# Towards Characterizing Securely Computable Two-Party Randomized Functions 

Deepesh Data ${ }^{(\boxtimes)}$ and Manoj Prabhakaran<br>Department of Computer Science and Engineering, Indian Institute of Technology Bombay, Mumbai, India<br>deepesh.data@gmail.com


#### Abstract

A basic question of cryptographic complexity is to combinatorially characterize all randomized functions which have informationtheoretic semi-honest secure 2-party computation protocols. The corresponding question for deterministic functions was answered almost three decades back, by Kushilevitz [Kus89]. In this work, we make progress towards understanding securely computable randomized functions. We bring tools developed in the study of completeness to bear on this problem. In particular, our characterizations are obtained by considering only symmetric functions with a combinatorial property called simplicity [MPR12].

Our main result is a complete combinatorial characterization of randomized functions with ternary output kernels, that have informationtheoretic semi-honest secure 2-party computation protocols. In particular, we show that there exist simple randomized functions with ternary output that do not have secure computation protocols. (For deterministic functions, the smallest output alphabet size of such a function is 5 , due to an example given by Beaver [Bea89].)

Also, we give a complete combinatorial characterization of randomized functions that have 2-round information-theoretic semi-honest secure 2party computation protocols.

We also give a counter-example to a natural conjecture for the full characterization, namely, that all securely computable simple functions have secure protocols with a unique transcript for each output value. This conjecture is in fact true for deterministic functions, and - as our results above show - for ternary functions and for functions with 2-round secure protocols.


## 1 Introduction

Understanding the nature of secure multiparty computation has been a key problem in modern cryptography, ever since the notion was introduced. While
D. Data-This work was done while the author was a Ph.D. student at the Tata Institute of Fundamental Research in Mumbai.
this has been a heavily researched area, some basic problems have remained open. In this work we explore the following fundamental question:

Which randomized functions have information-theoretic semi-honest secure 2-party computation protocols?

The corresponding question for deterministic functions was answered almost three decades back, by Kushilevitz [Kus89] (originally, restricted to symmetric functions, in which both parties get the same output). The dual question of which functions are complete, initiated by Kilian [Kil88], has also been fully resolved, for semi-honest [MPR12] and even active security [KMPS14]. However, the above question itself has seen little progress since 1989.

In this work, we make progress towards understanding securely computable randomized functions. (Throughout this paper, security will refer to informationtheoretic semi-honest security.) We bring tools developed in the study of completeness to bear on this problem. In particular, our characterizations are obtained by considering only symmetric functions with a combinatorial property called simplicity [MPR12]. (As shown in [MPR12], a function is semi-honest securely computable if and only if it is simple, and a related function called its "kernel" - which is always a simple function - is securely computable.)

One may start off by attempting to generalize the result of Kushilevitz [Kus89] so that it applies to randomized functions as well. This characterization showed that any securely computable deterministic function has a secure protocol in which the two parties take turns to progressively reveal more and more information about their respective inputs - by restricting each input to smaller and smaller subsets - until there is exactly enough information to evaluate the function. However, a naïve generalization of this result fails for randomized functions, as it is possible for a securely computable function to have every output value in the support of every input combination; thus the input spaces cannot be shrunk at all during a secure protocol.

A more fruitful approach would be to consider the protocol for deterministic functions as partitioning the output space at each step, and choosing one of the parts. This is indeed true when considering deterministic functions which are simple. Such a protocol results in a unique transcript for each output value. An a priori promising conjecture would be that every securely computable simple function - deterministic or randomized - has such a unique-transcript protocol. Unfortunately, this conjecture turns out to be false.

However, for small output alphabets, we can prove that this conjecture holds. Indeed, we show that the exact threshold on the alphabet size where this conjecture breaks down is 4 . When the output alphabet size of a simple function is at most 3, we give a combinatorial characterization of secure computability; our characterization implies that such functions do have unique-transcript secure protocols. We also characterize simple functions which have two-round secure protocols, which again turn out to all have unique-transcript protocols.

We leave the full characterization as an important open problem.

## Our Results

- Our main result is a complete combinatorial characterization of randomized functions with ternary output kernels, that have information-theoretic semihonest secure 2-party computation protocols. In particular, we show that there exist simple randomized functions with ternary output that do not have secure computation protocols. (For deterministic functions, the smallest output alphabet size of such a function is 5 , due to an example given by Beaver [Bea89].)
- We also give a complete combinatorial characterization of randomized functions that have 2-round information-theoretic semi-honest secure 2-party computation protocols.
- We also give a counter-example to a natural conjecture for the full characterization, namely, that all securely computable simple functions have secure protocols with a unique transcript for each output value. This conjecture is in fact true for deterministic functions, and - as our results above show - for ternary functions and for functions with 2-round secure protocols.


### 1.1 Technical Overview

Prior work [MPR12, MPR13] lets us focus on symmetric functions: A randomized function $F$ is securely realizable if and only if it is "isomorphic" - i.e., essentially equivalent, up to sampling additional local outputs - to a symmetric function $G$ called its kernel, and $G$ itself has a secure protocol; see Theorem 3. Further the kernel of $F$ is easy to find and explicitly defined in Definition 2. Hence the problem of characterizing secure computability of general randomized functions reduces to the problem of characterizing secure computability of randomized functions which are kernels. Such functions are symmetric and simple (a symmetric function $G$ is simple, if $\operatorname{Pr}[G(x, y)=z]=\rho(x, z) \cdot \sigma(y, z)$ for some fixed functions $\rho: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}_{+}$and $\sigma: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathbb{R}_{+}$, where $\mathcal{X}$ and $\mathcal{Y}$ are Alice's and Bob's input domains and $\mathcal{Z}$ is their common output domain). As such, we work with symmetric and simple functions.

Characterizing Ternary-Kernel Functions. Our main result could be stated as follows:

Theorem 1. If a randomized function $F$ has a kernel $G$ with an output alphabet of size at most 3, then $F$ is securely computable if and only if $F$ is simple and there is some ordering of $G$ 's output alphabet $\mathcal{Z}$ as $\left(z_{1}, z_{2}, z_{3}\right)$ and two functions $q: \mathcal{X} \rightarrow[0,1]$ and $r: \mathcal{Y} \rightarrow[0,1]$, such that one of the following holds:

$$
\begin{aligned}
& \operatorname{Pr}\left[G(x, y)=z_{1}\right]=q(x) \\
& \operatorname{Pr}\left[G(x, y)=z_{2}\right]=(1-q(x)) \cdot r(y) \quad \text { or } \operatorname{Pr}\left[G(x, y)=z_{2}\right]=(1-r(y)) \cdot q(x) \\
& \operatorname{Pr}\left[G(x, y)=z_{3}\right]=(1-q(x)) \cdot(1-r(y)) \quad \operatorname{Pr}\left[G(x, y)=z_{3}\right]=(1-r(y)) \cdot(1-q(x))
\end{aligned}
$$

Note that if the first set of conditions holds, there is a secure protocol in which Alice either sends $z_{1}$ as the output to Bob (with probability $q(x)$ ) or asks Bob to pick the output; if Bob is asked to pick the output he sends back either $z_{2}$ as the output (with probability $r(y)$ ) or $z_{3}$ otherwise. If the second condition holds there is a symmetric protocol with Bob making the first move.

Surprisingly, these are the only two possibilities for $G$ to have a secure protocol. To prove this, however, takes a careful analysis of the linear-algebraic constraints put on a protocol by the security definition and the fact that the function is simple. We start by observing that a secure protocol for a symmetric simple function must have a simulator that can simulate the transcript of an execution merely from the output (rather than the corrupt party's input and the output). ${ }^{1}$ Then, supposing that the first message in the protocol is a single bit sent by Alice, we identify that there is a quantity independent of either party's input, denoted by $\phi(z)$, that gives the probability of the first message being 0 , conditioned on the output being $z$. Specifying these quantities, $\phi(z)$ at each round fully specifies the protocol. We show that $\sum_{z \in \mathcal{Z}} \operatorname{Pr}[G(x, y)=z] \cdot \phi(z)=q(x)$, a quantity independent of $y$. By carefully analyzing the constraints arising from these equations, we prove Theorem 1.

Characterizing Functions Having 2-Round Protocols. Our second result is as follows:

Theorem 2. A function $F$ has a 2-round secure protocol iff its kernel has a 2-round unique-transcript protocol.

Observe that $F$ has a 2-round secure protocol iff its kernel has one too, as $F$ is isomorphic to its kernel and a secure protocol for a function can be transformed to one for an isomorphic function without changing the communication involved. What needs to be proven is that if the kernel (or any symmetric simple function) has a 2 -round secure protocol, then it has a 2-round unique-transcript protocol. We do this by identifying an equivalence relation among the outputs, such that any 2-round protocol with (say) Alice making the first move will have Alice's input only influencing which equivalence class of the output is chosen, and then Bob's input influences which output is chosen, given its equivalence class.

### 1.2 Related Work

There has been a large body of work regarding the complexity of secure multiparty and 2-party computation. [MPR13] surveys many of the important results in the area. Kushilevitz [Kus89] gave a combinatorial characterization of securely computable two-party deterministic functions (with perfect security), along with

[^0]a generic round-optimal secure protocol for functions that satisfy the characterization condition. This was later extended to statistical security and also to security against active corruption [MPR09, KMQR09].

Among randomized functions in which only Bob gets any output, simple functions have a deterministic kernel corresponding to a function of Alice's input alone, and hence are always securely computable. This observation was already made by Kilian [Kil00]. Recently, Data [Dat16] considered the same problem with a probability distribution on the inputs, and gave communication-optimal protocols in different security settings.

## 2 Preliminaries

A general randomized function is denoted by a conditional probability distribution $p_{Z_{A} Z_{B} \mid X Y}$, where $X, Y, Z_{A}, Z_{B}$ take values in finite alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}_{A}, \mathcal{Z}_{B}$, respectively. In a protocol for computing this function, when Alice and Bob are given inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, respectively, they should output $z_{A} \in$ $\mathcal{Z}_{A}$ and $z_{B} \in \mathcal{Z}_{B}$ respectively, such that $\left(z_{A}, z_{B}\right)$ is distribute according to $p_{Z_{A} Z_{B} \mid X=x, Y=y}$.

Notation. We shall consider various discrete random variables (inputs, outputs, protocol messages). We denote the probability mass function of a random variable $U$ by $p_{U}$. For random variables $(U, V)$, we denote the conditional probability mass function of $U$ conditioned on $V$ by $p_{U \mid V} \cdot[n]$ denotes the set $\{1, \cdots, n\}$.

