin{aligned}
&n(R_{\mathrm{w}_2} + \delta_n) \\
&\geq \sum_{i=1}^{n} \Big[ I(V_{2,i}; \widetilde{X}_{2,i}|Y_i, V_{1,i}) - I(V_{2,i}; V_{1,i}|\widetilde{X}_{2,i}, Y_i) \\
&\qquad + I(U_{2,i}; \widetilde{X}_{2,i}|Y_i, V_{2,i}, U_{1,i}) - I(U_{2,i}; U_{1,i}|\widetilde{X}_{2,i}, Y_i, V_{2,i}) \Big].
\end{aligned} \tag{104}
$$

Now we consider the sum-rate bound such that

$$
\begin{aligned}
&n(R_{\mathrm{w}_1} + \delta_n) + n(R_{\mathrm{w}_2} + \delta_n) \\
&\overset{(a)}{\geq} \log(|\mathcal{W}_1| \cdot |\mathcal{W}_2|) \geq H(W_1, W_2) \\
&\geq I(W_1, W_2; \widetilde{X}_1^n, \widetilde{X}_2^n) - I(W_1, W_2; Y^n) \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}|\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, Y_{i+1}^n) - I(W_1, W_2; Y_i|\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, Y_{i+1}^n) \Big] \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2, \widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, Y_{i+1}^n; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - I(W_1, W_2, \widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, Y_{i+1}^n; Y_i) \Big]
\end{aligned}
$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$
\begin{aligned}
&\overset{(d)}{\geq} \sum_{i=1}^{n} \Big[ I(W_1, W_2, X^{i-1}, Z^{i-1}, Y_{i+1}^n; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - I(W_1, W_2, X^{i-1}, Z^{i-1}, Y_{i+1}^n; Y_i) \Big] \\
&\overset{(e)}{=} \sum_{i=1}^{n} \Big[ I(U_{1,i}, U_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - I(U_{1,i}, U_{2,i}; Y_i) \Big] \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(U_{1,i}, U_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i) \\
&\overset{(g)}{=} \sum_{i=1}^{n} \Big[ I(U_{1,i}, U_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i, V_{1,i}, V_{2,i}) + I(V_{1,i}, V_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i) \Big] \\
&\overset{(h)}{=} \sum_{i=1}^{n} \Big[ I(U_{1,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i, V_{1,i}, V_{2,i}) + I(U_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i, U_{1,i}, V_{2,i}) \\
&\qquad\qquad + I(V_{1,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i) + I(V_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i} | Y_i, V_{1,i}) \Big] \\
&\geq \sum_{i=1}^{n} \Big[ I(U_{1,i}; \widetilde{X}_{1,i} | Y_i, V_{1,i}, V_{2,i}) + I(U_{2,i}; \widetilde{X}_{2,i} | Y_i, U_{1,i}, V_{2,i}) \\
&\qquad\qquad + I(V_{1,i}; \widetilde{X}_{1,i} | Y_i) + I(V_{2,i}; \widetilde{X}_{2,i} | Y_i, V_{1,i}) \Big] \qquad (105)
\end{aligned}
$$

where $(a)$ follows by (4) and (5), $(b)$ follows from Csiszár's sum identity [39], $(c)$ follows because $(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)$ are i.i.d., $(d)$ follows from the data processing inequality applied to the Markov chains

$$(X^{i-1}, Z^{i-1}) - (\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, W_1, W_2, Y_{i+1}^n) - (\widetilde{X}_{1,i}, \widetilde{X}_{2,i}) \qquad (106)$$

$$(\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}) - (X^{i-1}, Z^{i-1}, W_1, W_2, Y_{i+1}^n) - Y_i \qquad (107)$$

$(e)$ follows from the definitions of $U_{1,i}$ and $U_{2,i}$, $(f)$ and $(g)$ follow from the Markov chain

$$(V_{1,i}, V_{2,i}) - (U_{1,i}, U_{2,i}) - (\widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - Y_i \qquad (108)$$

