Oblivious-Transfer Complexity of Noisy Coin-Toss via Secure Zero Communication Reductions

Saumya Goyal[1⋆], Varun Narayanan[2⋆⋆], and Manoj Prabhakaran[3⋆⋆⋆]

[1] Stanford University
saumyagoyal01@gmail.com
[2] IIT Bombay
mp@cse.iitb.ac.in
[3] Technion
varunnkv@gmail.com

Abstract. In p-noisy coin-tossing, Alice and Bob obtain fair coins which are of opposite values with probability p. Its Oblivious-Transfer (OT) complexity refers to the least number of OTs required by a semi-honest perfectly secure 2-party protocol for this task. We show a tight bound of $\Theta(\log 1/p)$ for the OT complexity of p-noisy coin-tossing. This is the first instance of a lower bound for OT complexity that is independent of the input/output length of the function.
We obtain our result by providing a general connection between the OT complexity of randomized functions and the complexity of Secure Zero Communication Reductions (SZCR), as recently defined by Narayanan et al. (TCC 2020), and then showing a lower bound for the complexity of an SZCR from noisy coin-tossing to (a predicate corresponding to) OT.

1 Introduction

Consider two parties trying to do a "p-noisy coin-toss" such that each one gets a uniformly random bit, but with probability $p < 1/2$ the bits they obtain are different.[4] They would like to do this with semi-honest information-theoretic security (so that each one has no information about the other's bit, beyond what it learns from its own bit), using as few instances of Oblivious Transfer (OT) as possible.

An easy upper bound on the number of OTs needed is $O(\log 1/p)$, because they can obtain the desired outputs by evaluating a boolean circuit with that many binary gates on $O(\log 1/p)$ uniformly random bits from each party; the upper bound follows from the semi-honest GMW protocol [13,14,15] (requiring a couple of OTs for each non-linear gate). But it is *a priori* not at all clear if this

⋆ Work done while at IIT Bombay
⋆⋆ Supported by ERC Project NTSC (742754) and ISF Grants 1709/14 & 2774/20.
⋆⋆⋆ Supported by IITB Trust Lab.

[4] This functionality is sometimes referred to as sampling from a *binary symmetric source*. Note that for semi-honest security, as we consider, this is a cryptographically trivial task without any noise (i.e., when $p = 0$).

is the only way to carry out this computation. In particular, a protocol can rely on the semi-honest parties to sample non-uniform bits and use them as inputs in a protocol, and more generally, employ a protocol that does not involve a circuit evaluation at all.

Information-theoretic measures have been used to reason about the complexity of randomized functions in cryptographic and non-cryptographic settings. The most relevant technique to lower bound the OT complexity of general randomized functions is to use the "tension" of the resulting correlation [24]. However, it only yields a lower bound of *one* OT for sampling a noisy coin. Further, for the amortized setting, the lower bound on the rate degrades as the noise decreases.

In this work, we present for the first time an OT complexity lower bound that goes beyond the input/output length of a function, by showing that the number of OTs required for noisy coin-tossing is $\Theta(\log 1/p)$. Further, our lower bound also has a "direct sum" version, showing that tossing n such coins has OT complexity $\Theta(n \log 1/p)$. Remarkably – and in contrast to the information-theoretically derived lower bounds – our result shows that *OT complexity increases as p decreases*, although at the limit when $p = 0$, the OT complexity is 0. Indeed, an information-theoretic complexity measure like tension is unlikely to uncover this non-monotonic behavior of OT complexity.

Our main tool is *Secure Zero Communication Reductions* (SZCR) as defined recently in [23]. We extend the connection between SZCR complexity and OT complexity to randomized functions (in [23] this was limited to deterministic functions), and then show that the noisy coin-flip functionality has a large SZCR complexity of $\Omega(\log 1/p)$. Along the way, we develop a relaxation of SZCR complexity – which we term the *balanced embedding complexity* of a function – which is easier to interpret (especially for randomized functions) and which is sufficient to derive our lower bound.

OT Complexity and Randomized Functions. OT complexity of a (two-input) function – namely, the minimum number of instances of OT that is required by an information-theoretically secure[5] two-party computation protocol for evaluating the function – is a fundamental complexity measure. It follows from the results in the pioneering work in the 80's [13,15,14] that the OT complexity is upper bounded by the circuit complexity of the function. More recently, Beimel et al. [5] gave non-trivial upper bounds on OT complexity of all functions based on Private Information Retrieval (PIR) protocols, which become sub-exponential when instantiated using state-of-the art PIR results [11]. On the other hand, the few lower bounds that we do have – in terms of communication complexity [6] and "tension" [24] – are no larger than the (smaller) input and output length. Making further progress on OT complexity lower bounds faces major barriers, by implying lower bounds for circuit complexity (for explicit functions) or PIR (even existentially). Showing an existential lower bound that is super polynomial in the input length will imply super-logarithmic lower bounds for

[5] Throughout this paper, we consider semi-honest and perfect security, which arguably gives the cleanest notion of OT complexity.

the client computational complexity of 2-server PIR [5], and consequently lower bounds for codes on which PIR can be based. However, these barriers do not apply to *randomized functions*, motivating the current work.

