# On the Communication Complexity of Secure Computation 

Deepesh Data ${ }^{1, \star}$, Manoj M. Prabhakaran ${ }^{2, \star \star}$, and Vinod M. Prabhakaran ${ }^{1, \star \star \star}$<br>${ }^{1}$ School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India<br>\{deepeshd, vinodmp\}@tifr.res.in<br>${ }^{2}$ Department of Computer Science, University of Illinois, Urbana-Champaign, USA<br>mmp@illinois.edu


#### Abstract

Information theoretically secure multi-party computation (MPC) is a central primitive of modern cryptography. However, relatively little is known about the communication complexity of this primitive.

In this work, we develop powerful information theoretic tools to prove lower bounds on the communication complexity of MPC. We restrict ourselves to a concrete setting involving 3 -parties, in order to bring out the power of these tools without introducing too many complications. Our techniques include the use of a data processing inequality for residual information - i.e., the gap between mutual information and Gács-Körner common information, a new information inequality for 3-party protocols, and the idea of distribution switching by which lower bounds computed under certain worst-case scenarios can be shown to apply for the general case.

Using these techniques we obtain tight bounds on communication complexity by MPC protocols for various interesting functions. In particular, we show concrete functions that have "communication-ideal" protocols, which achieve the minimum communication simultaneously on all links in the network. Also, we obtain the first explicit example of a function that incurs a higher communication cost than the input length in the secure computation model of Feige, Kilian and Naor [17, who had shown that such functions exist. We also show that our communication bounds imply tight lower bounds on the amount of randomness required by MPC protocols for many interesting functions.


## 1 Introduction

Information theoretically secure multi-party computation has been a central primitive of modern cryptography. The seminal results of Ben-Or, Goldwasser,

[^0]and Wigderson [3] and Chaum, Crépeau, and Damgård [9] showed that information theoretically secure function computation is possible between parties connected by pairwise, private links as long as only a strict minority may collude in the honest-but-curious model (and a strictly less than one-third minority may collude in the malicious model). Since then, several protocols have improved the efficiency of these protocols.

However, relatively less is known about lower bounds on the amount of communication required by a secure multi-party computation protocol, with a few notable exceptions [31|21|10|17]. In fact, [28] shows that establishing strong communication lower bounds (even with restrictions on the number of rounds) would imply breakthrough lower bound results for other well-studied problems like private-information retrieval and locally decodable codes. Further, due to the standard upper bounds on the communication needed in a secure multi-party computation protocol [39, such lower bounds would imply non-trivial circuit complexity lower bounds - a notoriously hard problem in theoretical computer science. The goal of this work is to develop tools to tackle the difficult problem of lower bounds for communication in secure multi-party computation, even if they do not immediately have direct implications to circuit complexity or locally decodable codes.

In this work we develop novel information-theoretic tools to obtain lower bounds on the communication complexity of secure computation. Our tools have connections with information-complexity techniques developed in the context of communication complexity and related problems. In particular, all these tools are related to notions of "common information" introduced by Gács-Körner [22] and Wyner [43], ${ }^{1}$

We shall restrict our study to a concrete setting that brings out the power of these tools without introducing too many additional complications. Our setting involves 3 parties (with security against corruption of any single party) of which only two parties have inputs, $X$ and $Y$, and only the third party produces an output $Z$ as a (possibly randomized) function of the inputs. This class of functions is similar to that studied in [17], but our protocol model is more general (since it allows fully interactive communication), making it harder to establish lower bounds. Also, our lower bounds apply to the semi-honest setting, where security is required only against passive corruption.

Results and Techniques. We study the setting shown in Figure 1. We obtain lower bounds on the expected number of bits that need to be exchanged between each pair of parties when securely evaluating a (possibly randomized) function of two inputs, so that Alice and Bob feed the inputs to the function, and Charlie receives the output from the function. In fact, our bounds are on the entropy
${ }^{1}$ In communication complexity and related problems, the lower bound techniques relate to Wyner common information 3916, whereas the tools in this work are more directly related to Gács-Körner common information. Wyner common information and Gács-Körner common information have been generalized to a measure of correlation represented as the "tension region" in (40).


Fig. 1. A three-party secure computation problem. Alice (party-1) has input $X$ and Bob (party-2) has $Y$. We require that (i) Charlie (party-3) obtains as output a randomized function of the other two parties' inputs, distributed as $p_{Z \mid X Y}$, (ii) Alice and Bob learn no additional information about each other's inputs, and (iii) Charlie learns nothing more about $X, Y$ than what is revealed by $Z$. All parties can talk to each other, over multiple rounds over bidirectional pairwise private links.
of the transcript between each pair 2 and hence hold even when the protocol is amortized over several instances with independent inputs. Further, often these bounds do not depend on the input distribution (as long as the distribution has full support), and hold even if the protocol is allowed to depend on the input distribution.

At a high-level, the ingredients in deriving our lower bounds are the following:

- Firstly, we observe that, since Alice and Bob do not obtain any outputs, they are both forced to reveal their inputs fully (up to equivalent inputs) to the rest of the system, and further, Charlie's output depends on the inputs only through all the communication he has with the rest of the system. Combined with the privacy requirements, this immediately leads to a naïve lower bound: specifically, writing $X, Y, Z$ as $X_{1}, X_{2}, X_{3}$, we have $H\left(M_{i j}\right) \geq H\left(X_{i}, X_{j} \mid X_{k}\right)$, where $\{i, j, k\}=\{1,2,3\}$, ${ }^{3}$
We strengthen the naïve lower bounds by relying on a "data-processing inequality" for residual information - i.e., the gap between mutual-information and (Gács-Körner) common information - which lets us relate the residual information between the messages to the residual information between the inputs/outputs. This bound is given in Theorem 1 .
- We can further improve the above lower bounds using a new tool, called distribution switching. The key idea is that the security requirement forces the
${ }^{2}$ The entropy bounds translate to bounds on the expected number of bits communicated, when we require that the messages on the individual links are encoded using (possibly adaptively chosen) prefix-free codes. See the full version [14] for details.
${ }^{3}$ We point out a simple example for which one can obtain a tight bound from this naïve bound: addition (in any group) requires one group element to be communicated between every pair of players, even with amortization over several independent instances. Previous lower bounds for secure evaluation of addition, while considering an arbitrary number of parties, either restricted themselves to bounding the number of messages required [21|10], or relied on non-standard security requirements (like "unstoppability" 21].
distribution of the transcript on certain links to be independent of the inputs. Hence, we can optimize our bounds over all input distributions having full support. Further, this shows that even if the protocol is allowed to depend on the input distribution, our bounds (which depend only on the function being evaluated) hold for every input distribution that has full support over the input domain. The resulting bound is summarised in Theorem 2
- A different improvement comes from exploiting the fact that in a protocol, the transcripts have to be generated by the parties interactively, rather than be created by an omniscient "dealer." An important technical contribution of this work is to provide a new tool towards this, in the form of a new information inequality for 3-party interactive protocols (Lemma 4). We use this along with the idea of distribution switching to significantly improve the above lower bounds by optimizing them using appropriate distributions of inputs. In fact, we can take the different terms in our bounds and optimize each of them separately using different distributions over the inputs. The resulting bounds (Theorem 3 and Theorem 4) are often stronger than what can be obtained by considering a single input distribution for the entire expression.