Protocols. We consider computationally unbounded two-party protocols, without any setup. Such a protocol $\Pi$ is fully defined by the input and output domains, the next message functions and the output functions for Alice and Bob. Let $\left(\mathcal{X}, \mathcal{Z}_{A}, \operatorname{next}_{\Pi}^{A}\right.$, out $\left.A_{\Pi}^{A}\right)$ and $\left(\mathcal{Y}, \mathcal{Z}_{B}, \operatorname{next}_{\Pi}^{B}\right.$, out $\left.{ }_{\Pi}^{B}\right)$ denote the input domain, output domain, the next message function and the output function, for Alice and Bob, respectively. The functions are all potentially randomized. next ${ }_{\Pi}^{A}$ and next ${ }_{I}^{B}$ output the next message given the transcript so far and the local input. (Note that in the information-theoretic setting, protocols need not maintain local state.) Similarly, out ${ }_{\Pi}^{A}$ and out ${ }_{I}^{B}$ map the transcript and local input to a local output in $\mathcal{Z}_{A}$ and $\mathcal{Z}_{B}$, respectively.

In running a protocol, Alice and Bob are first given inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, respectively. Then they take turns sending messages to each other according to their next message functions, and in the end (recognizable by both parties) each party produces an output according to its output function. We assume that the protocol terminates with probability one after a finite number rounds.

Running a protocol induces a distribution over the set of all possible (complete) transcripts, $\mathcal{M}$. The protocol $\Pi$ induces a conditional probability distribution $\operatorname{Pr}_{\Pi}[m \mid x, y]$, for every input $x \in \mathcal{X}, y \in \mathcal{Y}$ and every $m \in \mathcal{M}$. Suppose that $m=\left(m_{1}, m_{2}, \ldots,\right)$ is the transcript generated by $\Pi$ when parties have inputs $x$ and $y$, where $m_{i}$ 's, for odd $i$, are sent by Alice, and $m_{i}$ 's, for even $i$, are sent by Bob. Note that during the execution of a protocol, a message sent
by any party is determined by its input and all the messages it has exchanged so far; and conditioned on these two, the message is independent of the other party's input. Using this we can write $\operatorname{Pr}_{\Pi}[m \mid x, y]=\alpha(m, x) \beta(m, y)$, where $\alpha(m, x)=\Pi_{i: i}$ is odd $\operatorname{Pr}_{\Pi}\left[m_{i} \mid m_{<i}, x\right]$, and $\beta(m, y)=\Pi_{i: i}$ is even $\operatorname{Pr}_{\Pi}\left[m_{i} \mid m_{<i}, y\right]$.

Secure Protocols. Throughout this paper security refers to information-theoretic semi-honest security. We restrict ourselves to finite functions and perfect security. We remark that all our results can be extended to statistical security, as we shall show in the full version.

Definition 1. A protocol $\Pi$ for computing a function $p_{Z_{A} Z_{B} \mid X Y}$, with transcript space $\mathcal{M}$, is said to be (perfectly semi-honest) secure iff there exist functions $S_{1}: \mathcal{M} \times \mathcal{X} \times \mathcal{Z}_{A} \rightarrow[0,1]$ and $S_{2}: \mathcal{M} \times \mathcal{Y} \times \mathcal{Z}_{B} \rightarrow[0,1]$ such that $\operatorname{Pr}_{\Pi}\left[m \mid x, y, z_{A}, z_{B}\right]=S_{1}\left(m, x, z_{A}\right)=S_{2}\left(m, y, z_{B}\right)$ for all $m \in \mathcal{M}$ and $x, y, z_{A}, z_{B}$ such that $p_{Z_{A} Z_{B} \mid X Y}\left(z_{A}, z_{B} \mid x, y\right)>0$.

Kernel and Simple Functions. Maji et al. [MPR12] simplified the secure computation of a general randomized function $p_{Z_{A} Z_{B} \mid X Y}$ to secure computation of a symmetric randomized function $p_{Z \mid X Y}$. For that they defined weighted characteristic bipartite graph of a randomized function $p_{Z_{A} Z_{B} \mid X Y}$ as $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)=$ ( $V, E, \mathrm{wt}$ ), where
$-V=\left(\mathcal{X} \times \mathcal{Z}_{A}\right) \cup\left(\mathcal{Y} \times \mathcal{Z}_{B}\right)$,

- $E=\left\{\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right): p_{Z_{A} Z_{B} \mid X Y}\left(z_{A}, z_{B} \mid x, y\right)>0\right\}$, and
- the weight function wt: $\left(\mathcal{X} \times \mathcal{Z}_{A}\right) \times\left(\mathcal{Y} \times \mathcal{Z}_{B}\right) \rightarrow[0,1]$ is defined as

$$
\mathrm{wt}\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right):=\frac{p_{Z_{A} Z_{B} \mid X Y}\left(z_{A}, z_{B} \mid x, y\right)}{|\mathcal{X}| \times|\mathcal{Y}|}
$$

Note that if $\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right) \notin E$, then $\mathrm{wt}\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)=0$.
Let $k$ be the number of connected components in the above-defined graph. We say that $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)=(V, E, \mathrm{wt})$ is a product distribution graph, if there exist probability distributions $p$ over $\mathcal{X} \times \mathcal{Z}_{A}, q$ over $\mathcal{Y} \times \mathcal{Z}_{B}$, and $c$ over $[k]$, such that for all $\left(x, z_{A}\right) \in \mathcal{X} \times \mathcal{Z}_{A}$ and $\left(y, z_{B}\right) \times \mathcal{Y} \times \mathcal{Z}_{B}$, if $\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)$ lies in the $j$ th connected component of $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)=(V, E, \mathrm{wt})$, then $\mathrm{wt}\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)=$ $p\left(x, z_{A}\right) \cdot q\left(y, z_{B}\right) / c_{j}$, otherwise $\operatorname{wt}\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)=0$.

Definition 2 (Kernel - Common-information in a randomized function [MPR12]). The kernel of a randomized function $p_{Z_{A} Z_{B} \mid X Y}$ is a symmetric randomized function, which takes $x$ and $y$ from the parties and samples $\left(z_{A}, z_{B}\right)$ according to $p_{Z_{A} Z_{B} \mid X=x, Y=y}$. Then it outputs to both parties the connected component of $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)$ which contains the edge $\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)$.

Note that the kernel of $p_{Z_{A} Z_{B} \mid X Y}$ is a symmetric randomized function. We denote it by $p_{Z \mid X Y}$, where the alphabet of $Z$ is $\mathcal{Z}=\left\{z_{1}, z_{2}, \ldots, z_{k}\right\}$, where $k$ is the number of connected components in $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)$. The kernel $p_{Z \mid X Y}$ is defined as follows: for every $j \in[k], p_{Z \mid X Y}\left(z_{j} \mid x, y\right):=$
$\sum_{\left(z_{A}, z_{B}\right)} p_{Z_{A} Z_{B} \mid X Y}\left(z_{A}, z_{B} \mid x, y\right)$, where summation is taken over all $\left(z_{A}, z_{B}\right)$ 's such that $\left(\left(x, z_{A}\right),\left(y, z_{B}\right)\right)$ lies in the $j^{\text {th }}$ connected component in $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)$. The following theorem was proved in [MPR12].

Theorem 3 [MPR12, Theorem 3]. A randomized function $p_{Z_{A} Z_{B} \mid X Y}$ is securely computable if and only if $\mathcal{G}\left(p_{Z_{A} Z_{B} \mid X Y}\right)$ is a product distribution graph and the kernel of $p_{Z_{A} Z_{B} \mid X Y}$ is securely computable.

Theorem 3 reduces secure computablity of $p_{Z_{A} Z_{B} \mid X Y}$ to secure computability of the kernel of $p_{Z_{A} Z_{B} \mid X Y}$ and a simple combinatorial check on $p_{Z_{A} Z_{B} \mid X Y}$ (which is to check whether the weighted characteristic bipartite graph of $p_{Z_{A} Z_{B} \mid X Y}$ is a product distribution graph or not). ${ }^{2}$
Definition 3. A symmetric randomized function $p_{Z \mid X Y}$ is said to be simple if there exist two functions $\rho: \mathcal{X} \times \mathcal{Z} \rightarrow \mathbb{R}_{+}$and $\sigma: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathbb{R}_{+}$such that for all $x \in \mathcal{X}, y \in \mathcal{Y}$, and $z \in \mathcal{Z}, p_{Z \mid X Y}(z \mid x, y)=\rho(x, z) \cdot \sigma(y, z)$.

Here $\mathbb{R}_{+}$denotes the set of non-negative real numbers. For deterministic functions, instead of $\mathbb{R}_{+}$, one can take $\{0,1\}$ in the above definition.
Remark 1. The original definition of a simple function given in [MPR12] seems to be different from our definition. There it was defined for a general randomized function $p_{Z_{A} Z_{B} \mid X Y}$, which defined simplicity in terms of isomorphism between a function and its kernel, whereas we defined simplicity for symmetric functions only. Since isomorphic functions are essentially equivalent - up to sampling additional local outputs - and the kernel of a general randomized function is a symmetric function, our definition of simplicity is equivalent to the one given in [MPR12].

Note that the Kernel of a securely computable randomized function $p_{Z_{A} Z_{B} \mid X Y}$ is a simple function. As shown in [MPR12], a secure protocol for the kernel can be transformed to one for the original function itself, and vice versa, without changing the communication involved. Thus we shall focus on characterizing secure computability for kernel functions, which are all symmetric, simple functions.

The combinatorial definition of simplicity above will be crucially used in our analysis. Indeed, the factorization property clarifies otherwise obscure connections and elusive constraints.

For protocols for simple and symmetric functions, the output can be written as a function merely of the transcript and we can simulate the transcript just based on the (common) output, without needing either party's input. We prove the following lemma in Appendix A.
Lemma 1. If $\Pi$ is a perfectly semi-honest secure protocol for a simple symmetric function $p_{Z \mid X Y}$, then there are (deterministic) functions out ${ }_{\Pi}: \mathcal{M} \rightarrow \mathcal{Z}$ and $S: \mathcal{M} \rightarrow[0,1]$ such that for all $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$ and $m \in \mathcal{M}$,

$$
\begin{array}{rlrl}
\operatorname{out}_{\Pi}^{A}(m, x) & =\operatorname{out}_{\Pi}^{B}(m, y)=\operatorname{out}_{\Pi}(m) \quad \text { if } \quad \operatorname{Pr}_{\Pi}[m \mid x, y]>0 \\
\operatorname{Pr}_{\Pi}[m \mid x, y, z] & =S(m, z) & \text { if } \quad p_{Z \mid X Y}(z \mid x, y)>0
\end{array}
$$

[^1]Note that above, if $z \neq \operatorname{out}_{\Pi}(m), S(m, z)=\operatorname{Pr}_{\Pi}[m \mid x, y, z]=0$. By writing $\mu(m)=S\left(m\right.$, out $\left._{\Pi}(m)\right)$ we have the following: for all $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$ s.t. $p_{Z \mid X Y}(z \mid x, y)>0$, and all $m \in \mathcal{M}$, we have

$$
\operatorname{Pr}_{\Pi}[m \mid x, y, z]= \begin{cases}\mu(m) & {\text { if } \operatorname{out}_{\Pi}(m)=z}^{0}  \tag{1}\\ \text { otherwise }\end{cases}
$$