Unfortunately, secure computation of randomized functions is relatively less well-understood, compared to deterministic functions. Indeed, even the characterization of which randomized functions are trivial (i.e., have 0 OT complexity) remains open.

While upper bounds on OT complexity of randomized functions can be obtained via upper bounds on OT complexity of appropriate deterministic functions (evaluated on randomized inputs), *this connection does not apply to lower bounds*. As an illustrative example, we present an inputless randomized function f which corresponds to evaluating a deterministic function g on random inputs, such that g has a positive OT complexity and f has 0 OT complexity![6] For $x, y \in [3]$, let $g(x, y) = M_{x,y}$ where $M = \begin{bmatrix} 1 & 1 & 2 \\ 4 & 5 & 2 \\ 4 & 3 & 3 \end{bmatrix}$. One way to compute this function would be for Alice and Bob to pick x and y respectively, and then use secure function evaluation to compute $g(x, y)$. Now, being an "undecomposable function", the function g cannot be securely computed without using any OTs [21,3]. However, f has a protocol that uses no OTs at all: one party can sample $M_{x,y}$ (without sampling x, y) and send it to the other one; then, independently, Alice samples x and Bob samples y conditioned on $M_{x,y}$.

Our result establishes, for the first time, a non-trivial lower bound technique for OT complexity of randomized functions. While this possibility was alluded to as a motivation in [23], the actual connection between SZCR and OT complexity established there was restricted to deterministic functions.

Our Contributions. We summarize our contributions as follows:

- The main result of this work is to show, for the first time, that the OT complexity of a randomized function can grow independent of the input/output size of the function. Specifically, we show that the OT complexity of securely sampling a noisy coin with flip probability p is $\Theta(\log 1/p)$. Further, this result has a "direct sum" version, so that sampling n independent copies of such a coin has OT complexity $\Theta(n \log 1/p)$.
- While proving this, we develop a more generally applicable tool, which shows that the complexity of an SZCR for a randomized function is a lower bound on the OT complexity of that function (denoted as $|f|_{\text{SZCR}} \leq |f|_{\mathsf{OT}}$). We do this by carefully generalizing the analysis in [23] where the same result was shown for deterministic functions.
- As a contribution towards facilitating future work on SZCR, we present a relaxation of SZCR complexity of randomized functions, namely, *balanced embedding* complexity, so that our result can be summarized as

$$|f|_{\text{EMB}} \leq |f|_{\text{SZCR}} \lesssim |f|_{\mathsf{OT}},$$

[6] This phenomenon occurs whenever g is undecomposable [21] but "simple" [22].

where the balanced embedding complexity $|f|_{\text{EMB}}$ is simpler to reason about. Indeed, our tight result on noisy coin-tossing is obtained by establishing a lower bound for balanced embedding complexity.

Related Work. There is a rich line of work in information-theoretic cryptography that studies the complexity of functions through the lens of secure 2-party computation. Starting with the seminal results in the 80's [20,21], complete and trivial functionalities for 2-party computation have been thoroughly characterized, for various levels of security (semi-honest, standalone, UC-secure) (see [22] for a survey). However, quantitative complexity results have been much sparser. The question of OT complexity was explicitly discussed by [4]. [6] presented a general lower bound in terms of the one-way communication complexity of the function. Important upper bounds of OT complexity follow from the semi-honest GMW protocol [13,14,15] and via PIR protocols [5]. Separate from the lower bound arguments in [6], a long line of works used information-theoretic tools for showing various complexity lower bounds for reductions in information-theoretic cryptography [4,10,26,19,27,18,16,17,9,25,24]; however, we may not expect such information-theoretic tools to uncover the non-monotonic behavior of OT complexity that we report here.

A similar sounding concept, called *Secure Non-Interactive Reduction* (SNIR) was introduced in [1] (also called Secure Non-Interactive Simulation or SNIS in [2]). It is instructive to compare both SNIR and SZCR with the standard notion of (semi-honest) secure reduction (SR) to a correlation like OT (i.e., the notion of OT complexity). Roughly put,

$$\text{SNIR} \Rightarrow \text{SR} \Rightarrow \text{SZCR}$$

indicating that SNIR is a "stronger" primitive than SR, which is in turn stronger than SZCR. While every function has an SR to the OT correlation (i.e., it is a complete correlation), that is not the case for SNIR: Indeed, there are no complete correlations for SNIR [1]. Both SNIR and SZCR are motivated by approaching the notoriously hard lower bound questions for SR, but they do it in different ways.