The resulting bounds are summarized in Theorem 1 Theorem 2 Theorem 3 and Theorem 4. We remark that unlike most of the existing results (for e.g. the bounds in [21] for summation $(\bmod 2)$ ), our lower bounds are not restricted to specific functions, but are applicable to all 3-party functions (except Theorem 2 and Theorem 4, which place some restrictions on the functions). To illustrate the use of our lower bounds, we apply them to several interesting example functions. In particular, we show the following:

- We analyze secure protocols for a few functions - GROUP-ADD, CONTROLLEDERASURE and REMOTE-OT - and, applying our lower bounds, show that these protocols achieve optimal communication complexity simultaneously on each link. We call such a protocol a communication-ideal protocol. We leave it open to characterize which functions have communication-ideal protocols.
- We show an explicit deterministic function $f:\{0,1\}^{n} \times\{0,1\}^{n} \rightarrow\{0,1\}^{n-1}$ which has a communication-ideal protocol in which Charlie's total communication cost is (and must be at least) $3 n-1$ bits. In contrast, [17] showed that there exist functions $f:\{0,1\}^{n} \times\{0,1\}^{n} \rightarrow\{0,1\}$, for which Charlie must receive at least $3 n-4$ bits, if the protocol is required to be in their non-interactive model. (Note that our bound is incomparable to that of [17], since we require the output of our function to be longer; on the other hand, our bound uses an explicit function, and continues to hold even if we allow unrestricted interaction.)

In the full version of this paper [14], we extend the above results to lower bounds on a couple of related quantities. Firstly, we identify a multi-secret sharing primitive that is interesting on its own right, but also has the property that lower bounds on its share sizes serve as lower bounds for communication complexity of MPC protocols; some of our preliminary lower bounds are, in fact, bounds on the share sizes for such a multi-secret sharing scheme. Secondly, we show
that our lower bounds for communication complexity also yield lower bounds on the amount of randomness needed in secure computation protocols. We analyze secure protocols for several natural functions, and prove that these protocols are randomness-optimal, i.e., they use the least amount of randomness.

Related Work. Communication complexity of multi-party computation without security requirements has been widely studied since [44] (see [33]), and more recently has seen the use of information-theoretic tools as well, in [7] and subsequent works. Independently, in the information theory literature communication requirements of interactive function computation have been studied (e.g. [37]).

In secure multi-party computation, there has been a vast literature on information-theoretic security, focusing on building efficient protocols, as well as characterizing various aspects like corruption models that admit secure protocols (e.g. [3|9|8|27|20|26]) and the number of rounds of interaction needed (e.g. [18|24|19|38,30]). Among other things, these results upper-bounded the communication complexity of multi-party secure computation in terms of the circuit complexity of the computation. Recently, [1] showed that, in general this upper bound is not tight by showing that all functions can be securely evaluated with sub-exponential communication (in our model of 3-party computation protocols), whereas most functions have exponential circuit complexity.

But lower-bounding communication complexity has received much less attention. For 2-party secure computation with security against passive corruption of one party (when the function admits such a protocol), communication complexity was combinatorially characterized in [31]. 21|10] gave tight lower and upper bounds on the number of messages needed for secure computation of addition $(\bmod 2)$ by $n$ parties. Further, relying on a stronger corruption model (fail-stop corruption), [21] also argued a lower bound for the amortized communication complexity of secure addition over any finite field. Feige et al. 17] obtained a lower bound on the communication complexity for a restricted class of 3-party protocols; along with positive results, they gave a modest lower bound for communication needed for evaluating random functions in this model. The difficulty of obtaining general lower bounds was pointed out by Ishai and Kushilevitz [28], who related such lower bounds to lower bounds for locally decodable codes and private information retrieval protocols. The connection to private information retrieval protocols was recently used in [1] to, among other things, derive the best known general upper bound on communication for Boolean functions in the model of [17]. The related question of how much randomness is required for secure computation seems to have received even less attention, but again, with some notable exceptions $32|5| 23 \mid 34$.

We remark that in a model with computational security, under computational hardness assumptions, the communication complexity of secure computation is linear in the input size, relying on fully homomorphic encryption ([25] and subsequent works) or exponential computation by the parties [36]. Also, in a model with exponential amount of correlated randomness shared among the parties, such a result was obtained in [29].

Information-theoretic tools have been successfully used in deriving bounds in various cryptographic problems like key agreement (e.g. 35[11), secure 2-party computation (e.g. [15]) and secret-sharing and its variants (e.g. [2] and [4]). In this work, we rely on information-theoretic tools developed in 4240, which also considered cryptographic problems. Some preliminary observations leading to this work appeared in [13].

## 2 Preliminaries

Notation. We write $p_{X}$ to denote the distribution of a discrete random variable $X ; p_{X}(x)$ denotes $\operatorname{Pr}[X=x]$. When clear from the context, the subscript of $p_{X}$ will be omitted. The conditional distribution denoted by $p_{Z \mid U}$ specifies $\operatorname{Pr}[Z=$ $z \mid U=u$ ], for each value $z$ that $Z$ can take and each value $u$ that $U$ can take. A randomized function of two variables, is specified by a probability distribution $p_{Z \mid X Y}$, where $X, Y$ denote the two input variables, and $Z$ denotes the output variable.

For random variables $T, U, V$, we write the Markov chain $T-U-V$ to indicate that $T$ and $V$ are conditionally independent conditioned on $U$ : i.e., $I(T ; V \mid U)=$ 0 . All logarithms are to the base 2 and entropies are in bits.
Protocols. A 3-party protocol $\Pi$ is specified by a collection of "next message functions" ( $\Pi_{1}, \Pi_{2}, \Pi_{3}$ ) which probabilistically map a state of the protocol to the next state (in a restricted manner), and output functions ( $\Pi_{1}^{\text {out }}, \Pi_{2}^{\text {out }}, \Pi_{3}^{\text {out }}$ ) used to define the outputs of the parties as probabilistic functions of their views. We shall also allow the protocol to depend on the distribution of the inputs to the parties. (This would allow one to tune a protocol to be efficient for a suitable input distribution. Allowing this makes our lower bounds stronger; on the other hand, none of the protocols we give for our examples require this flexibility.)

Without loss of generality, the state of the protocol consists only of the inputs received by each party and the transcript of the messages exchanged so far $4^{4}$ We denote the final transcripts on the three links, after executing protocol $\Pi$ on its specified input distribution by $M_{12}^{\Pi}, M_{23}^{\Pi}$ and $M_{31}^{\Pi}$. When $\Pi$ is clear from the context, we simply write $M_{12}$ etc. We define $M_{1}=\left(M_{12}, M_{31}\right)$ as the transcripts that party 1 can see; $M_{2}$ and $M_{3}$ are defined similarly. We define the view of the $i^{\text {th }}$ party, $V_{i}$ to consist of $M_{i}$ and that party's inputs and outputs (if any).

It is easy to see that a protocol, along with an input distribution, fully defines a joint distribution over all the inputs, outputs and the joint transcripts on all the links.
Secure Computation. We consider three party computation functionalities, in which Alice and Bob (parties 1 and 2) receive as inputs the random variables $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$, respectively, and Charlie (party 3) produces an output $Z \in \mathcal{Z}$ distributed according to a specified distribution $p_{Z \mid X Y}$. The sets $\mathcal{X}, \mathcal{Y}$

[^1]and $\mathcal{Z}$ are always finite. In secure computation, we shall consider the inputs to the computation to come from a distribution $p_{X Y}$ over $\mathcal{X} \times \mathcal{Y}$.