$(h)$ follows from the Markov chain

$$V_{1,i} - (U_{1,i}, Y_i, V_{2,i}) - (U_{2,i}, \widetilde{X}_{1,i}, \widetilde{X}_{2,i}). \qquad (109)$$

**Privacy Leakage to Decoder**: We have

$$
\begin{aligned}
&n(R_{\ell,\mathrm{Dec}} + \delta_n) \qquad (110) \\
&\overset{(a)}{\geq} H(W_1, W_2 | Y^n) - H(W_1, W_2 | X^n) \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; X_i | X^{i-1}, Y_{i+1}^n) - I(W_1, W_2; Y_i | Y_{i+1}^n, X^{i-1}) \Big] \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; X_i | X^{i-1}, Z^{i-1}, Y_{i+1}^n) - I(W_1, W_2; Y_i | Y_{i+1}^n, X^{i-1}, Z^{i-1}) \Big]
\end{aligned}
$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$\begin{aligned}
&\overset{(d)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2, X^{i-1}, Z^{i-1}, Y_{i+1}^n; X_i) - I(W_1, W_2, Y_{i+1}^n, X^{i-1}, Z^{i-1}; Y_i) \Big] \\
&\overset{(e)}{=} \sum_{i=1}^{n} \Big[ I(U_{1,i}, U_{2,i}; X_i) - I(U_{1,i}, U_{2,i}; Y_i) \Big] \\
&\overset{(f)}{=} \sum_{i=1}^{n} I(U_{1,i}, U_{2,i}; X_i | Y_i)
\end{aligned} \tag{111}$$

where $(a)$ follows by (6) and from the Markov chain $(W_1, W_2) - X^n - Y^n$, $(b)$ follows from Csiszár's sum identity, $(c)$ follows from the Markov chain

$$Z^{i-1} - (X^{i-1}, Y_{i+1}^n) - (X_i, Y_i, W_1, W_2) \tag{112}$$

$(d)$ follows because $(X^n, Y^n, Z^n)$ are i.i.d., $(e)$ follows from the definitions of $U_{1,i}$ and $U_{2,i}$, and $(f)$ follows from the Markov chain

$$(U_{1,i}, U_{2,i}) - X_i - Y_i. \tag{113}$$

**Privacy Leakage to Eve**: We have

$$\begin{aligned}
&n(R_{\ell,\text{Eve}} + \delta_n) \\
&\overset{(a)}{\geq} [H(W_1, W_2|Z^n) - H(W_1, W_2|Y^n)] + [H(W_1, W_2|Y^n) - H(W_1, W_2|X^n)] \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; Y_i|Y_{i+1}^n, Z^{i-1}) - I(W_1, W_2; Z_i|Z^{i-1}, Y_{i+1}^n) \Big] \\
&\quad + \sum_{i=1}^{n} \Big[ I(W_1, W_2; X_i|X^{i-1}, Y_{i+1}^n) - I(W_1, W_2; Y_i|Y_{i+1}^n, X^{i-1}) \Big] \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; Y_i|Y_{i+1}^n, Z^{i-1}) - I(W_1, W_2; Z_i|Z^{i-1}, Y_{i+1}^n) \Big] \\
&\quad + \sum_{i=1}^{n} \Big[ I(W_1, W_2; X_i|X^{i-1}, Y_{i+1}^n, Z^{i-1}) - I(W_1, W_2; Y_i|Y_{i+1}^n, X^{i-1}, Z^{i-1}) \Big] \\
&\overset{(d)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2, Y_{i+1}^n, Z^{i-1}; Y_i) - I(W_1, W_2, Z^{i-1}, Y_{i+1}^n; Z_i) \Big] \\
&\quad + \sum_{i=1}^{n} \Big[ I(W_1, W_2, X^{i-1}, Y_{i+1}^n, Z^{i-1}; X_i) - I(W_1, W_2, Y_{i+1}^n, X^{i-1}, Z^{i-1}; Y_i) \Big] \\
&\overset{(e)}{=} \sum_{i=1}^{n} \Big[ I(V_{1,i}, V_{2,i}; Y_i) - I(V_{1,i}, V_{2,i}; Z_i) + I(U_{1,i}, U_{2,i} V_{1,i}, V_{2,i}; X_i) \\
&\qquad - I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; Y_i) \Big] \\
&= \sum_{i=1}^{n} \Big[ - I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; Z_i) + I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; X_i) \\
&\qquad + I(U_{1,i}, U_{2,i}; Z_i|V_{1,i}, V_{2,i}) - I(U_{1,i}, U_{2,i}; Y_i|V_{1,i}, V_{2,i}) \Big]
\end{aligned}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$$\overset{(f)}{\geq} \sum_{i=1}^{n} \left[ I(U_{1,i}, U_{2,i}; X_i|Z_i) + \left[ I(U_{1,i}, U_{2,i}; Z_i|V_{1,i}, V_{2,i}) - I(U_{1,i}, U_{2,i}; Y_i|V_{1,i}, V_{2,i}) \right]^{-} \right] \tag{114}$$