Thus, for each $z \in \mathcal{Z}$ (such that for some $\left.(x, y), p_{Z \mid X Y}(z \mid x, y)>0\right), \mu$ defines a probability distribution over $\left\{m \in \mathcal{M}: \operatorname{out}_{\Pi}(m)=z\right\}$. Also, since $\operatorname{Pr}_{\Pi}[m \mid x, y]=\operatorname{Pr}_{\Pi}[m, z \mid x, y]$, for $z=$ out $_{\Pi}(m)$, and $\operatorname{Pr}_{\Pi}[m, z \mid x, y]=$ $\operatorname{Pr}_{\Pi}[m \mid x, y, z] \cdot \operatorname{Pr}_{\Pi}[z \mid x, y]=\mu(m) \cdot p_{Z \mid X Y}(z \mid x, y)$ we have, for all $x \in \mathcal{X}, y \in$ $\mathcal{Y}, m \in \mathcal{M}$,

$$
\begin{equation*}
\operatorname{Pr}_{\Pi}[m \mid x, y]=\mu(m) \cdot p_{Z \mid X Y}\left(\operatorname{out}_{\Pi}(m) \mid x, y\right) . \tag{2}
\end{equation*}
$$

A Normal Form for $p_{Z \mid X Y}$. For a symmetric randomized functionality $p_{Z \mid X Y}$, we define the relation $x \equiv x^{\prime}$ for $x, x^{\prime} \in \mathcal{X}$ to hold, if $\forall y \in \mathcal{Y}, z \in \mathcal{Z}, p(z \mid x, y)=$ $p\left(z \mid x^{\prime}, y\right)$; similarly we define $y \equiv y^{\prime}$ for $y, y^{\prime} \in \mathcal{Y}$. We define $z \equiv z^{\prime}$ for $z, z^{\prime} \in \mathcal{Z}$, if there exists a constant $c>0$ such that $\forall x \in \mathcal{X}, y \in \mathcal{Y}, p(z \mid x, y)=c \cdot p\left(z^{\prime} \mid x, y\right)$. We say that $p_{Z \mid X Y}$ is in normal form if $x \equiv x^{\prime} \Rightarrow x=x^{\prime}, y \equiv y^{\prime} \Rightarrow y=y^{\prime}$, and $z \equiv z^{\prime} \Rightarrow z=z^{\prime}$.

It is easy to see that any $p_{Z \mid X Y}$ can be transformed into one in normal form $p_{Z^{*} \mid X^{*} Y^{*}}$ with possibly smaller alphabets, so that $p_{Z \mid X Y}$ is securely computable if and only if $p_{Z^{*} \mid X^{*} Y^{*}}$ is securely computable. We will assume in this paper that $p_{Z \mid X Y}$ is in normal form.

For the ease of notation, in this paper we often denote a randomized function $p_{Z \mid X Y}$ by an equivalent function $f$, such that $f(x, y, z)=p_{Z \mid X Y}(z \mid x, y)$, for every $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$. We may use $f$ and $p_{Z \mid X Y}$ interchangeably.

Unique-Transcript Protocols. A unique transcript protocol $\Pi$ for a symmetric function is one in which each output $z \in \mathcal{Z}$ has a unique transcript that can result in it: i.e., for every $z \in \mathcal{Z}$, there is at most one $m \in \mathcal{M}$ such that $\operatorname{out}_{\Pi}(m)=z$. Such a protocol is always a secure protocol for the function it computes (and has a deterministic simulator, which when given an output $z$ assigns probability 1 to the unique transcript $m$ such that out $\left.{ }_{\Pi}(m)=z\right)$.

It follows from [Kus89] that every securely computable simple deterministic function has a unique-transcript secure protocol. The subset of securely computable simple randomized functions that we characterize also have uniquetranscript protocols. However, we shall show an example of a securely computable function (just outside the sets we characterize) which does not have any unique-transcript protocol. Understanding such functions remains the next step in fully characterizing securely computable randomized functions.

## 3 Characterization of Functions up to Ternary Output Alphabet

As mentioned earlier, we shall represent a randomized function $p_{Z \mid X Y}$ by an equivalent function $f: \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \rightarrow[0,1]$, such that $f(x, y, z)=p_{Z \mid X Y}(z \mid x, y)$.

For a simple function $f$, we shall write $f=(\rho, \sigma)$ where $f(x, y, z)=\rho(x, z)$. $\sigma(y, z)$. Note that for every $x \in \mathcal{X}, y \in \mathcal{Y}$ we have $\sum_{z \in \mathcal{Z}} f(x, y, z)=1$.

### 3.1 Compact Representation of Secure Protocols

We assume, w.l.o.g., that both Alice and Bob send only binary messages to each other in every round, and only one party sends a message in one round. Suppose $\Pi$ is a protocol that securely computes a simple function $f=(\rho, \sigma)$. In this section we consider protocols of an arbitrary number of rounds, and so, w.l.o.g., we assume that in $\Pi$ Alice and Bob take turns exchanging single bits and that Alice sends the first message in $\Pi$.

Let $q: \mathcal{X} \rightarrow[0,1]$ be such that, Alice, given an input $x$ chooses 0 as her first message with probability $q(x)$ (and 1 with probability $1-q(x)$ ). We define $\Pi^{(0)}$ to be the protocol, with input spaces $\mathcal{X}^{(0)}=\{x \in \mathcal{X}: q(x)>0\}$ and $\mathcal{Y}$, in which Alice's first message is redefined to be 0 with probability 1 for every input $x \in \mathcal{X}^{(0)}$; otherwise $\Pi^{(0)}$ has identical next message and output functions as $\Pi$. Let $f^{(0)}$ be the function computed by $\Pi^{(0)}$ : i.e., $f^{(0)}(x, y, z)=\operatorname{Pr}_{\Pi^{(0)}}[z \mid x, y]$ for all $x \in \mathcal{X}^{(0)}, y \in \mathcal{Y}, z \in \mathcal{Z} . f^{(1)}$ is defined symmetrically.

Claim 1. There exists a function $\phi: \mathcal{Z} \rightarrow[0,1]$ such that for all $x \in \mathcal{X}, y \in \mathcal{Y}$,

$$
\begin{equation*}
q(x)=\sum_{z \in \mathcal{Z}} \phi(z) f(x, y, z) \tag{3}
\end{equation*}
$$

Further, $f^{(0)}(x, y, z)=\frac{\phi(z)}{q(x)} \cdot f(x, y, z)$, for all $x \in \mathcal{X}^{(0)}, y \in \mathcal{Y}, z \in \mathcal{Z}$.
Proof. We define

$$
\begin{equation*}
\phi(z):=\sum_{\substack{m: m_{1}=0, \\ \text { out }_{I}(m)=z}} \mu(m) \tag{4}
\end{equation*}
$$

where $\mu(m)$ is as defined in (1). Here $m_{1}$ denotes the first bit in the transcript $m$. Note that we have $\phi(z) \in[0,1]$ because for each $z, \mu$ defines a probability distribution over $\mathcal{M}_{z}:=\left\{m\right.$ : out $\left.{ }_{\Pi}(m)=z\right\}$ and $\phi(z)$ sums up the probabilities for a subset of $\mathcal{M}_{z}$.

To see that $\phi$ satisfies the claim, note that $\sum_{z \in \mathcal{Z}} \phi(z) f(x, y, z)=$ $\sum_{m: m_{1}=0} \mu(m) \cdot f\left(x, y\right.$, out $\left._{\Pi}(m)\right)$. Now, from (2), $\mu(m) \cdot f\left(x, y\right.$, out $\left._{\Pi}(m)\right)=$ $\operatorname{Pr}_{\Pi}[m \mid x, y]$. Hence,

$$
\sum_{z \in \mathcal{Z}} \phi(z) f(x, y, z)=\sum_{m: m_{1}=0} \operatorname{Pr}_{\Pi}[m \mid x, y]=q(x) .
$$

Also, for all $x \in \mathcal{X}^{(0)}, y \in \mathcal{Y}, z \in \mathcal{Z}$,

$$
\begin{aligned}
f^{(0)}(x, y, z) & =\sum_{\begin{array}{c}
m: \\
\operatorname{out}_{\Pi}(m)=z
\end{array}} \operatorname{Pr}_{\Pi(0)}[m \mid x, y]=\sum_{\substack{m: m_{1}=0, \operatorname{out}_{\Pi}(m)=z}} \frac{\operatorname{Pr}_{\Pi}[m \mid x, y]}{q(x)}=\sum_{\begin{array}{c}
m: m 1_{1}=0, \\
\text { out }_{\Pi}(m)=z
\end{array}} \mu(m) \frac{f(x, y, z)}{q(x)} \\
& =\frac{f(x, y, z)}{q(x)} \phi(z) .
\end{aligned}
$$

Note that since $\mu(m)$ is the probability of the transcript being $m$ given that the output is out ${ }_{\Pi}(m)$, (4) gives that for every $z \in \mathcal{Z}, \phi(z)=\operatorname{Pr}\left[m_{1}=\right.$ $0 \mid$ out $\left._{\Pi}(m)=z\right]$, i.e., the probability of the first message being 0 , conditioned on the output being $z$. It follows from Claim 1 that a secure protocol is completely and compactly described by the values of $(\phi(z))_{z \in \mathcal{Z}}$ similarly defined in every round.

Remark 2. If $\phi(z)=c$ for all $z \in \mathcal{Z}$, then $q(x)=c$ for all $x \in \mathcal{X}$ (because $\sum_{z \in \mathcal{Z}} f(x, y, z)=1$ for all $\left.(x, y)\right)$. Further, if $c>0, f^{(0)}=f$, and if $c<1$, $f^{(1)}=f$. This corresponds to a protocol in which Alice sends an inconsequential first message, which neither depends on her input, nor influences the output. Hence, if $\Pi$ is round-optimal, it cannot be the case that $\phi(z)=c$ for all $z \in \mathcal{Z}$.

Note that (3) holds for every $y \in \mathcal{Y}$. Hence, we obtain

$$
\begin{equation*}
\sum_{z \in \mathcal{Z}}\left(f(x, y, z)-f\left(x, y^{\prime}, z\right)\right) \phi(z)=0 \quad \forall x \in \mathcal{X}, y, y^{\prime} \in \mathcal{Y} \tag{5}
\end{equation*}
$$

For each $x \in \mathcal{X}$, this gives a system of $|\mathcal{Y}|-1$ equations, by choosing a fixed $y \in \mathcal{Y}$ and all $y^{\prime} \in \mathcal{Y} \backslash\{y\}$ (one may write the equations resulting from other choices of $y, y^{\prime}$ as linear combinations of these $|\mathcal{Y}|-1$ equations).

### 3.2 Binary Output Alphabet

Theorem 4. If $f: \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \rightarrow[0,1]$ is a simple function with $|\mathcal{Z}|=2$, then $f$ is securely computable.
Proof. In fact, we can prove a stronger statement: if $f$ is as above, then $f(x, y, z)$ is either independent of $x$ or independent of $y$ (or both). In that case, clearly $f$ is securely computable by a protocol in which one party computes the output and sends it to the other party.