A (perfectly) secure computation protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)=\left(\Pi_{1}, \Pi_{2}, \Pi_{3}, \Pi_{3}^{\text {out }}\right)$ for $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is a protocol which satisfies the following conditions:

- Correctness: Output of Charlie, is distributed according to $p_{Z \mid X=x, Y=y}$, where $x, y$ are the inputs to Alice and Bob
- Privacy: The privacy condition corresponds to "1-privacy", wherein at most one party is passively corrupt. Corresponding to security against Alice, Bob and Charlie, respectively, we have the following three Markov chains. $V_{1}-X-$ $(Y, Z), V_{2}-Y-(X, Z)$ and $V_{3}-Z-(X, Y)$. Equivalently (see Footnote 4), $I\left(M_{1} ;(Y, Z) \mid X\right)=I\left(M_{2} ;(X, Z) \mid Y\right)=I\left(M_{3} ;(X, Y) \mid Z\right)=0$.
Intuitively, the privacy condition guarantees that even if one party (say Alice) is curious, and retains its view from the protocol (in particular, $M_{1}$ ), this view reveals nothing more to it about the inputs and outputs of the other parties (namely, $Y, Z$ ), than what its own inputs and outputs reveal (as long as the other parties erase their own views). In other words, a curious party may as well simulate a view for itself based on just its inputs and outputs, rather than retain the actual view it obtained from the protocol execution.

For simplicity, we prove all our results for perfect security as defined above; this is also the setting for classical positive results like that of 33. But in fact, we expect all our bounds to extend to the setting of statistical security as well (following 41|40] who extend similar results to the statistical security case) 5 Also, the above security requirements are for an honest execution of the protocol (corresponding to honest-but-curious or passive corruption of at most one party). The lower bounds derived in this model typically continue to hold for active corruption as well (since for many functionalities, every protocol secure against active corruption is a protocol secure against passive corruption); in this case, when a party uses a broadcast channel (as would be necessary in our setting, where 1 out of 3 parties is corrupted), it is counted as sending individual messages to every other party.
Communication Complexity and Entropy. A standard approach to lowerbounding the number of bits in a string is to lower-bound its entropy. However, in an interactive setting, a party sees the messages in each round, rather than just a concatenation of all the bits sent over the entire protocol. In a setting where we allow variable length messages, this would seem to allow communicating more bits of information than the length of the transcript itself. But this allows the parties to learn when the message transmitted in a round ends, implicitly inserting an end-of-message marker into the bit stream. To account for this, one can require that the message sent in every round is a codeword in a prefixfree code. (The code itself can be dynamically determined based on previous

[^2]messages exchanged over the link.) It can be shown that, with this requirement, the number of bits communicated in each link is indeed lower-bounded by the entropy of the transcript in that link.
Normal Form. For a pair $\left(p_{X Y}, p_{Z \mid X Y}\right)$, define the relations $x \cong x^{\prime}, y \cong y^{\prime}$ and $z \cong z^{\prime}$ as follows.

1. For any $x, x^{\prime} \in \mathcal{X}$, let $\mathcal{S}_{x, x^{\prime}}=\left\{y \in \mathcal{Y}: p_{X Y}(x, y)>0, p_{X Y}\left(x^{\prime}, y\right)>0\right\}$. We say that $x \cong x^{\prime}$, if $\forall y \in \mathcal{S}_{x, x^{\prime}}$ and $z \in \mathcal{Z}$, we have $p_{Z \mid X Y}(z \mid x, y)=p_{Z \mid X Y}\left(z \mid x^{\prime}, y\right)$.
2. For any $y, y^{\prime} \in \mathcal{Y}$, let $\mathcal{S}_{y, y^{\prime}}=\left\{x \in \mathcal{X}: p_{X Y}(x, y)>0, p_{X Y}\left(x, y^{\prime}\right)>0\right\}$. We say that $y \cong y^{\prime}$, if $\forall x \in \mathcal{S}_{y, y^{\prime}}$ and $z \in \mathcal{Z}$, we have $p_{Z \mid X Y}(z \mid x, y)=p_{Z \mid X Y}\left(z \mid x, y^{\prime}\right)$.
3. Let $\mathcal{S}=\left\{(x, y): p_{X Y}(x, y)>0\right\}$. For any $z, z^{\prime} \in \mathcal{Z}$, we say that $z \cong z^{\prime}$, if $\exists c \geq 0$ such that $\forall(x, y) \in \mathcal{S}$, we have $p_{Z \mid X Y}(z \mid x, y)=c \cdot p_{Z \mid X Y}\left(z^{\prime} \mid x, y\right)$.

A pair $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is said to be in normal form if $x \cong x^{\prime} \Rightarrow x=x^{\prime}, y \cong y^{\prime} \Rightarrow$ $y=y^{\prime}$, and $z \cong z^{\prime} \Rightarrow z=z^{\prime}$.

It is easy to show (as we do in the full version) that we may assume without loss of generality that $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is in normal form since any given $\left(p_{X Y}, p_{Z \mid X Y}\right)$ can be transformed to a ( $p_{X^{\prime} Y^{\prime}}, p_{Z^{\prime} \mid X^{\prime} Y^{\prime}}$ ) in normal form so that any secure computation protocol for the former can be transformed to one for the latter with the same communication costs, and vice versa.

Communication-Ideal Protocol. We say that a protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$ for securely computing a randomized function $p_{Z \mid X Y}$, for a distribution $p_{X Y}$ is communication-ideal if for each $i j \in\{12,23,31\}$,

$$
H\left(M_{i j}^{\Pi}\right)=\inf _{\Pi^{\prime}\left(p_{X Y}, p_{Z \mid X Y}\right)} H\left(M_{i j}^{\Pi^{\prime}}\right)
$$

where the infimum is over all secure protocols for $p_{Z \mid X Y}$ with the same distribution $p_{X Y}$. That is, a communication-ideal protocol achieves the least entropy possible for every link, simultaneously. We remark that it is not clear, a priori, how to determine if a given function $p_{Z \mid X Y}$ has a communication-ideal protocol for a given distribution $p_{X Y}$.

## Common Information and Residual Information

Gács and Körner [22] introduced the notion of common information to measure a certain aspect of correlation between two random variables. The Gács-Körner common information of a pair of correlated random variables $(U, V)$ can be defined as $H(U \sqcap V)$, where $U \sqcap V$ is a random variable with maximum entropy among all random variables $Q$ that are determined both by $U$ and by $V$ (i.e., there are functions $f$ and $g$ such that $Q=f(U)=g(V))$. In [40, the gap between mutual information and common information was termed residual information: $R I(U ; V):=I(U ; V)-H(U \sqcap V)$.

In [42], Wolf and Wullschleger identified (among other things) the following important data processing inequality for residual information.

Lemma 1 ([42]). If $T, U, V, W$ are jointly distributed random variables such that the following two Markov chains hold: (i) $U-T-W$, and (ii) $T-W-V$, then

$$
R I(T ; W) \leq R I((U, T) ;(V, W))
$$

The Markov chain conditions above correspond to the requirement that it is secure (against honest-but-curious adversaries) to require a pair of parties holding the views $(U, T)$ and $(V, W)$, to produce outputs $T, W$, respectively, because for the first party, the rest of its view, $U$, can be simulated based on the output $T$, independent of the output $W$ (and similarly, for the second party). The lemma states that under such a secure transformation from views to outputs, the residual information can only decrease.

It is easy to see that the following is an equivalent definition of residual information (see [40]).

$$
\begin{equation*}
R I(U ; V)=\min _{\substack{Q: \exists f, g \text { s.t. } \\ Q=f(U)=g(V)}} I(U ; V \mid Q) \text {. } \tag{1}
\end{equation*}
$$

The random variable $Q$ which achieves the minimum is, in fact, $U \sqcap V$. Note that the residual information is always non-negative.