where $(a)$ follows by (7) and from the Markov chain $(W_1, W_2) - X^n - Z^n$, $(b)$ follows from Csiszár's sum identity, $(c)$ follows from the Markov chain in (112), $(d)$ follows because $(X^n, Y^n, Z^n)$ are i.i.d., $(e)$ follows from the definitions of $V_{1,i}$, $V_{2,i}$, $U_{1,i}$ and $U_{2,i}$, and $(f)$ follows from the Markov chain

$$(V_{1,i}, V_{2,i}) - (U_{1,i}, U_{2,i}) - X_i - Z_i. \tag{115}$$

**Secrecy Leakage (to Eve)**: We obtain

$$\begin{aligned}
&n(R_s + \delta_n) \\
&\overset{(a)}{\geq} [H(W_1, W_2|Z^n) - H(W_1, W_2|Y^n)] + [H(W_1, W_2|Y^n) - H(W_1, W_2|\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)] \\
&\overset{(b)}{=} \sum_{i=1}^{n} \Big[ I(W_1, W_2; Y_i|Y_{i+1}^n, Z^{i-1}) - I(W_1, W_2; Z_i|Z^{i-1}, Y_{i+1}^n) \\
&\qquad + H(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}|Y_i) - H(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}|\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, W_1, W_2, Y_{i+1}^n, Y_i) \Big] \\
&\overset{(c)}{\geq} \sum_{i=1}^{n} \Big[ I(W_1, W_2, Y_{i+1}^n, Z^{i-1}; Y_i) - I(W_1, W_2, Z^{i-1}, Y_{i+1}^n; Z_i) \\
&\qquad + H(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}|Y_i) - H(\widetilde{X}_{1,i}, \widetilde{X}_{2,i}|X^{i-1}, Z^{i-1}, W_1, W_2, Y_{i+1}^n, Y_i) \Big] \\
&\overset{(d)}{=} \sum_{i=1}^{n} \Big[ I(V_{1,i}, V_{2,i}; Y_i) - I(V_{1,i}, V_{2,i}; Z_i) + I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}|Y_i) \Big] \\
&\overset{(e)}{=} \sum_{i=1}^{n} \Big[ I(V_{1,i}, V_{2,i}; Y_i) - I(V_{1,i}, V_{2,i}; Z_i) \\
&\qquad + I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; Y_i) \Big] \\
&= \sum_{i=1}^{n} \Big[ -I(U_{1,i}, U_{2,i}, V_{1,i}, V_{2,i}; Z_i) + I(U_{1,i}U_{2,i}, V_{1,i}, V_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}) \\
&\qquad + I(U_{1,i}, U_{2,i}; Z_i|V_{1,i}, V_{2,i}) - I(U_{1,i}, U_{2,i}; Y_i|V_{1,i}, V_{2,i}) \Big] \\
&\overset{(f)}{\geq} \sum_{i=1}^{n} \left[ I(U_{1,i}, U_{2,i}; \widetilde{X}_{1,i}, \widetilde{X}_{2,i}|Z_i) + \left[ I(U_{1,i}, U_{2,i}; Z_i|V_{1,i}, V_{2,i}) - I(U_{1,i}, U_{2,i}; Y_i|V_{1,i}, V_{2,i}) \right]^{-} \right]
\end{aligned} \tag{116}$$