Since $f$ is simple, we have $f(x, y, z)=\rho(x, z) \cdot \sigma(y, z)$, for all $x \in \mathcal{X}, y \in$ $\mathcal{Y}, z \in \mathcal{Z}$. Since $|\mathcal{Z}|=2$, we let $\mathcal{Z}=\{0,1\}$, and abbreviate $\rho(x, z)$ as $\rho_{z}(x)$ and $\sigma(y, z)$ as $\sigma_{z}(y)$ (for $z \in\{0,1\}$ ). Note that we have the following system of equations:

$$
\begin{equation*}
\rho_{0}(x) \sigma_{0}(y)+\rho_{1}(x) \sigma_{1}(y)=1 \quad \forall(x, y) \in \mathcal{X} \times \mathcal{Y} \tag{6}
\end{equation*}
$$

We consider 3 cases:
Case 1: $\exists x, x^{\prime} \in \mathcal{X}, x \neq x^{\prime}, \rho_{0}(x) \rho_{1}\left(x^{\prime}\right) \neq \rho_{0}\left(x^{\prime}\right) \rho_{1}(x)$. In this case, one can solve the system in (6) to get two values $s_{0}, s_{1}$ such that $\sigma_{0}(y)=s_{0}$ and $\sigma_{1}(y)=s_{1}$ for all $y \in \mathcal{Y}$. (Concretely, $s_{0}=\frac{\rho_{1}\left(x^{\prime}\right)-\rho_{1}(x)}{\Delta}$ and $s_{1}=\frac{\rho_{0}(x)-\rho_{0}\left(x^{\prime}\right)}{\Delta}$, where $\Delta=$ $\rho_{0}(x) \rho_{1}\left(x^{\prime}\right)-\rho_{0}\left(x^{\prime}\right) \rho_{1}(x) \neq 0$.) Hence $f(x, y, z)=\rho(x, z) \cdot s_{z}$, depends only on $x$.

Case 2: $\forall x, x^{\prime} \in \mathcal{X}, \rho_{0}(x) \rho_{1}\left(x^{\prime}\right)=\rho_{0}\left(x^{\prime}\right) \rho_{1}(x)$, and $\exists x \in \mathcal{X}, \rho_{0}(x)=0$. In this case we shall show that, $\forall x^{\prime} \in \mathcal{X}, \rho_{0}\left(x^{\prime}\right)=0$. Hence the function is the constant, deterministic function, with $f(x, y, 0)=0$ for all $(x, y)$.

Let $x$ be such that $\rho_{0}(x)=0$. Since $\rho_{0}(x) \sigma_{0}(y)+\rho_{1}(x) \sigma_{1}(y)=1$ (for any $y \in \mathcal{Y})$, we have $\rho_{1}(x) \neq 0$. Then, since $\forall x^{\prime} \in \mathcal{X}, \rho_{0}(x) \rho_{1}\left(x^{\prime}\right)=\rho_{0}\left(x^{\prime}\right) \rho_{1}(x)$, we have $0=\rho_{0}\left(x^{\prime}\right) \rho_{1}(x)$ which implies $\rho_{0}\left(x^{\prime}\right)=0$, as claimed.

Case 3: $\forall x, x^{\prime} \in \mathcal{X}, \rho_{0}(x) \rho_{1}\left(x^{\prime}\right)=\rho_{0}\left(x^{\prime}\right) \rho_{1}(x)$, and $\forall x \in \mathcal{X}, \rho_{0}(x) \neq 0$. In this case, $\forall x, x^{\prime} \in \mathcal{X}, \frac{\rho_{1}(x)}{\rho_{0}(x)}=\frac{\rho_{1}\left(x^{\prime}\right)}{\rho_{0}\left(x^{\prime}\right)}=\theta$, say. Then, from (6), we have that $\forall x \in \mathcal{X}, y \in \mathcal{Y}, \rho_{0}(x)=\frac{1}{\sigma_{0}(y)+\theta \sigma_{1}(y)}$ and $\rho_{1}(x)=\frac{\theta}{\sigma_{0}(y)+\theta \sigma_{1}(y)}$. Since the RHS in these two expressions do not depend on $x$, there are constants $r_{0}, r_{1}$ such that for all $x \in \mathcal{X}, \rho_{0}(x)=r_{0}$ and $\rho_{1}(x)=r_{1}$. Hence $f(x, y, z)=r_{z} \cdot \sigma(y, z)$, depends only on $y$.

### 3.3 Ternary Output Alphabet

To prove Theorem 1, we can focus on kernel functions, or simple symmetric functions. Let $\mathcal{Z}=\left\{z_{1}, z_{2}, z_{3}\right\}$. For a given symmetric function $f$ with $\mathcal{Z}$ as its output alphabet, we define two binary functions $\hat{f}_{i}$ and $f_{i}$, for any $i \in[3]$, as follows:

1. Output alphabet of $\hat{f}_{i}$ is $\left\{z_{i}, z_{*}\right\}$. For every $x \in \mathcal{X}, y \in \mathcal{Y}$, we define $\hat{f}_{i}\left(x, y, z_{i}\right):=f\left(x, y, z_{i}\right)$ and $\hat{f}_{i}\left(x, y, z_{*}\right):=\sum_{j \neq i} f\left(x, y, z_{j}\right)$.
2. Output alphabet of $f_{i}$ is $\mathcal{Z} \backslash\left\{z_{i}\right\}$. For every $x \in \mathcal{X}, y \in \mathcal{Y}$ for which $f\left(x, y, z_{i}\right)<1$, we define, $f_{i}\left(x, y, z_{j}\right):=f\left(x, y, z_{j}\right) /\left(1-f\left(x, y, z_{i}\right)\right)$, for $j \in[3] \backslash\{i\}$. If $f\left(x, y, z_{i}\right)=1$ for some $x \in \mathcal{X}, y \in \mathcal{Y}$, we leave $f_{i}\left(x, y, z_{i}\right)$ undefined.

Note that if for some $\tilde{x} \in \mathcal{X}, f\left(\tilde{x}, y, z_{i}\right)=1, \forall y \in \mathcal{Y}$, then we need not define $f_{i}$ for this particular $\tilde{x}$, because $f\left(\tilde{x}, y, z_{j}\right)=0, \forall j \neq i$ and for all $y \in \mathcal{Y}$. Similarly, if for some $\tilde{y} \in \mathcal{Y}, f\left(x, \tilde{y}, z_{i}\right)=1, \forall x \in \mathcal{X}$, then we need not define $f_{i}$ for this particular $\tilde{y}$, because $f\left(x, \tilde{y}, z_{j}\right)=0, \forall j \neq i$ and for all $x \in \mathcal{X}$. In the following, when we say that $f_{i}$ is securely computable, it must be defined for all inputs.

Theorem 5. Suppose $f$ is simple and in normal form. Then, $f$ is securely computable if and only if there exists an $i \in[3]$ such that both $\hat{f}_{i}$ and $f_{i}$ are simple.

Remark 3. Since $\hat{f}_{i}$ and $f_{i}$ are binary functions, being simple implies that they are functions of only one party's input (see proof of Theorem 4). Theorem 1 follows by considering the different possibilities of which party's input they can each depend on.

Remark 4. In the case of functions with a binary output alphabet, we proved in Subsect. 3.2 that if a function is simple, it is securely computable. However, Theorem 5 lets us show that this is not true in general (see in Sect.5.1).

Proof of Theorem 5. If $|\mathcal{X}|=1$ or $|\mathcal{Y}|=1$, then the theorem is trivially true. So, in the following we assume that $|\mathcal{X}|,|\mathcal{Y}| \geq 2$. First we show the only if $(\Rightarrow)$ part, and then the if $(\Leftarrow)$ part.
$\Rightarrow$ : Suppose $f$ is securely computable. Since $f$ is simple, we can write $f(x, y, z)=\rho(x, z) \sigma(y, z)$. Fix a round-optimal secure protocol $\Pi$ for $f$. Below, we assume that Alice sends the first message in $\Pi$ (the case when Bob sends the
first message being symmetric). Let $\phi$ be as in Claim 1. Since $\Pi$ is round-optimal, as discussed in Remark 2,

$$
\begin{equation*}
\exists i, j \in[3] \text { s.t. } \phi\left(z_{i}\right) \neq \phi\left(z_{j}\right) . \tag{7}
\end{equation*}
$$

For $i \in[3]$ and $x \in \mathcal{X}, y, y^{\prime} \in \mathcal{Y}$, we define

$$
\begin{equation*}
\nabla f_{i}^{x, y, y^{\prime}}:=f\left(x, y, z_{i}\right)-f\left(x, y^{\prime}, z_{i}\right) \tag{8}
\end{equation*}
$$

It follows from (5) and (8) that, $\forall x \in \mathcal{X}, y, y^{\prime} \in \mathcal{Y}$ :

$$
\begin{equation*}
\nabla f_{1}^{x, y, y^{\prime}} \phi\left(z_{1}\right)+\nabla f_{2}^{x, y, y^{\prime}} \phi\left(z_{2}\right)+\nabla f_{3}^{x, y, y^{\prime}} \phi\left(z_{3}\right)=0 \tag{9}
\end{equation*}
$$

Since $\sum_{z \in \mathcal{Z}} f(x, y, z)=1$ holds for every $x \in \mathcal{X}, y \in \mathcal{Y}$, we have that $\sum_{i=1}^{3} \nabla f_{i}^{x, y, y^{\prime}}=0$, for every $x \in \mathcal{X}$ and $y, y^{\prime} \in \mathcal{Y}$. Using this to replace $\nabla f_{3}^{x, y, y^{\prime}}$ in (9), we can write, $\forall x \in \mathcal{X}, y, y^{\prime} \in \mathcal{Y}$ :

$$
\begin{equation*}
\nabla f_{1}^{x, y, y^{\prime}}\left(\phi\left(z_{1}\right)-\phi\left(z_{3}\right)\right)+\nabla f_{2}^{x, y, y^{\prime}}\left(\phi\left(z_{2}\right)-\phi\left(z_{3}\right)\right)=0 \tag{10}
\end{equation*}
$$

(and two similar equations, replacing $\nabla f_{1}^{x, y, y^{\prime}}$ and $\nabla f_{2}^{x, y, y^{\prime}}$, respectively).
We define a function type ${ }_{f}: \mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{T}$, which classifies $\left(x, y, y^{\prime}\right)$ into one of 5 possible types in the set $\mathbb{T}=\left\{\mathrm{T}_{1}, \mathrm{~T}_{2: 1}, \mathrm{~T}_{2: 2}, \mathrm{~T}_{2: 3}, \mathrm{~T}_{3}\right\}$, depending on for which $i, \nabla f_{i}^{x, y, y^{\prime}}=0$ :

1. $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{1}$ if $\nabla f_{1}^{x, y, y^{\prime}}=\nabla f_{2}^{x, y, y^{\prime}}=\nabla f_{3}^{x, y, y^{\prime}}=0$.
2. $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{2: i}$ if $\nabla f_{i}^{x, y, y^{\prime}}=0$ and $\forall j \in[3] \backslash\{i\}, \nabla f_{j}^{x, y, y^{\prime}} \neq 0$.
3. $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{3}$ if $\forall i \in[3], \nabla f_{i}^{x, y, y^{\prime}} \neq 0$.