## 3 Lower Bounds on Communication Complexity

This section is divided into three parts. In Section 3.1, we derive preliminary lower bounds for secure computation. In each of the subsequent subsections, we give different improvements of the lower bounds derived in Section 3.1. Omitted proofs are available in the full version [14].

### 3.1 Preliminary Lower Bounds

We first state the following basic lemma for any protocol for secure computation. Similar results have appeared in the literature earlier (for instance, special cases of Lemma 2 appear in (16/41|13).
Lemma 2. Suppose $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is in normal form. Then, in any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, the cut isolating Alice from Bob and Charlie must reveal Alice's input $X$, i.e., $H\left(X \mid M_{12}, M_{31}\right)=0$. Similarly, $H\left(Y \mid M_{12}, M_{23}\right)=0$ and $H\left(Z \mid M_{23}, M_{31}\right)=0$.

Lemma 2 states the simple fact that, for $\left(p_{X Y}, p_{Z \mid X Y}\right)$ in normal form, the information about a party's input must flow out through the links she/he is part of, and the information about Charlie's output must flow in through the links he is part of. This crucially relies on the fact that Alice and Bob obtain no output, and Charlie has no input in our model.

We obtain a preliminary lower bound in Theorem 1 below by using the above lemma and the data-processing inequality for residual information in Lemma 1 Recall that the assumption below of $\left(p_{X Y}, p_{Z \mid X Y}\right)$ being in normal form is without loss of generality.

Theorem 1. Any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, where $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is in normal form, should satisfy the following lower bounds on the entropy of the transcripts on each link.

$$
\begin{align*}
& H\left(M_{23}\right) \geq \max \{R I(X ; Z), R I(X ; Y)\}+H(Y, Z \mid X),  \tag{2}\\
& H\left(M_{31}\right) \geq \max \{R I(Y ; Z), R I(X ; Y)\}+H(X, Z \mid Y),  \tag{3}\\
& H\left(M_{12}\right) \geq \max \{R I(X ; Z), R I(Y ; Z)\}+H(X, Y \mid Z) \tag{4}
\end{align*}
$$

Proof: We shall prove (2). The other two inequalities follow similarly.

$$
\begin{align*}
H\left(M_{23}\right) & \geq \max \left\{H\left(M_{23} \mid M_{31}\right), H\left(M_{23} \mid M_{12}\right)\right\} \\
& =\max \left\{I\left(M_{23} ; M_{12} \mid M_{31}\right), I\left(M_{23} ; M_{31} \mid M_{12}\right)\right\}+H\left(M_{23} \mid M_{12}, M_{31}\right) \tag{5}
\end{align*}
$$

We can bound the last term of (5) as follows (to already get a naïve bound):

$$
\begin{aligned}
H\left(M_{23} \mid M_{12}, M_{31}\right) & \stackrel{(\mathrm{a})}{=} H\left(M_{23}, Y, Z \mid M_{12}, M_{31}, X\right) \\
& \geq H\left(Y, Z \mid M_{12}, M_{31}, X\right) \stackrel{(\mathrm{b})}{=} H(Y, Z \mid X)
\end{aligned}
$$

where (a) follows from Lemma 2 and (b) follows from the privacy against Alice. Next, we lower bound the first term inside max of (5) by $R I(X ; Z)$ as follows.

$$
\begin{equation*}
I\left(M_{23} ; M_{12} \mid M_{31}\right)=I\left(M_{23} M_{31} ; M_{12} M_{31} \mid M_{31}\right) \geq R I\left(M_{23}, M_{31} ; M_{12}, M_{31}\right) \tag{6}
\end{equation*}
$$

where the inequality follows from (1) by taking $Q=M_{31}$. Now, by privacy against Charlie, we have $\left(M_{23}, M_{31}\right)-Z-X$ and by privacy against Alice, we have $\left(M_{12}, M_{31}\right)-X-Z$. Applying Lemma 1 with the above Markov chains, together with Lemma 2, we get

$$
R I\left(M_{23}, M_{31} ; M_{12}, M_{31}\right) \geq R I(Z ; X)=R I(X ; Z)
$$

Similarly, we can lower-bound the second term inside max of (5) by $R I(X ; Y)$, completing the proof.
In the rest of the paper we will restrict our attention to $p_{X Y}$ which have full support. This will allow us to strengthen the preliminary bounds in Theorem 1

### 3.2 Improved Lower Bounds via Distribution Switching

To improve the bounds in Theorem 1, we will use a technique we call distribution switching. This significantly improves the above bounds and leads to one of our main theorems.

The following lemma states that privacy requirements imply that the transcript $M_{12}$ generated by a secure protocol computing $p_{Z \mid X Y}$ is independent of both the inputs. Moreover, if the function $p_{Z \mid X Y}$ satisfies some additional constraints, then the other two transcripts also become independent of the inputs. The characteristic bipartite graph of a distribution $p_{X Y}$ is defined as a bipartite graph on vertex set $\mathcal{X} \cup \mathcal{Y}$ such that $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ have an edge between them whenever $p_{X Y}(x, y)>0$. The proof of the following lemma is along the lines of a similar lemma in [13].

Lemma 3. Consider a function $p_{Z \mid X Y}$.

1. Suppose that $p_{X Y}$ is such that the characteristic bipartite graph of $p_{X Y}$ is connected. Then, for any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, we have $I\left(X, Y, Z ; M_{12}\right)=0$.
2. Suppose $p_{X Y}$ has full support and $p_{Z \mid X Y}$ satisfies the following condition:

Condition 1. There is no non-trivial partition $\mathcal{X}=\mathcal{X}_{1} \cup \mathcal{X}$ (i.e., $\mathcal{X}_{1} \cap \mathcal{X}_{2}=$ $\varnothing$ and neither $\mathcal{X}_{1}$ nor $\mathcal{X}_{2}$ is empty), such that if $\mathcal{Z}_{k}=\left\{z \in \mathcal{Z}: x \in \mathcal{X}_{k}, y \in\right.$ $\mathcal{Y}, p(z \mid x, y)>0\}, k=1,2$, their intersection $\mathcal{Z}_{1} \cap \mathcal{Z}_{2}$ is empty.
Then, for any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, we have $I\left(X, Y, Z ; M_{31}\right)=0$.
3. Suppose $p_{X Y}$ has full support and $p_{Z \mid X Y}$ satisfies the following condition:

Condition 2. There is no non-trivial partition $\mathcal{Y}=\mathcal{Y}_{1} \cup \mathcal{Y}_{2}$ such that if $\mathcal{Z}_{k}=\left\{z \in \mathcal{Z}: x \in \mathcal{X}, y \in \mathcal{Y}_{k}, p(z \mid x, y)>0\right\}, k=1,2$, their intersection $\mathcal{Z}_{1} \cap \mathcal{Z}_{2}$ is empty.
Then, for any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, we have $I\left(X, Y, Z ; M_{23}\right)=0$.
We note that $p_{X Y}$ will have a connected characteristic bipartite graph if it has full support.

We will now strengthen the lower bounds from Theorem 1. Specifically, we will argue that even if the protocol is allowed to depend on the input distribution (as we do here), correctness and privacy conditions will require that the lower bounds derived for when the distributions of the inputs are changed continue to hold for the original setting.

Theorem 2. Consider any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, where $p_{X Y}$ has full support and ( $p_{X Y}, p_{Z \mid X Y}$ ) is in normal form.