where $(a)$ follows by (3), $(b)$ follows because $(\widetilde{X}_1^n, \widetilde{X}_2^n, Y^n)$ are i.i.d., and from Csiszár's sum identity and the Markov chain

$$Y^{i-1} - (\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, W_1, W_2, Y_{i+1}^n, Y_i) - (\widetilde{X}_{1,i}, \widetilde{X}_{2,i}) \tag{117}$$

$(c)$ follows because $(Y^n, Z^n)$ are i.i.d. and from the data processing inequality applied to the Markov chain

$$(X^{i-1}, Z^{i-1}) - (\widetilde{X}_1^{i-1}, \widetilde{X}_2^{i-1}, W_1, W_2, Y_{i+1}^n, Y_i) - (\widetilde{X}_{1,i}, \widetilde{X}_{1,i}) \tag{118}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2021 doi:10.20944/preprints202112.0135.v1

$(d)$ follows from the definitions of $V_{1,i}$, $V_{2,i}$, $U_{1,i}$, and $U_{2,i}$, $(e)$ follows from the Markov chain given in (108), and $(f)$ follows from the Markov chain

$$(V_{1,i}, V_{2,i}) - (U_{1,i}, U_{2,i}) - (\widetilde{X}_{1,i}, \widetilde{X}_{2,i}) - Z_i. \tag{119}$$

Introduce a uniformly distributed time-sharing random variable $Q \sim \text{Unif}[1:n]$ that is independent of other random variables, and define $X = X_Q$, $\widetilde{X}_1 = \widetilde{X}_{1,Q}$, $\widetilde{X}_2 = \widetilde{X}_{2,Q}$, $Y = Y_Q$, $Z = Z_Q$, $V_1 = V_{1,Q}$, $V_2 = V_{2,Q}$, $U_1 = (U_{1,Q}, Q)$, $U_2 = (U_{2,Q}, Q)$, and $f = f_Q$, so

$$(Q, V_1) - U_1 - \widetilde{X}_1 - X - (\widetilde{X}_2, Y, Z) \tag{120}$$

$$(Q, V_2) - U_2 - \widetilde{X}_2 - X - (\widetilde{X}_1, Y, Z) \tag{121}$$

form Markov chains. The proof of the outer bound follows by letting $\delta_n \to 0$.

**Cardinality Bounds**: We use the support lemma [39, Lemma 15.4] to prove the cardinality bounds and apply similar steps as in [20,23], so we omit the proof. □

## 6. 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, and two storage constraints with operational meanings on this function computation problem to define a lossless rate region by considering an eavesdropper that observes a correlated random variable. The lossless function computation problem was extended by allowing the function computed to be a distorted version of the target function, which defined the lossy function computation problem.

We proposed inner and outer bounds for the lossless and lossy rate regions. 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 characterized the exact rate region for functions that are partially invertible with respect to one of the transmitting node observations as well as for invertible functions. Moreover, we considered two different physical-degradation cases for the measurement channels of the eavesdropper and fusion center when the function computed was invertible. We derived the corresponding exact rate regions, one of which is evaluated for an example scenario, and proved that no auxiliary or time-sharing random variable is necessary to characterize these regions.

In future work, we will propose inner and outer bounds for the lossless and lossy multi-function computation problems with multiple transmitting nodes and characterize the rate regions for multi-function computations when the function computed is invertible.