Note that since $\sum_{i=1}^{3} \nabla f_{i}^{x, y, y^{\prime}}=0$, it cannot be the case that exactly one of $\nabla f_{i}^{x, y, y^{\prime}} \neq 0$.

We define the type of $f$ itself as the set:

$$
\begin{equation*}
T(f):=\left\{\tau \in \mathbb{T}: \exists\left(x, y, y^{\prime}\right) \text { s.t. type }\left(x, y, y^{\prime}\right)=\tau\right\} \tag{11}
\end{equation*}
$$

We prove several claims regarding $T(f)$ before showing that $\hat{f}_{i}$ and $f_{i}$ are simple. In proving these claims, we shall use the fact that $f$ is simple, is in normal-form, and $|\mathcal{X}|,|\mathcal{Y}|>1$.

Claim 2. $\mathrm{T}_{3} \notin T(f)$.
Proof. For the sake of contradiction, suppose there exists $\left(x, y, y^{\prime}\right)$ such that $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{3}$. Consider any $x^{\prime} \in \mathcal{X} \backslash\{x\}$. From (10), we have

$$
\begin{aligned}
\nabla f_{1}^{x, y, y^{\prime}}\left(\phi\left(z_{1}\right)-\phi\left(z_{3}\right)\right)+\nabla f_{2}^{x, y, y^{\prime}}\left(\phi\left(z_{2}\right)-\phi\left(z_{3}\right)\right) & =0 \\
\nabla f_{1}^{x^{\prime}, y, y^{\prime}}\left(\phi\left(z_{1}\right)-\phi\left(z_{3}\right)\right)+\nabla f_{2}^{x^{\prime}, y, y^{\prime}}\left(\phi\left(z_{2}\right)-\phi\left(z_{3}\right)\right) & =0
\end{aligned}
$$

We shall show that $\nabla f_{1}^{x, y, y^{\prime}} \nabla f_{2}^{x^{\prime}, y, y^{\prime}} \neq \nabla f_{2}^{x, y, y^{\prime}} \nabla f_{1}^{x^{\prime}, y, y^{\prime}}$. Then, the above system can be uniquely solved for $\phi\left(z_{1}\right)-\phi\left(z_{3}\right)$ and $\phi\left(z_{2}\right)-\phi\left(z_{3}\right)$, to yield 0 as the solution in each case. That is, $\phi\left(z_{1}\right)=\phi\left(z_{2}\right)=\phi\left(z_{3}\right)$, contradicting (7).

To complete the proof, we argue that $\nabla f_{1}^{x, y, y^{\prime}} \nabla f_{2}^{x^{\prime}, y, y^{\prime}} \neq \nabla f_{2}^{x, y, y^{\prime}} \nabla f_{1}^{x^{\prime}, y, y^{\prime}}$. Suppose not. Then, $\frac{\nabla f_{1}^{x^{\prime}, y, y^{\prime}}}{\nabla f_{1}^{x, y, y^{\prime}}}=\frac{\nabla f_{2}^{x^{\prime}, y, y^{\prime}}}{\nabla f_{2}^{x, y, y^{\prime}}}=\theta$, say (the denominators being nonzero, since type $\left.{ }_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{3}\right)$. Then, $\frac{\nabla f_{3}^{x^{\prime}, y, y^{\prime}}}{\nabla f_{3}^{x, y, y^{\prime}}}=\frac{-\nabla f_{1}^{x^{\prime}, y, y^{\prime}}-\nabla f_{2}^{x^{\prime}, y, y^{\prime}}}{-\nabla f_{1}^{x, y, y^{\prime}}-\nabla f_{2}^{x, y, y^{\prime}}}=\theta$. Invoking the simplicity of $f$, we get that $\forall j \in[3], \frac{\rho\left(x, z_{j}\right)}{\rho\left(x^{\prime}, z_{j}\right)}=\theta$. However, since for any $y$ we have $\sum_{j} \rho\left(x, z_{j}\right) \sigma\left(y, z_{j}\right)=\sum_{j} \rho\left(x^{\prime}, z_{j}\right) \sigma\left(y, z_{j}\right)$, we get $\theta=1$. Then, $x \equiv x^{\prime}$, contradicting the normal form of $f$. This completes the proof.

Claim 3. There can be at most one $i \in[3]$ such that $\mathrm{T}_{2: i} \in T(f)$.
Proof. For the sake of contradiction, suppose $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{2: j}$, and $\operatorname{type}_{f}\left(\tilde{x}, \tilde{y}, \tilde{y}^{\prime}\right)=\mathrm{T}_{2: k}$, for $j \neq k$. We consider the case $j=1, k=2$, as the other cases are symmetric. Now, type ${ }_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{2: 1}$, implies that $\nabla f_{i}^{x, y, y^{\prime}}=0$ only for $i=1$ and hence, by (10), we have $\nabla f_{2}^{x, y, y^{\prime}}\left(\phi\left(z_{2}\right)-\phi\left(z_{3}\right)\right)=0$, and so $\phi\left(z_{2}\right)=\phi\left(z_{3}\right)$. Similarly, $\operatorname{type}_{f}\left(\tilde{x}, \tilde{y}, \tilde{y}^{\prime}\right)=\mathrm{T}_{2: 2}$ implies that $\phi\left(z_{1}\right)=\phi\left(z_{3}\right)$. Thus we have $\phi\left(z_{1}\right)=\phi\left(z_{2}\right)=\phi\left(z_{3}\right)$, contradicting (7).

Claim 4. If from some $x \in \mathcal{X}$ and distinct $y, y^{\prime} \in \mathcal{Y}$, $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{1}$, then there exists $x^{\prime} \in \mathcal{X}$ such that type ${ }_{f}\left(x^{\prime}, y, y^{\prime}\right) \neq \mathrm{T}_{1}$.

Proof. Since $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{1}$, we have that for all $j \in[3], f\left(x, y, z_{j}\right)=$ $f\left(x, y^{\prime}, z_{j}\right)$. Since $f$ is in normal form, $y \not \equiv y^{\prime}$, and hence there exists $x^{\prime} \in \mathcal{X}$ such that for some $j \in[3], f\left(x^{\prime}, y, z_{j}\right) \neq f\left(x^{\prime}, y^{\prime}, z_{j}\right)$. Hence type ${ }_{f}\left(x^{\prime}, y, y^{\prime}\right) \neq \mathrm{T}_{1}$.

Claim 5. Suppose $T(f)=\left\{\mathbf{T}_{1}, \mathrm{~T}_{2: i}\right\}$ for some $i \in[3]$. Then, if for some $x \in \mathcal{X}$, and distinct $y, y^{\prime} \in \mathcal{Y}$, $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{1}$, then for all $\tilde{y} \in \mathcal{Y}, f\left(x, \tilde{y}, z_{i}\right)=1$.

Proof. By Claim 4, we have $x^{\prime}$ such that type ${ }_{f}\left(x^{\prime}, y, y^{\prime}\right) \neq \mathrm{T}_{1}$; since $T(f)=$ $\left\{\mathrm{T}_{1}, \mathrm{~T}_{2: i}\right\}$, we have type ${ }_{f}\left(x^{\prime}, y, y^{\prime}\right)=\mathrm{T}_{2: i}$. Then, for both values of $j \neq i$ in [3], we have $f\left(x^{\prime}, y, z_{j}\right) \neq f\left(x^{\prime}, y^{\prime}, z_{j}\right)$.

Invoking the simplicity of $f$, we have that for $j \neq i, \rho\left(x^{\prime}, z_{j}\right)\left(\sigma\left(y, z_{j}\right)-\right.$ $\left.\sigma\left(y^{\prime}, z_{j}\right)\right) \neq 0$, but $\rho\left(x, z_{j}\right)\left(\sigma\left(y, z_{j}\right)-\sigma\left(y^{\prime}, z_{j}\right)\right)=0$. Hence $\rho\left(x, z_{j}\right)=0$ for $j \neq i$. That is, for all $\tilde{y} \in \mathcal{Y}, f\left(x, \tilde{y}, z_{j}\right)=0$ for $j \neq i$ and hence $f\left(x, \tilde{y}, z_{i}\right)=1$.

From the above claims, we have two possibilities for $T(f)$ : either $\left\{\mathrm{T}_{2: i}\right\}$ or $\left\{\mathbf{T}_{1}, \mathbf{T}_{2: i}\right\}$ for some $i \in[3]$. Then we shall prove that $\hat{f}_{i}$ and $f_{i}$ are simple. Note that both these functions are binary functions. For our case where $f$ has a minimal-round protocol with Alice sending the first bit, and with $|\mathcal{X}|,|\mathcal{Y}|>1$, we show that this means that $\hat{f}_{i}$ is a function of only Alice's input, and $f_{i}$ is a function of only Bob's input.

Claim 6. If $T(f) \subseteq\left\{\mathrm{T}_{1}, \mathrm{~T}_{2: i}\right\}$, then $\hat{f}_{i}$ is simple.

Proof. We have that for all $\left(x, y, y^{\prime}\right), \nabla f_{i}^{x, y, y^{\prime}}=0$, i.e., $\rho\left(x, z_{i}\right)\left(\sigma\left(y, z_{i}\right)-\right.$ $\left.\sigma\left(y^{\prime}, z_{i}\right)\right)=0$. Now, if $\rho\left(x, z_{i}\right)=0$ for all $x$, then $\hat{f}_{i}$ is a constant function with $\hat{f}_{i}\left(x, y, z_{i}\right)=0$ and $\hat{f}_{i}\left(x, y, z_{*}\right)=1$. Otherwise, there is some $x \in \mathcal{X}$ such that $\rho\left(x, z_{i}\right) \neq 0$. Hence for all $y, y^{\prime}$ we have $\sigma\left(y, z_{i}\right)=\sigma\left(y^{\prime}, z_{i}\right)=s$, say. Then $\hat{f}_{i}\left(x, y, z_{i}\right)=\rho\left(x, z_{i}\right) \cdot s$, which is independent of $y$. Thus, in either case, $\hat{f}_{i}$ is simple.

Claim 7. If $T(f) \subseteq\left\{\mathrm{T}_{1}, \mathrm{~T}_{2: i}\right\}$, then $f_{i}(x, y, z)$ is simple.
Proof. Recall that $f_{i}$ has input spaces $\mathcal{X}_{i}$ and $\mathcal{Y}$, where $\mathcal{X}_{i}=\{x \in \mathcal{X}: \exists y \in$ $\mathcal{Y}$ s.t. $\left.f\left(x, y, z_{i}\right)<1\right\}$. We shall prove that for all $x, x^{\prime} \in \mathcal{X}_{i}$ and $y \in \mathcal{Y}$, $f_{i}\left(x, y, z_{j}\right)=f_{i}\left(x^{\prime}, y, z_{j}\right)$ for $j \neq i$.