1. We have the following strengthening of (4):

$$
\begin{align*}
& H\left(M_{12}\right) \geq \max \left\{\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(X^{\prime} ; Z^{\prime}\right)+H\left(X^{\prime}, Y^{\prime} \mid Z^{\prime}\right)\right)\right. \\
&\left.\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(Y^{\prime} ; Z^{\prime}\right)+H\left(X^{\prime}, Y^{\prime} \mid Z^{\prime}\right)\right)\right\} \tag{7}
\end{align*}
$$

where the sup operations are over $p_{X^{\prime} Y^{\prime}}$ having full support and the objective functions are evaluated using $p_{X^{\prime} Y^{\prime} Z^{\prime}}(x, y, z)=p_{X^{\prime} Y^{\prime}}(x, y) p_{Z \mid X Y}(z \mid x, y)$.
2. If $p_{Z \mid X Y}$ satisfies Condition 1 of Lemma 3, we have the following strengthening of (3):

$$
\begin{align*}
& H\left(M_{31}\right) \geq \max \left\{\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(Y^{\prime} ; Z^{\prime}\right)+H\left(X^{\prime}, Z^{\prime} \mid Y^{\prime}\right)\right)\right. \\
&\left.\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(X^{\prime} ; Y^{\prime}\right)+H\left(X^{\prime}, Z^{\prime} \mid Y^{\prime}\right)\right)\right\} \tag{8}
\end{align*}
$$

where the sup operations are over the same set of $p_{X^{\prime} Y^{\prime}}$ as in (7).
3. If $p_{Z \mid X Y}$ satisfies Condition 2 of Lemma 3, we have the following strengthening of (2):

$$
\begin{align*}
& H\left(M_{23}\right) \geq \max \left\{\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(X^{\prime} ; Z^{\prime}\right)+H\left(Y^{\prime}, Z^{\prime} \mid X^{\prime}\right)\right),\right. \\
&\left.\sup _{p_{X^{\prime} Y^{\prime}}}\left(R I\left(X^{\prime} ; Y^{\prime}\right)+H\left(Y^{\prime}, Z^{\prime} \mid X^{\prime}\right)\right)\right\}, \tag{9}
\end{align*}
$$

where the sup operations are over the same set of $p_{X^{\prime} Y^{\prime}}$ as in (7).
Proof: Notice that any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, where distribution $p_{X Y}$ has full support, continues to be a secure protocol even if we switch the input distribution to a different one $p_{X^{\prime} Y^{\prime}}$. This follows directly from examining the correctness and privacy conditions required for a protocol to be secure.

By Lemma 3 it follows that the transcript $M_{12}$ of the protocol (under both the original and the switched input distributions) must remain independent of the input data $X, Y$. Furthermore, since $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is in normal form and $p_{X^{\prime} Y^{\prime}}$ has full support, $\left(p_{X^{\prime} Y^{\prime}}, p_{Z \mid X Y}\right)$ is also in normal form. Hence, (7) follows from (4) of Theorem 1. Similarly, if the function $p_{Z \mid X Y}$ satisfies the condition 1 (resp. 2) of Lemma 3, we can show (8) (resp. (9)) as well.

### 3.3 An Information Inequality for Protocols and Improved Lower Bounds

We can give a different improvement to Theorem 1 by exploiting the fact that, in a protocol, transcripts are generated by the parties interactively rather than by an omniscient dealer. Towards this, we derive an information inequality relating the transcripts on different links in general 3-party protocols, in which parties do not share any common or correlated randomness or correlated inputs at the beginning of the protocol.

Lemma 4. In a 3-party protocol, if the inputs to the parties are independent of each other, then, for $\{\alpha, \beta, \gamma\}=\{1,2,3\}$,

$$
I\left(M_{\gamma \alpha} ; M_{\beta \gamma}\right) \geq I\left(M_{\gamma \alpha} ; M_{\beta \gamma} \mid M_{\alpha \beta}\right)
$$

Further, as in (6), $I\left(M_{\gamma \alpha} ; M_{\beta \gamma} \mid M_{\alpha \beta}\right) \geq R I\left(M_{\gamma \alpha}, M_{\alpha \beta} ; M_{\beta \gamma}, M_{\alpha \beta}\right)$. Hence, if the inputs are independent of each other,

$$
\begin{equation*}
I\left(M_{\gamma \alpha} ; M_{\beta \gamma}\right) \geq I\left(M_{\gamma \alpha} ; M_{\beta \gamma} \mid M_{\alpha \beta}\right) \geq R I\left(M_{\gamma \alpha}, M_{\alpha \beta} ; M_{\beta \gamma}, M_{\alpha \beta}\right) \tag{10}
\end{equation*}
$$

This inequality provides us with a means to exploit the protocol structure behind the transcripts. For instance, consider a secure protocol $\Pi\left(p_{X} p_{Y}, p_{Z \mid X Y}\right)$, where $p_{X}, p_{Y}$ have full support and ( $p_{X} p_{Y}, p_{Z \mid X Y}$ ) is in normal form. We have,

$$
\begin{aligned}
H\left(M_{12}\right) & =I\left(M_{12} ; M_{23}\right)+H\left(M_{12} \mid M_{23}\right) \\
& =I\left(M_{12} ; M_{23}\right)+I\left(M_{12} ; M_{31} \mid M_{23}\right)+H\left(M_{12} \mid M_{23}, M_{31}\right) \\
& \geq R I(X ; Z)+R I(Y ; Z)+H(X, Y \mid Z)
\end{aligned}
$$

where the last inequality used $H\left(M_{12} \mid M_{23}, M_{31}\right) \geq H(X, Y \mid Z)$ and $I\left(M_{12} ; M_{31} \mid M_{23}\right) \geq R I(Y ; Z)$ (both as in the proof of Theorem 1) as well as $I\left(M_{12} ; M_{23}\right) \geq R I(X ; Z)$. Thus the term $\max \{R I(X ; Z), R I(Y ; Z)\}$ in (4) can be replaced by $R I(X ; Z)+R I(Y ; Z)$ for independent inputs.

In the full version we prove the following two theorems by making use of Lemma 4 and distribution switching:

Theorem 3. The following communication complexity bounds hold for any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, where $p_{X Y}$ has full support and $\left(p_{X Y}, p_{Z \mid X Y}\right)$ is in normal form:

$$
\begin{align*}
H\left(M_{23}\right) \geq & \left(\sup _{p_{X^{\prime}}} R I\left(X^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{X^{\prime \prime}}} H\left(Y, Z^{\prime \prime} \mid X^{\prime \prime}\right)\right)  \tag{11}\\
H\left(M_{31}\right) \geq & \left(\sup _{p_{Y^{\prime}}} R I\left(Y^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{Y^{\prime \prime}}} H\left(X, Z^{\prime \prime} \mid Y^{\prime \prime}\right)\right),  \tag{12}\\
H\left(M_{12}\right) \geq & \max \left\{\sup _{p_{X^{\prime}}}\left(\sup _{p_{Y^{\prime}}} R I\left(Y^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{Y^{\prime}}} R I\left(X^{\prime} ; Z^{\prime \prime}\right)+H\left(X^{\prime}, Y^{\prime \prime} \mid Z^{\prime \prime}\right)\right),\right. \\
& \left.\sup _{p_{Y^{\prime}}}\left(\sup _{p_{X^{\prime}}} R I\left(X^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{X^{\prime}}} R I\left(Y^{\prime} ; Z^{\prime \prime}\right)+H\left(X^{\prime \prime}, Y^{\prime} \mid Z^{\prime \prime}\right)\right)\right\}, \tag{13}
\end{align*}
$$

where the sup operations are over distributions $p_{X^{\prime}}, p_{X^{\prime \prime}}, p_{Y^{\prime}}, p_{Y^{\prime \prime}}$ having full support. The terms in the right hand side of (11) are evaluated using the distribution $p_{Y}$ of the data $Y$ of Bob, i.e.,

$$
\begin{aligned}
p_{X^{\prime} Y Z^{\prime}}(x, y, z) & =p_{X^{\prime}}(x) p_{Y}(y) p_{Z \mid X Y}(z \mid x, y), \\
p_{X^{\prime \prime} Y Z^{\prime \prime}}(x, y, z) & =p_{X^{\prime \prime}}(x) p_{Y}(y) p_{Z \mid X Y}(z \mid x, y)
\end{aligned}
$$