**Author Contributions:** Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization,supervision, project administration, funding acquisition, O.G. The author has read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the German Federal Ministry of Education and Research (BMBF) within the national initiative for "Post Shannon Communication (NewCom)" under the Grant 16KIS1004 and by the German Research Foundation (DFG) under the Grant SCHA 1944/9-1.

**Acknowledgments:** O. Günlü thanks Matthieu Bloch and Rafael F. Schaefer for their contributions to the conference papers used in this work.

**Conflicts of Interest:** The author declares no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References


1. Günlü, O.; Bloch, M.; Schaefer, R.F. Secure Lossy Function Computation with Multiple Private Remote Source Observations. Int. ITG Workshop Smart Antennas; , 2021. to appear.
2. Günlü, O.; Bloch, M.; Schaefer, R.F. Multiple Noisy Private Remote Source Observations for Secure Function Computation. Asilomar Conf. Signals, Syst., Comput.; , 2021. to appear.
3. Mijumbi, R.; Serrat, J.; Gorricho, J.L.; Bouten, N.; De Turck, F.; Boutaba, R. Network Function Virtualization: State-of-the-Art and Research Challenges. *IEEE Commun. Surveys Tuts.* **2016**, *18*, 236–262.
4. Predd, J.B.; Kulkarni, S.B.; Poor, H.V. Distributed learning in wireless sensor networks. *IEEE Sign. Process. Mag.* **2006**, *23*, 56–69.
5. Yao, A.C. Protocols for secure computations. IEEE Symp. Foundations Comp. Sci.; , 1982; pp. 160–164.
6. Yao, A.C. How to generate and exchange secrets. IEEE Symp. Foundations Comp. Sci.; , 1986; pp. 162–167.
7. Tyagi, H.; Narayan, P.; Gupta, P. When Is a Function Securely Computable? *IEEE Trans. Inf. Theory* **2011**, *57*, 6337–6350.
8. Orlitsky, A.; Roche, J.R. Coding for computing. *IEEE Trans. Inf. Theory* **2001**, *47*, 903–917.
9. Bloch, M.; Günlü, O.; Yener, A.; Oggier, F.; Poor, H.V.; Sankar, L.; Schaefer, R.F. An Overview of Information-Theoretic Security and Privacy: Metrics, Limits and Applications. *IEEE J. Sel. Areas Inf. Theory* **2021**, *2*, 5–22.
10. Ma, N.; Ishwar, P. Some Results on Distributed Source Coding for Interactive Function Computation. *IEEE Trans. Inf. Theory* **2011**, *57*, 6180–6195.
11. Sefidgaran, M.; Tchamkerten, A. Computing a function of correlated Sources: A rate region. IEEE Int. Symp. Inf. Theory; , 2011; pp. 1856–1860.
12. Nazer, B.; Gastpar, M. Computation Over Multiple-Access Channels. *IEEE Trans. Inf. Theory* **2007**, *53*, 3498–3516.
13. Kowshik, H.; Kumar, P.R. Optimal Function Computation in Directed and Undirected Graphs. *IEEE Trans. Inf. Theory* **2012**, *58*, 3407–3418.
14. Kannan, S.; Viswanath, P. Multi-Session Function Computation and Multicasting in Undirected Graphs. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 702–713.
15. Slepian, D.; Wolf, J. Noiseless coding of correlated information sources. *IEEE Trans. Inf. Theory* **1973**, *19*, 471–480.
16. Wyner, A.D.; Ziv, J. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inf. Theory* **1976**, *22*, 1–10.
17. Goldenbaum, M.; Boche, H.; Poor, H.V. On secure computation over the binary modulo-2 adder multiple-access wiretap channel. IEEE Inf. Theory Workshop; , 2016; pp. 21–25.
18. Tyagi, H.; Watanabe, S. Converses For Secret Key Agreement and Secure Computing. *IEEE Trans. Inf. Theory* **2015**, *61*, 4809–4827.