By Claim 5, for all $x \in \mathcal{X}_{i}$ and distinct $y, y^{\prime} \in \mathcal{Y}, \operatorname{type}_{f}\left(x, y, y^{\prime}\right) \neq \mathrm{T}_{1}$, and hence type ${ }_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{2: i}$. Therefore, for all $x \in \mathcal{X}_{i}$ and $j \neq i, \rho\left(x, z_{j}\right) \neq 0$.

Let $\{j, \bar{\jmath}\}=[3] \backslash\{i\}$. Now,

$$
f_{i}\left(x, y, z_{j}\right)=\frac{f\left(x, y, z_{j}\right)}{1-f\left(x, y, z_{i}\right)}=\frac{\rho\left(x, z_{j}\right) \sigma\left(y, z_{j}\right)}{\rho\left(x, z_{j}\right) \sigma\left(y, z_{j}\right)+\rho\left(x, z_{\bar{\jmath}}\right) \sigma\left(y, z_{\bar{\jmath}}\right)}=\frac{\sigma\left(y, z_{j}\right)}{\sigma\left(y, z_{j}\right)+\gamma(x) \sigma\left(y, z_{\bar{\jmath}}\right)},
$$

where $\gamma(x):=\frac{\rho\left(x, z_{\bar{j}}\right)}{\rho\left(x, z_{j}\right)}$. Here we have used the fact that $\rho\left(x, z_{j}\right) \neq 0$ for all $x \in \mathcal{X}_{i}$. Thus to prove the claim, it is enough to show that $\gamma(x)=\gamma\left(x^{\prime}\right)$ for all $x, x^{\prime} \in \mathcal{X}_{i}$.

For any $x \in \mathcal{X}_{i}$, as mentioned above, for any distinct $y, y^{\prime} \in \mathcal{Y}$, we have $\operatorname{type}_{f}\left(x, y, y^{\prime}\right)=\mathrm{T}_{2: i}$, which implies that $\nabla f_{j}^{x, y, y^{\prime}}+\nabla f_{\bar{\jmath}}^{x, y, y^{\prime}}=-\nabla f_{i}^{x, y, y^{\prime}}=0$. That is,

$$
\rho\left(x, z_{j}\right)\left(\sigma\left(y, z_{j}\right)-\sigma\left(y^{\prime}, z_{j}\right)\right)+\rho\left(x, z_{\bar{\jmath}}\right)\left(\sigma\left(y, z_{\bar{\jmath}}\right)-\sigma\left(y^{\prime}, z_{\bar{\jmath}}\right)\right)=0
$$

Writing the above equation for any $x, x^{\prime} \in \mathcal{X}_{i}$, if $\rho\left(x, z_{j}\right) \rho\left(x^{\prime}, z_{\bar{\jmath}}\right) \neq$ $\rho\left(x, z_{\bar{\jmath}}\right) \rho\left(x^{\prime}, z_{j}\right)$, we will be able to solve that $\left(\sigma\left(y, z_{j}\right)-\sigma\left(y^{\prime}, z_{j}\right)\right)=\left(\sigma\left(y, z_{\bar{\jmath}}\right)-\right.$ $\left.\sigma\left(y^{\prime}, z_{\bar{\jmath}}\right)\right)=0$. But this implies that $\nabla f_{j}^{x, y, y^{\prime}}=\nabla f_{\bar{\jmath}}^{x, y, y^{\prime}}=0$ (for any $x \in \mathcal{X}_{i}$ ), contradicting the fact that type ${ }_{f}\left(x, y, y^{\prime}\right) \neq \mathrm{T}_{1}$ for all $x \in \mathcal{X}_{i}$ and distinct $y, y^{\prime} \in \mathcal{Y}$. Thus we should have $\rho\left(x, z_{j}\right) \rho\left(x^{\prime}, z_{\bar{\jmath}}\right)=\rho\left(x, z_{\bar{\jmath}}\right) \rho\left(x^{\prime}, z_{j}\right)$, or (dividing by $\left.\rho\left(x, z_{j}\right) \cdot \rho\left(x^{\prime}, z_{j}\right) \neq 0\right), \gamma(x)=\gamma\left(x^{\prime}\right)$.

Taken together, the above claims prove the only if $(\Rightarrow)$ part of Theorem 5.
$\Leftarrow$ : Let $i \in[3]$ be such that $\hat{f}_{i}$ and $f_{i}$ are simple, and therefore, securely computable. A 2-round secure protocol for $f$ is as follows. If both $\hat{f}_{i}$ and $f_{i}$ are independent of $x$ or $y$ (or both), then clearly $f$ is securely computable by a protocol in which one party computes the output and sends it to the other party. The only interesting case is when $\hat{f}_{i}$ is independent of $y$ and $f_{i}$ is independent of $x$ (the other case when $\hat{f}_{i}$ is independent of $x$ and $f_{i}$ is independent of $y$ being symmetric): Alice picks $j \in\{i, *\}$ with probability $\hat{f}_{i}\left(x, y_{1}, z_{j}\right)$ and sends $j$ to Bob. If $j=i$, both Alice and Bob output $z_{i}$ with probability 1 ; otherwise, Bob picks $k \in[3] \backslash\{i\}$ with probability $f_{i}\left(x_{1}, y, z_{k}\right)$ and sends $k$ to Alice. Now both Alice and Bob output $z_{k}$ with probability 1. It is clear from the definitions of $\hat{f}_{i}$ and $f_{i}$ that the output from this protocol is correctly distributed. It is easy to see that this is a unique-transcript protocol, and therefore, is perfectly private.

## 4 Functions with 2-Round Secure Protocols

We prove the following theorem which, when applied to the kernel of a given function, implies Theorem 2.

Theorem 6. A simple symmetric function $p_{Z \mid X Y}$ with $X, Y, Z$ over alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, has a two-round protocol with Alice making the first move iff there exists a surjective map $g: \mathcal{Z} \rightarrow \mathcal{W}$ to some set $\mathcal{W}$ and probability distributions $p_{W \mid X}$ and $p_{Z \mid W Y}$ where $W$ is over the alphabet $\mathcal{W}$, such that $p_{Z \mid W Y}(z \mid w, y)=0$ if $w \neq g(z)$, and for all $x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z}$,

$$
p_{Z \mid X Y}(z \mid x, y)=p_{W \mid X}(g(z) \mid x) \cdot p_{Z \mid W Y}(z \mid g(z), y)
$$

In that case, $p_{Z \mid X Y}$ has a unique-transcript secure protocol in which Alice sends $w$ sampled according to $p_{W \mid X=x}$ and Bob sends back $z$ according to $p_{Z \mid W=w, Y=y}$.

Proof. It is easy to see that if $g, p_{W \mid X}, p_{Z \mid W Y}$ as in the statement exist, then the protocol described is indeed a unique-transcript protocol for $p_{Z \mid X Y}$ : its output is distributed correctly, and since $p_{Z \mid W Y}(z \mid w, y)=0$ if $w \neq g(z)$, the only transcript resulting in the output $z$ is of the form $(g(z), z)$. A unique-transcript protocol is always a secure protocol since the transcript can be simulated from the output by a (deterministic) simulator.

We prove the other direction below. Suppose we are given a two-round protocol $\Pi_{0}$ for $p_{Z \mid X Y}$, with the two messages denoted by $a$ and $b$. Then, we can construct a secure protocol $\Pi$ in which Bob computes the second message $b$ as before, but sends out $z=\operatorname{out}_{\Pi_{0}}(a, b)$, and both Alice and Bob output $z$ : clearly $\Pi$ has the same output as $\Pi_{0}$ and is also secure since the transcript of $\Pi$ can be simulated from a (simulated) transcript of $\Pi_{0}$ by applying a deterministic function to it.
$\Pi$ is defined by probability distributions $\operatorname{Pr}_{\Pi}[a \mid x]$ and $\operatorname{Pr}_{\Pi}[z \mid a, y]$. For convenience, we define $\alpha(a, x):=\operatorname{Pr}_{\Pi}[a \mid x]$ and $\beta(a, z, y):=\operatorname{Pr}_{\Pi}[z \mid a, y]$. Also, since $p_{Z \mid X Y}$ is simple, let us write $p_{Z \mid X Y}(z \mid x, y)=\rho(x, z) \cdot \sigma(y, z)$.

Before proceeding further, note that for a transcript $m=(a, z)$, from (2), we have

$$
\operatorname{Pr}_{\Pi}[m \mid x, y]=\mu(m) \rho(x, z) \sigma(y, z)=\alpha(a, x) \beta(a, z, y)
$$

If $\operatorname{Pr}_{\Pi}[m \mid x, y]>0$ for some $x, y$, then by considering the above equality for $(x, y)$ as well as $\left(x^{\prime}, y\right)$, and dividing the latter by the former (which is non-zero), we get that for all $x^{\prime} \in \mathcal{X}$,

$$
\begin{equation*}
\frac{\rho\left(x^{\prime}, z\right)}{\rho(x, z)}=\frac{\alpha\left(a, x^{\prime}\right)}{\alpha(a, x)} \tag{12}
\end{equation*}
$$

We define an equivalence relation $\equiv$ over $\mathcal{Z}$ as follows:

$$
z_{1} \equiv z_{2} \text { if } \exists c>0, \forall x \in \mathcal{X}, \rho\left(x, z_{1}\right)=c \rho\left(x, z_{2}\right)
$$

We let $\mathcal{W}$ be the set of equivalence classes of $\equiv$, and define $g: \mathcal{Z} \rightarrow \mathcal{W}$ which maps $z$ to its equivalence class. Thus, $z_{1} \equiv z_{2}$ iff $g\left(z_{1}\right)=g\left(z_{2}\right)$. We also define a
function $h$ that maps the first message $a$ in a transcript to an element in $\mathcal{W}$, as follows:

$$
h(a)=g(z) \text { if } \exists x, y \text { s.t. } \operatorname{Pr}_{\Pi}[a, z \mid x, y]>0 .
$$

For $h$ to be well-defined, we need that each $a$ has a unique value $h(a)$ that satisfies the above condition. Suppose $z_{1}, z_{2} \in \mathcal{Z}$ are such that $\operatorname{Pr}_{\Pi}\left[a, z_{1} \mid x_{1}, y_{1}\right]>0$ and $\operatorname{Pr}_{\Pi}\left[a, z_{2} \mid x_{2}, y_{2}\right]>0$ (from some $\left.\left(x_{1}, y_{1}\right),\left(x_{2}, y_{2}\right) \in \mathcal{X} \times \mathcal{Y}\right)$. By applying (12) to these, we get that for all $x^{\prime} \in \mathcal{X}$

$$
\frac{\rho\left(x^{\prime}, z_{1}\right)}{\rho\left(x_{1}, z_{1}\right)}=\frac{\alpha\left(a, x^{\prime}\right)}{\alpha\left(a, x_{1}\right)} \quad \text { and } \quad \frac{\rho\left(x^{\prime}, z_{2}\right)}{\rho\left(x_{2}, z_{2}\right)}=\frac{\alpha\left(a, x^{\prime}\right)}{\alpha\left(a, x_{2}\right)} .
$$