Similarly, the terms in (12) are evaluated using the distribution $p_{X}$ of the data $X$ of Alice. The lower bound in (13) does not depend on the distributions $p_{X}$ and $p_{Y}$ of the data. The terms on the top row of (13), for instance, are evaluated using

$$
\begin{aligned}
p_{X^{\prime} Y^{\prime} Z^{\prime}}(x, y, z) & =p_{X^{\prime}}(x) p_{Y^{\prime}}(y) p_{Z \mid X Y}(z \mid x, y) \\
p_{X^{\prime} Y^{\prime \prime} Z^{\prime \prime}}(x, y, z) & =p_{X^{\prime}}(x) p_{Y^{\prime \prime}}(y) p_{Z \mid X Y}(z \mid x, y)
\end{aligned}
$$

When the function satisfies certain additional constraints, we can strengthen the lower bounds on the $H\left(M_{23}\right)$ and $H\left(M_{31}\right)$ as follows:

Theorem 4. Consider any secure protocol $\Pi\left(p_{X Y}, p_{Z \mid X Y}\right)$, where $p_{X Y}$ has full support and ( $p_{X Y}, p_{Z \mid X Y}$ ) is in normal form.

1. Suppose the function $p_{Z \mid X Y}$ satisfies Condition 1 of Lemma 3. Then, we have the following strengthening of (12).

$$
\begin{equation*}
H\left(M_{31}\right) \geq \sup _{p_{X^{\prime}}}\left(\left(\sup _{p_{Y^{\prime}}} R I\left(Y^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{Y^{\prime \prime}}} H\left(X^{\prime}, Z^{\prime \prime} \mid Y^{\prime \prime}\right)\right)\right) \tag{14}
\end{equation*}
$$

where the sup operations are over distributions $p_{X^{\prime}}, p_{Y^{\prime}}, p_{Y^{\prime \prime}}$ having full support and the terms in the right hand side are evaluated using the distribution

$$
\begin{aligned}
& p_{X^{\prime} Y^{\prime} Z^{\prime} Y^{\prime \prime} Z^{\prime \prime}}\left(x^{\prime}, y^{\prime}, z^{\prime}, y^{\prime \prime}, z^{\prime \prime}\right)= \\
& \quad p_{X^{\prime}}\left(x^{\prime}\right) p_{Y^{\prime}}\left(y^{\prime}\right) p_{Z \mid X Y}\left(z^{\prime} \mid x^{\prime}, y^{\prime}\right) p_{Y^{\prime \prime}}\left(y^{\prime \prime}\right) p_{Z \mid X Y}\left(z^{\prime \prime} \mid x^{\prime}, y^{\prime \prime}\right)
\end{aligned}
$$

2. Suppose the function $p_{Z \mid X Y}$ satisfies Condition 2 of Lemma 3. Then, we have the following strengthening of (11).

$$
\begin{equation*}
H\left(M_{23}\right) \geq \sup _{p_{Y^{\prime}}}\left(\left(\sup _{p_{X^{\prime}}} R I\left(X^{\prime} ; Z^{\prime}\right)\right)+\left(\sup _{p_{X^{\prime \prime}}} H\left(Y^{\prime}, Z^{\prime \prime} \mid X^{\prime \prime}\right)\right)\right) \tag{15}
\end{equation*}
$$

where the sup operations are over distributions $p_{X^{\prime}}, p_{X^{\prime \prime}}, p_{Y^{\prime}}$ having full support and the terms in the right hand side are evaluated using the distribution

$$
\begin{aligned}
& p_{X^{\prime} Y^{\prime} Z^{\prime} X^{\prime \prime} Z^{\prime \prime}}\left(x^{\prime}, y^{\prime}, z^{\prime}, x^{\prime \prime}, z^{\prime \prime}\right)= \\
& \quad p_{X^{\prime}}\left(x^{\prime}\right) p_{Y^{\prime}}\left(y^{\prime}\right) p_{Z \mid X Y}\left(z^{\prime} \mid x^{\prime}, y^{\prime}\right) p_{X^{\prime \prime}}\left(x^{\prime \prime}\right) p_{Z \mid X Y}\left(z^{\prime \prime} \mid x^{\prime \prime}, y^{\prime}\right)
\end{aligned}
$$

Note that in Theorem 2, Theorem 3 and Theorem 4 any choice of $p_{X^{\prime} Y^{\prime}}$, $p_{X^{\prime}}, p_{X^{\prime \prime}}, p_{Y^{\prime}}, p_{Y^{\prime \prime}}$ (with full support) will yield a lower bound. For a given function, while all choices do yield valid lower bounds, one is often able to obtain the best lower bound analytically (as in Theorem 5 where it is seen to be the best as it matches an upper bound) or numerically (as in Theorem 6).

To summarize, for any secure computation problem ( $p_{X Y}, p_{Z \mid X Y}$ ), expressed in the normal form, Theorem 1 gives lower bounds on entropies of transcripts on all three links. If, in addition, $p_{X Y}$ has full support, then for $H\left(M_{31}\right)$, our best lower bound is the larger of (3) and (12); for $H\left(M_{23}\right)$, it is the larger of (2) and (11); and for $H\left(M_{12}\right)$, it is the larger of (7) and (13). Further, if $p_{Z \mid X Y}$ satisfies condition 1 of Lemma 3 then for $H\left(M_{31}\right)$, our best lower bound is the larger of (8) and (14); if $p_{Z \mid X Y}$ satisfies condition 2 of Lemma 3 then for $H\left(M_{23}\right)$, our best lower bound is the larger of (9) and (15).

Our communication lower bounds were developed for protocols whose designs may take into account the joint distribution of $X$ and $Y$. However, the right hand sides of (7) and (13) do not depend on the distribution $p_{X Y}$ of the inputs. Thus, even though we allow the protocol to depend on the distributions, our lower bound on $H\left(M_{12}\right)$ does not. The same is true for (8) and (14) for $H\left(M_{31}\right)$ (resp. (19) and (15) for $H\left(M_{23}\right)$ ), which apply when the function $p_{Z \mid X Y}$ satisfies condition 1 (resp. 2) of Lemma 3. When these conditions are not satisfied, the communication complexity of the optimal protocol may indeed depend on the distribution of the input (see full version for an example).

## 4 Application to Specific Functions

In this section we consider a few important examples, and apply our generic lower bounds from above to these examples, to obtain interesting results. While
many of these results are natural to conjecture, they are not easy to prove (see, for instance, Footnote 3).
Optimality of the FKN Protocol. Feige et al. [17] provided a generic (noninteractive) secure computation protocol for all 3-party functions in our model. This protocol uses a straight-forward (but "inefficient") reduction from an arbitrary function to a variant of the oblivious transfer problem, which we shall call the remote OT function (defined below), and then gives a simple protocol for this new function. While the resulting protocol is inefficient for most functions, one could ask whether the protocol that [17] used for REMOTE OT itself is optimal. We use our lower bounds from above to show that this is indeed the case.