19. Prabhakaran, V.; Ramchandran, K. On Secure Distributed Source Coding. IEEE Inf. Theory Workshop; , 2007; pp. 442–447.
20. Tu, W.; Lai, L. On Function Computation With Privacy and Secrecy Constraints. *IEEE Trans. Inf. Theory* **2019**, *65*, 6716–6733.
21. Günlü, O.; Kramer, G. Privacy, Secrecy, and Storage With Multiple Noisy Measurements of Identifiers. *IEEE Trans. Inf. Forensics Security* **2018**, *13*, 2872–2883.
22. Günlü, O. Key Agreement with Physical Unclonable Functions and Biometric Identifiers. PhD thesis, TU Munich, Germany, 2018. published by Dr.-Hut Verlag in Feb. 2019.
23. Günlü, O.; Bloch, M.; Schaefer, R.F. Secure Multi-Function Computation with Private Remote Sources. [Online]. Available: arxiv.org/abs/2106.09485.
24. Wang, Y.; Rane, S.; Draper, S.C.; Ishwar, P. A Theoretical Analysis of Authentication, Privacy, and Reusability Across Secure Biometric Systems. *IEEE Trans. Inf. Forensics Security* **2012**, *7*, 1825–1840.
25. Günlü, O.; Kittichokechai, K.; Schaefer, R.F.; Caire, G. Controllable Identifier Measurements for Private Authentication With Secret Keys. *IEEE Trans. Inf. Forensics Security* **2018**, *13*, 1945–1959.
26. Günlü, O.; Schaefer, R.F.; Kramer, G. Private Authentication with Physical Identifiers Through Broadcast Channel Measurements. IEEE Inf. Theory Workshop; , 2019; pp. 1–5.
27. Li, N.; Zhang, Y.; Kuo, C.C.J. Explainable Machine Learning based Transform Coding for High Efficiency Intra Prediction. [Online]. Available: arxiv.org/abs/2012.11152.
28. Günlü, O.; Kernetzky, T.; İşcan, O.; Sidorenko, V.; Kramer, G.; Schaefer, R.F. Secure and Reliable Key Agreement with Physical Unclonable Functions. *Entropy* **2018**, *20*.
29. Voloshynovskiy, S.; Koval, O.; Holotyak, T.; Beekhof, F. Privacy enhancement of common randomness based authentication: Key rate maximized case. IEEE Int. Workshop Inf. Forensics Security; , 2009; pp. 86–90.
30. Campisi, P. *Security and Privacy in Biometrics*; London, U.K.: Springer-Verlag, 2013.
31. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-Efficient Learning of Deep Networks from Decentralized Data. Int. Conf. Artif. Intell. Statist.; , 2017; pp. 1273–1282.
32. Yassaee, M.H.; Aref, M.R.; Gohari, A. Achievability Proof via Output Statistics of Random Binning. *IEEE Trans. Inf. Theory* **2014**, *60*, 6760–6786.
33. Gamal, A.E.; Kim, Y.H. *Network Information Theory*; Cambridge, U.K.: Cambridge University Press, 2011.

 doi:10.20944/preprints202112.0135.v1


34. Ericson, T.; Körner, J. Successive encoding of correlated sources. *IEEE Trans. Inf. Theory* **1983**, *29*, 390–395.
35. Ahmadipour, M.; Wigger, M.; Kobayashi, M. Joint Sensing and Communication over Memoryless Broadcast Channels. IEEE Inf. Theory Workshop; , 2021; pp. 1–5.
36. Bloch, M. *Lecture Notes in Information-Theoretic Security*; Atlanta, GA: Georgia Inst. Technol., 2018.
37. Bloch, M.; Barros, J. *Physical-layer Security*; Cambridge, U.K.: Cambridge University Press, 2011.
38. Günlü, O.; Schaefer, R.F.; Poor, H.V. Biometric and Physical Identifiers with Correlated Noise for Controllable Private Authentication. [Online]. Available: arxiv.org/abs/2001.00847.
39. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2 ed.; Cambridge, U.K.: Cambridge University Press, 2011.