Hence $\forall x^{\prime} \in \mathcal{X}, \rho\left(x^{\prime}, z_{2}\right)=c \rho\left(x^{\prime}, z_{1}\right)$, where $c:=\frac{\rho\left(x_{2}, z_{2}\right) \alpha\left(a, x_{1}\right)}{\alpha\left(a, x_{2}\right) \rho\left(x_{1}, z_{1}\right)}$. All the factors in $c$ are positive (as they appear in $\operatorname{Pr}_{\Pi}\left[a, z_{1} \mid x_{1}, y_{1}\right] \cdot \operatorname{Pr}_{\Pi}\left[a, z_{2} \mid x_{2}, y_{2}\right]>0$ ), and also independent of $x^{\prime}$. Thus $z_{1} \equiv z_{2}$ and $g\left(z_{1}\right)=g\left(z_{2}\right)$, making $h(a)$ well defined.
$p_{W \mid X}$ is defined as follows: given $x$, sample $a$ as in $\Pi$, and output $w=h(a)$. That is, $p_{W \mid X}(w \mid x)=\sum_{a: h(a)=w} \alpha(a, x)$. Finally, we define $p_{Z \mid W Y}$ as follows: given $w$, we argue that we can reverse sample $a$ without access to $x$ (so that $w=h(a))$, and then use the protocol $\Pi$ to sample $z$ from $(a, y)$. That is, we define a distribution over $a$ given $w$, by the probability

$$
\eta(a, w):= \begin{cases}\frac{\alpha(a, x)}{\sum_{a^{\prime}: h\left(a^{\prime}\right)=w}^{\alpha\left(a^{\prime}, x\right)}} & \text { if } h(a)=w \\ 0 & \text { otherwise }\end{cases}
$$

where any $x \in \mathcal{X}$ such that $\sum_{a^{\prime}: h\left(a^{\prime}\right)=w} \alpha\left(a^{\prime}, x\right)>0$ is used. This is well defined because, by (12), switching from $x$ to $x^{\prime}$ amounts to multiplying both the numerator and the denominator by the same factor (namely, $\frac{\rho\left(x^{\prime}, z\right)}{\rho(x, z)}$ for any $z \in g^{-1}(w)$ ). Then, $p_{Z \mid W Y}(z \mid w, y)=\sum_{a} \eta(a, w) \beta(a, z, y)$. We verify that, when $w=g(z)$,

$$
\begin{aligned}
p_{W \mid X}(w \mid x) \cdot p_{Z \mid W Y}(z \mid w, y) & =\left(\sum_{a: h(a)=w} \alpha(a, x)\right)\left(\sum_{a} \eta(a, w) \beta(a, z, y)\right) \\
& =\sum_{a} \alpha(a, x) \beta(a, z, y)=\operatorname{Pr}_{\Pi}[z \mid x, y]=p_{Z \mid X Y}(z \mid x, y)
\end{aligned}
$$

## 5 Complexity of Randomized Functions

We point out a couple of complexity aspects in which randomized functions differ from deterministic functions. This also points to the difficulty in characterization of securely computable functions in the case of randomized functions.

### 5.1 Smaller Simple Functions Which Are Not Securely Computable

It is well-known that simple functions are not all securely computable, even for deterministic functions, with a first example given by Beaver [Bea89], with an output alphabet of size 5 . This turns out to be the smallest output alphabet for a simple deterministic function that is not securely realizable.

But for randomized functions, we see a higher level of complexity arising even with an output size of 3 . In Fig. 1 we show an example of a simple function with ternary output alphabet that is not securely computable, i.e., it does not satisfy the characterization given in Theorem 5 .

|  | $y_{1}$ |  | $y_{2}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | $z_{1}$ | $z_{2} \quad z_{3}$ | $z_{1}$ | $z_{2}$ | $z_{3}$ |
| $z_{1}$ | 2/9 |  | 5/18 |  |  |
| $x_{1} z_{2}$ |  | 4/9 |  | 2/9 |  |
| $z_{3}$ |  | 1/3 |  |  | 1/2 |
| $z_{1}$ | 1/3 |  | 5/12 |  |  |
| $x_{2} z_{2}$ |  | 5/12 |  | 5/24 |  |
| $z_{3}$ |  | 1/4 |  |  | 3/8 |

Fig. 1. This function is not securely computable as it does not satisfy the condition from Theorem 5. However, it is simple, with the functions $\rho$ and $\sigma$ given by \begin{tabular}{|l|lll|}
\hline \& $z_{1}$ \& $z_{2}$ \& $z_{3}$ <br>
\hline$x_{1}$ \& $1 / 3$ \& $2 / 3$ \& 1 <br>
$x_{2}$ \& $1 / 2$ \& $5 / 8$ \& $3 / 4$ <br>
\hline

 and 

\hline \& $z_{1}$ \& $z_{2}$ \& $z_{3}$ <br>
\hline$y_{1}$ \& $2 / 3$ \& $2 / 3$ \& $1 / 3$ <br>
$y_{2}$ \& $5 / 6$ \& $1 / 3$ \& $1 / 2$ <br>
\cline { 2 - 4 }
\end{tabular} .

### 5.2 Limits of Unique-Transcript Protocols

It follows from Sect. 3 that all securely computable randomized functions with a ternary output kernel can be computed using unique-transcript protocols. Also, it follows from Sect. 4 that all randomized functions securely computable by two-round protocols are in fact securely computable using two-round uniquetranscript protocols. Also, the characterization by Kushilevitz [Kus89] showed that all securely computable deterministic functions have unique-transcript secure protocols. Thus one may reasonably suspect that all securely computable functions have unique-transcript secure protocols.

However, we show that in some sense, the above results give the limits of unique-transcript protocols: If we go just beyond the above conditions - namely, ternary output, 2-round computable, deterministic - then we can indeed find securely computable functions that do not have any unique-transcript secure protocol. In Appendix B, we demonstrate a simple randomized function with an output alphabet of size 4 , securely computable by a 3 -round protocol, such that it has no unique-transcript secure protocols.

Acknowledgments. Deepesh Data's research was supported in part by a Microsoft Research India Ph.D. Fellowship. Manoj Prabhakaran wishes to thank Hemanta Maji and Amit Sahai for an earlier collaboration on the same problem. The observation in Sect. 5.2 was made (using a different example) during that work as well.

## A Security for Simple Symmetric Functions

In this section we prove Lemma 1 .
Proof of Lemma 1. Firstly, by perfect correctness, for any $x, y$, we require that with probability one, $\operatorname{out}_{\Pi}^{A}(m, x)=\operatorname{out}_{\Pi}^{B}(m, y)$, where $m$ is produced by $\Pi$ on input $(x, y)$. Suppose there is an $x_{0} \in \mathcal{X}, y_{0} \in \mathcal{Y}$ such that $\operatorname{Pr}_{\Pi}\left[m \mid x_{0}, y_{0}\right]>0$.

Since we can write $\operatorname{Pr}_{\Pi}[m \mid x, y]=\alpha(m, x) \beta(m, y)$, we have $\alpha\left(m, x_{0}\right)$ $\beta\left(m, y_{0}\right)>0$. Also, for all $x$ such that $\alpha(m, x)>0$, we have a positive probability of $\Pi$ producing $m$ on input ( $x, y_{0}$ ) and hence we must have out ${ }_{\Pi}^{A}(m, x)=$ out $_{\Pi}^{B}\left(m, y_{0}\right)$ with probability 1 (which requires them to be deterministic). Similarly, for all $y$ such that $\beta(m, y)>0$, we have $\operatorname{out}_{I I}^{B}(m, y)=\operatorname{out}_{\Pi}^{A}\left(m, x_{0}\right)$ with probability 1. Letting $\operatorname{out}_{\Pi}(m):=\operatorname{out}_{\Pi}^{A}\left(m, x_{0}\right)=\operatorname{out}_{\Pi}^{B}\left(m, y_{0}\right)$ (which must be deterministic), we have that out $A_{\Pi}^{A}(m, x)=\operatorname{out}_{\Pi}^{B}(m, y)=\operatorname{out}_{\Pi}(m)$ if $\operatorname{Pr}_{\Pi}[m \mid x, y]>0$.

Now we prove the second part. Note that since $\Pi$ computes $p_{Z \mid X Y}$, we have that for all $x, y, z, \operatorname{Pr}_{\Pi}[z \mid x, y]=p_{Z \mid X Y}(z \mid x, y)$. Consider any $x, y, z$ such that $p_{Z \mid X Y}(z \mid x, y)>0$. For all $m$ such that out ${ }_{\Pi}(m) \neq z$, we can set $S(m, z)=0$. So, suppose out ${ }_{\Pi}(m)=z$. Then

$$
\operatorname{Pr}_{\Pi}[m \mid x, y, z]=\frac{\operatorname{Pr}_{\Pi}[m, z \mid x, y]}{\operatorname{Pr}_{\Pi}[z \mid x, y]}=\frac{\operatorname{Pr}_{\Pi}[m \mid x, y]}{p_{Z \mid X Y}(z \mid x, y)}=\frac{\alpha(m, x) \beta(m, y)}{\rho(z, x) \sigma(z, y)},
$$

where we wrote $\operatorname{Pr}_{\Pi}[m, z \mid x, y]=\operatorname{Pr}_{\Pi}[m \mid x, y]$ (since $z=\operatorname{out}_{\Pi}(m)$ ), $\operatorname{Pr}_{\Pi}[m \mid x, y]=\alpha(m, x) \beta(m, y)$ (since $\Pi$ is a protocol) and $p_{Z \mid X Y}(z \mid x, y)=$ $\rho(z, x) \sigma(z, y)$ (since $p_{Z \mid X Y}$ is a simple function). Using the security guarantee for a symmetric function (Definition 1, with $z_{A}=z_{B}=z$ ), we get

$$
S_{1}(m, x, z)=S_{2}(m, y, z)=\frac{\alpha(m, x)}{\rho(z, x)} \cdot \frac{\beta(m, y)}{\sigma(z, y)}
$$

Now, fixing $(m, z)$ as above, consider all $(x, y)$ such that $p_{Z \mid X Y}(z \mid x, y)>0$. If the above expression is 0 for all choices of $(x, y)$, then we can simply set $S(m, z)=0$. Otherwise, there is some $x$ such that $S_{1}(m, x, z) \neq 0$. Then, considering the above expression for that $x$, we get that $\frac{\beta(m, y)}{\sigma(z, y)}$ equals a quantity that is independent of $y$ (and hence is a function of $(m, z)$ alone). Similarly, by considering $S_{2}(m, y, z)$, we get that $\frac{\alpha(m, x)}{\rho(z, x)}$ is a function of $(m, z)$ alone. Hence we have

$$
S_{1}(m, x, z)=S_{2}(m, y, z)=S(m, z)
$$

## B Example for Sect. 5.2

Theorem 7. There exists a randomized function that can be securely computed using a 3-round protocol, but cannot be securely computed using a uniquetranscript protocol with any number of rounds.