The Remote $\binom{m}{1}-\mathrm{OT}_{2}^{n}$ function, is defined as follows: Alice's input $X=$ $\left(X_{0}, X_{1}, \ldots, X_{m-1}\right)$ is made up of $m$ strings each of length $n$ bits, and Bob has an input $Y \in\{0,1, \ldots, m-1\}$. Charlie wants to compute $Z=f(X, Y)=X_{Y}$. The protocol of [17] requires $n m$ bits to be exchanged over the Alice-Charlie (31) link, $n+\log m$ bits over the Bob-Charlie (23) link and $n m+\log m$ bits over the Alice-Bob (12) link. In the full version, we prove the following theorem, which shows that this protocol is optimal and in fact, a communication-ideal protocol.

Theorem 5. Any secure protocol $\Pi$ ( $p_{X Y}$, REMOTE-OT) for computing REMOTE $\binom{m}{1}-\mathrm{OT}_{2}^{n}$ for inputs $X$ and $Y$, where $p_{X Y}$ has full support, must satisfy

$$
H\left(M_{31}\right) \geq n m, \quad H\left(M_{23}\right) \geq n+\log m, \quad \text { and } \quad H\left(M_{12}\right) \geq n m+\log m
$$

In the full version we also give two other examples (GROUP-ADD, CONTROLLEDERASURE) which have communication-ideal protocols.
Separating Secure and Insecure Computation. A basic question of secure computation is whether it needs more bits to be communicated than the inputsize itself (which suffices for insecure computation). While natural to expect, it is not easy to prove this. In their restricted model, [17] showed a non-explicit result, that for securely computing most Boolean functions on the domain $\{0,1\}^{n} \times$ $\{0,1\}^{n}$, Charlie is required to receive at least $3 n-4$ bits, which is significantly more than the $2 n$ bits sufficient for insecure computation.

REMOTE $\binom{2}{1}-\mathrm{OT}_{2}^{n}$ from above already gives us an explicit example of a function where this is true: the total input size is $2 n+1$, but the communication is at least $H\left(M_{31}\right)+H\left(M_{23}\right) \geq 3 n+1$. To present an easy comparison to the lower bound of [17], we can consider a symmetrized variant of REMOTE $\binom{2}{1}-\mathrm{OT}_{2}^{n}$, in which two instances of REMOTE $\binom{2}{1}-\mathrm{OT}_{2}^{n}$ are combined, one in each direction. More specifically, $X=\left(A_{0}, A_{1}, a\right)$ where $A_{0}, A_{1}$ are of length $(n-1) / 2$ (for an odd $n$ ) and $a$ is a single bit; similarly $Y=\left(B_{0}, B_{1}, b\right)$; the output of the function is defined as an $(n-1)$ bit string $f(X, Y)=\left(A_{b}, B_{a}\right)$. Considering (say) the uniform input distribution over $X, Y$, the bounds for REMOTE $\binom{2}{1}-\mathrm{OT}_{2}^{n}$ add up to give us $H\left(M_{31}\right) \geq 3(n-1) / 2+1$ and $H\left(M_{23}\right) \geq 3(n-1) / 2+1$, so that the communication with Charlie is lower-bounded by $H\left(M_{31}\right)+H\left(M_{23}\right) \geq 3 n-1$.

This compares favourably with the bound of [17] in many ways: our lower bound holds even in a model that allows interaction; in particular, this makes
the gap between insecure computation ( $n-1$ bits in our case, $2 n$ bits for [17]) and secure computation (about $3 n$ bits for both) somewhat larger. More importantly, our lower bound is explicit (and tight for the specific function we use), whereas that of [17] is existential. However, our bound does not subsume that of [17], who considered Boolean functions. Our results do not yield a bound significantly larger than the input size, when the output is a single bit. It appears that this regime is more amenable to combinatorial arguments, as pursued in [17], rather than information theoretic arguments.
Communication Complexity of Securely Computing and. We define the 3 -party and function as follows: Alice has an input bit $X$, Bob has an input bit $Y$ and Charlie should obtain $Z=f(X, Y)=X \wedge Y$. In the full version, we compute the following lower bound.

Theorem 6. Any secure protocol $\Pi$ ( $p_{X Y}$, AND) for computing AND for inputs $X$ and $Y$, where $p_{X Y}$ has full support over $\{0,1\}^{n} \times\{0,1\}^{n}$, must satisfy

$$
H\left(M_{31}\right) \geq n \log (3), \quad H\left(M_{23}\right) \geq n \log (3), \quad \text { and } \quad H\left(M_{12}\right) \geq n(1.826)
$$

The best known protocol for and (due to [17]) achieves $H\left(M_{12}\right)=1+\log (3)$, $H\left(M_{23}\right)=H\left(M_{31}\right)=\log (3)$. Our lower bounds on $H\left(M_{31}\right)$ and $H\left(M_{23}\right)$ match this, but there is a gap for $H\left(M_{12}\right)$ : an upper bound of $1+\log (3) \approx 2.585$ against a lower bound of 1.826. Closing this gap remains an open problem.

## 5 Conclusion

In this work we presented new tools to obtain lower bounds on the communication complexity of secure 3-party computation, and showed that they yield tight bounds for interesting examples. However, the general problem of obtaining tight lower bounds for communication complexity of secure computation is wide open; indeed, their implications to circuit lower bounds presents a "barrier" to obtaining super-linear bounds for explicit functions. We propose a possibly easier open problem: do there exist Boolean functions with super-linear communication complexity for secure computation? Note that lower bounds on circuit complexity do not directly translate to lower bounds on communication complexity of secure computation, as established by a sub-exponential upper bound of $2^{\widetilde{O}(\sqrt{n})}$ for the latter [1. Though it is plausible that for random Boolean functions, the actual communication cost is $2^{\Omega\left(n^{\epsilon}\right)}$ for some $\epsilon>0$, none of the current techniques appear capable of delivering such a result.