Fig. 2. This simple randomized function $p_{Z \mid X Y}$ is securely computable by a 3 round protocol (given in Fig. 3) that is not unique-transcript, but cannot be securely computed using any unique-transcript protocol (with any number of rounds).

Proof. Consider the simple randomized function $p_{Z \mid X Y}$ given in Fig. 2. This can be securely computed, and a 3-round protocol for that is given in Fig. 3. Note that this protocol is not unique-transcript. First we show that the protocol given in Fig. 3 is secure, i.e., it is correct and perfectly private; and later we show that no unique-transcript protocol can securely compute this function with any number of rounds.

Correctness: It follows from the fact that for every $x, y, z, p_{Z \mid X Y}(z \mid x, y)$ is equal to the sum of the probabilities on different paths leading to leaves labelled as $z$, where probability of a path is equal to the product of the probabilities (corresponding to the particular $x$ and $y$ ) appearing on the edges along that path.

Privacy: Consider an arbitrary $k \in[4]$. We need to show that for any transcript $m, p\left(m \mid x_{i}, y_{j}, z_{k}\right)$ must be the same for every $i, j \in[2], k \in[4]$. This trivially holds for every $z \in \mathcal{Z} \backslash\left\{z_{2}, z_{3}\right\}$, because there is a unique path from root to the leaf corresponding to $z$, which means that output being $z$ itself determines the whole transcript. But for $z_{2}$ and $z_{3}$ there are two possible transcripts. Fix any $z \in\left\{z_{2}, z_{3}\right\}$, say, $z_{2}$; a similar argument holds for $z=z_{3}$ as well. There are two transcripts $\left(m_{11}, m_{22}, m_{31}\right)$ and $\left(m_{12}, m_{23}, m_{33}\right)$ for $z_{2}$, implying two distinct paths. In order to show that the protocol is perfectly private, we need to show that $p\left(m_{11}, m_{22}, m_{31} \mid x_{i}, y_{j}, z_{2}\right)$ is the same for all $i, j \in[2]$. It can be easily verified that this is indeed the case, which implies that the protocol described in Fig. 3 is perfectly private.


Fig. 3. This describes a 3 -round protocol for securely computing $p_{Z \mid X Y}$, denoted by matrix $A$, where the first message is sent by Alice. This protocol is not uniquetranscript: both $A_{11}$ and $A_{12}$ have $z_{2}, z_{3}$ in common, and that is highlighted in red color. The meaning of the probabilities on the edges is as follows: If Alice's input is $x_{1}$, then she sends $m_{11}$ as the first message with probability $1 / 3$ and $m_{12}$ as the first message with probability $2 / 3$. If Alice's input is $x_{2}$, then she sends $m_{11}$ as the first message with probability $1 / 2$ and $m_{12}$ as the first message with probability $1 / 2$. If Alice sends $m_{11}$, then the problem reduces to securely computing $A_{11}$, and if Alice sends $m_{12}$, then the problem reduces to securely computing $A_{12}$. Suppose Alice reduces the problem to $A_{11}$. Now it is Bob's turn to send a message. If Bob's input is $y_{1}$, he sends $m_{21}$ with probability $2 / 3$ and $m_{22}$ with probability $1 / 3$, and so on $\ldots$ In the end, at the leaf nodes there is only one possible $z_{i}$ to output, and they output that element with probability 1. (Color figure online)

Now we show that no unique-transcript protocol (with any number of rounds) can securely compute $A:=p_{Z \mid X Y}$. Note that in a unique-transcript protocol, during any round, the party who is sending a message makes a partition of the output alphabet, and sends the other party the part (i.e., the reduced output alphabet) in which the output should lie. We show below that neither Alice nor Bob can make a partition in the first round itself. This implies that no uniquetranscript protocol exists for securely computing $A$ with any number of rounds.

- Alice cannot partition $\mathcal{Z}$ : Suppose, to the contrary, that Alice can partition $\mathcal{Z}$ in the first round; and, assume, w.l.o.g., that Alice makes two parts $\mathcal{Z}=$ $\mathcal{Z}_{1} \uplus \mathcal{Z}_{2}$.
Let $m_{i}, i=1,2$ denote the message that Alice sends to Bob in order to restrict the output to $\mathcal{Z}_{i}, i=1,2$. This implies that $p_{M_{1} \mid X Y Z}\left(m_{1} \mid x_{i}, y_{j}, z \in \mathcal{Z}_{1}\right)=1$ and $p_{M_{1} \mid X Y Z}\left(m_{2} \mid x_{i}, y_{j}, z \in \mathcal{Z}_{2}\right)=1$, for every $i, j \in\{1,2\}$. We show below that if Alice partitions $\mathcal{Z}=\mathcal{Z}_{1} \uplus \mathcal{Z}_{2}$, then the following must hold for every $i \in\{1,2\}$ :

$$
\begin{align*}
\sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) & =\sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{2}\right)  \tag{13}\\
\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) & =\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{2}\right) \tag{14}
\end{align*}
$$

It can be easily verified that the matrix $A$ does not satisfy the above two conditions for any non-trivial and disjoint $\mathcal{Z}_{1}, \mathcal{Z}_{2}$, which is a contradiction: since $\mathcal{Z}=\mathcal{Z}_{1} \uplus \mathcal{Z}_{2}$ and $|\mathcal{Z}|=4$, one of the following must hold: (i) either $\left|\mathcal{Z}_{1}\right|$ $=1$ or $\left|\mathcal{Z}_{2}\right|=1$, or (ii) either $\left|\mathcal{Z}_{1}\right|=2$ or $\left|\mathcal{Z}_{2}\right|=2$. This verification can be done easily even exhaustively. We show (13) and (14) below. In the following, $i$ belongs to $\{1,2\}$.

$$
\begin{align*}
p_{M_{1} \mid X Y}\left(m_{1} \mid x_{i}, y_{1}\right)= & \sum_{z \in \mathcal{Z}_{1}} p_{M_{1} Z \mid X Y}\left(m_{1}, z \mid x_{i}, y_{1}\right)+\sum_{z \in \mathcal{Z}_{2}} p_{M_{1} Z \mid X Y}\left(m_{1}, z \mid x_{i}, y_{1}\right) \\
= & \sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) \underbrace{p_{M_{1} \mid X Y Z}\left(m_{1} \mid x_{i}, y_{1}, z\right)}_{=1} \\
& +\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) \underbrace{p_{M_{1} \mid X Y Z}\left(m_{1} \mid x_{i}, y_{1}, z\right)}_{=0} \\
= & \sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) \tag{15}
\end{align*}
$$

Similarly we can show the following:

$$
\begin{align*}
& p_{M_{1} \mid X Y}\left(m_{1} \mid x_{i}, y_{2}\right)=\sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{2}\right),  \tag{16}\\
& p_{M_{1} \mid X Y}\left(m_{2} \mid x_{i}, y_{1}\right)=\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{1}\right) \tag{17}
\end{align*}
$$

$$
\begin{equation*}
p_{M_{1} \mid X Y}\left(m_{2} \mid x_{i}, y_{2}\right)=\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{i}, y_{2}\right) \tag{18}
\end{equation*}
$$

Since Alice sends the first message, which means that the Markov chain $M_{1}-X-Y$ holds. This implies that $p_{M_{1} \mid X Y}\left(m_{j} \mid x_{i}\right)=p_{M_{1} \mid X Y}\left(m_{j} \mid x_{i}, y_{1}\right)=$ $p_{M_{1} \mid X Y}\left(m_{j} \mid x_{i}, y_{2}\right)$ for every $j \in\{1,2\}$. Now comparing (15) and (16) gives (13), and (17) and (18) gives (14).

- Bob cannot partition $\mathcal{Z}$ : Switching the roles of Alice and Bob with eachother in the above argument and using the fact that for every partition $\mathcal{Z}=$ $\mathcal{Z}_{1} \uplus \mathcal{Z}_{2}$, the matrix $A$ does not satisfy the following two conditions, we can prove that Bob also cannot partition $\mathcal{Z}$. In the following, $i$ belongs to $\{1,2\}$.

$$
\begin{aligned}
\sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{1}, y_{i}\right) & =\sum_{z \in \mathcal{Z}_{1}} p_{Z \mid X Y}\left(z \mid x_{2}, y_{i}\right) \\
\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{1}, y_{i}\right) & =\sum_{z \in \mathcal{Z}_{2}} p_{Z \mid X Y}\left(z \mid x_{2}, y_{i}\right)
\end{aligned}
$$

This completes the proof of Theorem 7.

## References

[Bea89] Beaver, D.: Perfect privacy for two-party protocols. In: Feigenbaum, J., Merritt, M. (eds.) Proceedings of DIMACS Workshop on Distributed Computing and Cryptography, vol. 2, pp. 65-77. American Mathematical Society (1989)
[Dat16] Data, D.: Secure computation of randomized functions. In: IEEE International Symposium on Information Theory. ISIT 2016, Barcelona, Spain, 10-15 July 2016, pp. 3053-3057 (2016)
[Kil88] Kilian, J.: Founding cryptography on oblivious transfer. In: STOC, pp. 20-31. ACM (1988)
[Kil00] Kilian, J.: More general completeness theorems for secure two-party computation. In: Proceedings of 32 th STOC, pp. 316-324. ACM (2000)
[KMPS14] Kraschewski, D., Maji, H.K., Prabhakaran, M., Sahai, A.: A full characterization of completeness for two-party randomized function evaluation. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 659-676. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-55220-5_36
[KMQR09] Künzler, R., Müller-Quade, J., Raub, D.: Secure computability of functions in the IT setting with dishonest majority and applications to longterm security. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 238-255. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00457-5_15
[Kus89] Kushilevitz, E.: Privacy and communication complexity. In: FOCS, pp. 416-421. IEEE (1989)
[MPR09] Maji, H.K., Prabhakaran, M., Rosulek, M.: Complexity of multi-party computation problems: the case of 2-party symmetric secure function evaluation. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 256-273. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00457-5_16
[MPR12] Maji, H.K., Prabhakaran, M., Rosulek, M.: A unified characterization of completeness and triviality for secure function evaluation. In: Galbraith, S., Nandi, M. (eds.) INDOCRYPT 2012. LNCS, vol. 7668, pp. 40-59. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-34931-7_4
[MPR13] Maji, H.K., Prabhakaran, M., Rosulek, M.: Complexity of multi-party computation functionalities. In: Prabhakaran, M., Sahai, A. (eds.) Secure Multi-Party Computation. Cryptology and Information Security Series, vol. 10, pp. 249-283. IOS Press (2013)
[Rei09] Reingold, O. (ed.): TCC 2009. LNCS, vol. 5444. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00457-5


[^0]:    ${ }^{1}$ This is not true for every symmetric function. For instance, in a semi-honest secure protocol for XOR, the transcript must necessarily reveal both parties' inputs, but this cannot be simulated from the output without knowing one party's input. A function like XOR is not simple, though it is isomorphic to one.

[^1]:    ${ }^{2}$ There are other easier checks; see [MPR12, Lemma 1] for details.