## References

1. Beimel, A., Ishai, Y., Kumaresan, R., Kushilevitz, E.: On the cryptographic complexity of the worst functions. In: Lindell, Y. (ed.) TCC 2014. LNCS, vol. 8349, pp. 317-342. Springer, Heidelberg (2014)
2. Beimel, A., Orlov, I.: Secret sharing and non-Shannon information inequalities. IEEE Transactions on Information Theory 57(9), 5634-5649 (2011)
3. Ben-Or, M., Goldwasser, S., Wigderson, A.: Completeness theorems for noncryptographic fault-tolerant distributed computation. In: Proc. 20th STOC, pp. $1-10$. ACM (1988)
4. Blundo, C., De Santis, A., Di Crescenzo, G., Gaggia, A.G., Vaccaro, U.: Multisecret sharing schemes. In: Desmedt, Y.G. (ed.) Advances in Cryptology CRYPTO 1994. LNCS, vol. 839, pp. 150-163. Springer, Heidelberg (1994)
5. Blundo, C., Santis, A.D., Persiano, G., Vaccaro, U.: Randomness complexity of private computation. Computational Complexity 8(2), 145-168 (1999)
6. Braun, G., Pokutta, S.: Common information and unique disjointness. In: FOCS, pp. 688-697 (2013)
7. Chakrabarti, A., Shi, Y., Wirth, A., Yao, A.C.-C.: Informational complexity and the direct sum problem for simultaneous message complexity. In: FOCS, pp. 270278. IEEE (2001)
8. Chaum, D.: The spymasters double-agent problem: Multiparty computations secure unconditionally from minorities and cryptographically from majorities. In: Brassard, G. (ed.) Advances in Cryptology - CRYPTO 1989. LNCS, vol. 435, pp. 591-602. Springer, Heidelberg (1990)
9. Chaum, D., Crépeau, C., Damgård, I.: Multiparty unconditionally secure protocols. In: Proc. 20th STOC, pp. 11-19. ACM (1988)
10. Chor, B., Kushilevitz, E.: A communication-privacy tradeoff for modular addition. Inf. Process. Lett. 45(4), 205-210 (1993)
11. Csiszár, I., Narayan, P.: Secrecy capacities for multiple terminals. IEEE Transactions on Information Theory 50(12), 3047-3061 (2004)
12. Data, D., Dey, B.K., Mishra, M., Prabhakaran, V.M.: How to securely compute the modulo-two sum of binary sources, arXiv, 1405.2555 (preprint 2014)
13. Data, D., Prabhakaran, V.M.: Communication requirements for secure computation. In: Proc. 51st Annual Allerton Conference on Communication, Control, and Computing (2013)
14. Data, D., Prabhakaran, V.M., Prabhakaran, M.M.: On the communication complexity of secure computation, arXiv, 1311.7584 (preprint, 2014)
15. Dodis, Y., Micali, S.: Lower bounds for oblivious transfer reductions. In: Stern, J. (ed.) EUROCRYPT 1999. LNCS, vol. 1592, pp. 42-55. Springer, Heidelberg (1999)
16. Dodis, Y., Micali, S.: Parallel reducibility for information-theoretically secure computation. In: Bellare, M. (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 74-92. Springer, Heidelberg (2000)
17. Feige, U., Kilian, J., Naor, M.: A minimal model for secure computation (extended abstract). In: STOC, pp. 554-563. ACM (1994)
18. Fischer, M.J., Lynch, N.A.: A lower bound for the time to assure interactive consistency. Inf. Process. Lett. 14(4), 183-186 (1982)
19. Fitzi, M., Garay, J.A., Gollakota, S., Pandu Rangan, C., Srinathan, K.: Roundoptimal and efficient verifiable secret sharing. In: Halevi, S., Rabin, T. (eds.) TCC 2006. LNCS, vol. 3876, pp. 329-342. Springer, Heidelberg (2006)
20. Fitzi, M., Hirt, M., Maurer, U.M.: General adversaries in unconditional multiparty computation. In: Lam, K.-Y., Okamoto, E., Xing, C. (eds.) ASIACRYPT 1999. LNCS, vol. 1716, pp. 232-246. Springer, Heidelberg (1999)
21. Franklin, M.K., Yung, M.: Communication complexity of secure computation (extended abstract). In: STOC, pp. 699-710. ACM (1992)
22. Gács, P., Körner, J.: Common information is far less than mutual information. Problems of Control and Information Theory 2(2), 149-162 (1973)
23. Gál, A., Rosén, A.: Omega $(\log n)$ lower bounds on the amount of randomness in 2-private computation. SIAM J. Comput. 34(4), 946-959 (2005)
24. Gennaro, R., Ishai, Y., Kushilevitz, E., Rabin, T.: The round complexity of verifiable secret sharing and secure multicast. In: Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing, pp. 580-589. ACM (2001)
25. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC, pp. 169178. ACM (2009)
26. Hirt, M., Lucas, C., Maurer, U., Raub, D.: Passive corruption in statistical multiparty computation. In: Smith, A. (ed.) ICITS 2012. LNCS, vol. 7412, pp. 129-146. Springer, Heidelberg (2012), http://eprint.iacr.org/2012/272
27. Hirt, M., Maurer, U.M.: Complete characterization of adversaries tolerable in secure multi-party computation (extended abstract). In: PODC, pp. 25-34 (1997)
28. Ishai, Y., Kushilevitz, E.: On the hardness of information-theoretic multiparty computation. In: Cachin, C., Camenisch, J.L. (eds.) EUROCRYPT 2004. LNCS, vol. 3027, pp. 439-455. Springer, Heidelberg (2004)
29. Ishai, Y., Kushilevitz, E., Meldgaard, S., Orlandi, C., Paskin-Cherniavsky, A.: On the power of correlated randomness in secure computation. In: Sahai, A. (ed.) TCC 2013. LNCS, vol. 7785 , pp. 600-620. Springer, Heidelberg (2013)
30. Katz, J., Koo, C.-Y., Kumaresan, R.: Improving the round complexity of vss in point-to-point networks. Inf. Comput. 207(8), 889-899 (2009)
31. Kushilevitz, E.: Privacy and communication complexity. In: FOCS, pp. 416-421. IEEE (1989)
32. Kushilevitz, E., Mansour, Y.: Randomness in private computations. SIAM J. Discrete Math. 10(4), 647-661 (1997)
33. Kushilevitz, E., Nisan, N.: Communication complexity. Cambridge University Press, New York (1997)
34. Lee, E.J., Abbe, E.: A Shannon approach to secure multi-party computations, arXiv, 1401.7360 (preprint, 2014)
35. Maurer, U.M., Wolf, S.: Secret-key agreement over unauthenticated public channels iii: Privacy amplification. IEEE Transactions on Information Theory 49(4), 839851 (2003)
36. Naor, M., Nissim, K.: Communication preserving protocols for secure function evaluation. In: STOC, pp. 590-599 (2001)
37. Orlitsky, A., Roche, J.R.: Coding for computing. IEEE Transactions on Information Theory 47(3), 903-917 (2001)
38. Patra, A., Choudhary, A., Rabin, T., Rangan, C.P.: The round complexity of verifiable secret sharing revisited. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 487-504. Springer, Heidelberg (2009)
39. Prabhakaran, M.M., Prabhakaran, V.M.: Communication complexity lower bounds from assisted common information. Under Preparation
40. Prabhakaran, V.M., Prabhakaran, M.M.: Assisted common information with an application to secure two-party sampling. IEEE Transactions on Information Theory $60(6), 3413-3434$ (2014)
41. Winkler, S., Wullschleger, J.: On the efficiency of classical and quantum oblivious transfer reductions. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 707723. Springer, Heidelberg (2010); Full version, arXiv, 1205.5136
42. Wolf, S., Wullschleger, J.: New monotones and lower bounds in unconditional twoparty computation. IEEE Transactions on Information Theory 54(6), 2792-2797 (2008)
43. Wyner, A.D.: The wire-tap channel. The Bell System Technical Journal 54(8), 1355-1387 (1975)
44. Yao, A.C.-C.: Some complexity questions related to distributive computing (preliminary report). In: STOC, pp. 209-213. ACM (1979)

[^0]:    * Research supported in part by ITRA, Media Lab Asia, India.
    ** Research supported in part by NSF grants 1228856 and 0747027.
    *** Research funded in part by ITRA, Media Lab Asia, India and a Ramanujan fellowship from DST, India.

[^1]:    ${ }^{4}$ Since the parties are computationally unbounded, there is no need to allow private randomness as part of the state; randomness for a party can always be resampled at every round conditioned on the inputs, outputs and messages in that party's view.

[^2]:    ${ }^{5}$ We remark that our bounds do not apply to a relaxed security setting sometimes considered in the information theory literature: there the error in computation/security is only required to go to 0 as the size of the input grows to infinity. [12] gives an example where there is a strict gap between the communication complexity under this relaxed setting and the perfect security setting of this paper.