- Lower bounds (or impossibility results) for SNIR are an "easier" target than those for SR, and would provide a platform for nurturing new techniques; as and when we completely settle a question for SNIR (as is done in [7]),we can approach SR by relaxing the model (e.g., allow one-directional communication).
- Lower bounds for SZCR are formally (but not necessarily conceptually) harder than those for SR. In this case, one seeks to develop new techniques by asking simpler variants of the lower bound question: e.g., existential questions (a la the "invertible rank conjecture" of [23]) or lower bounds for randomized functions (as in this work) Also, the new perspective provided by SZCR may lead to fresh approaches to the original hard lower bound problems of SR.

2 Technical Overview

Our overall plan to obtain a lower bound on the OT complexity of a randomized function is to show that $|f|_{\text{EMB}} \lesssim |f|_{\mathsf{OT}}$, where $|f|_{\text{EMB}}$ is a new "balanced embedding complexity" that we define for functions, and then directly derive a lower bound for $|f|_{\text{EMB}}$. The significance of this connection is that, *a priori*, OT complexity is difficult to lower bound due to the complex possibilities in a protocol. On the other hand, a balanced embedding has a relatively simple structure that could allow us to easily derive a lower bound on $|f|_{\text{EMB}}$.

As such, the main technical contribution of this work is to define $|f|_{\text{EMB}}$ and to show that $|f|_{\text{EMB}} \lesssim |f|_{\mathsf{OT}}$. This involves a few different steps:

- Defining balanced embedding.
- Obtaining an easy lower bound on $|f|_{\text{EMB}}$, where f is the p-noisy coin-tossing functionality.
- Showing that $|f|_{\text{EMB}} \leq |f|_{\text{SZCR}}$, where $|f|_{\text{SZCR}}$ refers to the "SZCR complexity" of f.
- The final (and main) technical challenge is to show $|f|_{\text{SZCR}} \lesssim |f|_{\mathsf{OT}}$.

Below, we expand on each of these steps.

Balanced Embedding. Identifying randomized functions as weighted bipartite graphs, we define a form of weighted embedding of one such graph into another. The embedding is a "fuzzy" embedding that assigns weights relating how much one node in one graph is associated with a node in the other graph. In fact, there are two such weights (π and θ) which "balance" each other – hence the name balanced embedding.

Definition 1. *Let $G = (\mathcal{S}, \mathcal{T}, \boldsymbol{\omega})$ be a weighted bipartite graph, where $\mathcal{S}, \mathcal{T}$ form a bi-partition of the nodes of G and $\boldsymbol{\omega} : \mathcal{S} \times \mathcal{T} \to \mathbb{R}_{\geq 0}$ is the weight function. We define a* balanced embedding *of G into another bipartite graph $H = (\mathcal{U}, \mathcal{V}, \boldsymbol{\phi})$ as (π, θ) where $\pi, \theta : (\mathcal{U} \times \mathcal{S}) \cup (\mathcal{V} \times \mathcal{T}) \to \mathbb{R}_{\geq 0}$ are weight functions such that the following hold, for all $(\alpha, \beta) \in \mathcal{S} \times \mathcal{T}$:*

$$\sum_{v \in \mathcal{V}} \pi(v, \beta) \cdot \boldsymbol{\phi}(u, v) = \theta(u, \alpha) \cdot \boldsymbol{\omega}(\alpha, \beta) \qquad \forall u \in \mathcal{U} \quad (1)$$

$$\sum_{u \in \mathcal{U}} \pi(u, \alpha) \cdot \boldsymbol{\phi}(u, v) = \theta(v, \beta) \cdot \boldsymbol{\omega}(\alpha, \beta) \qquad \forall v \in \mathcal{V} \quad (2)$$

$$\sum_{u \in \mathcal{U}} \pi(u, \alpha) \cdot \theta(u, \alpha) = 1 \qquad \sum_{v \in \mathcal{V}} \pi(v, \beta) \cdot \theta(v, \beta) = 1 \quad \textit{if } \boldsymbol{\omega}(\alpha, \beta) > 0 \quad (3)$$

Given a randomized function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$, we define its characteristic bipartite graph as $G_f = (\mathcal{X} \times \mathcal{A}, \mathcal{Y} \times \mathcal{B}, \boldsymbol{\omega})$, where $\boldsymbol{\omega}((x, a), (y, b)) = \Pr[f(x, y) = (a, b)]$. We will be interested in the balanced embedding of G_f into the weighted graph $H_{\boldsymbol{\phi}} := (\mathcal{U}, \mathcal{V}, \boldsymbol{\phi})$, where $\boldsymbol{\phi} : \mathcal{U} \times \mathcal{V} \to \{0, 1\}$ is a predicate. In fact, we are specifically interested in predicates that correspond to multiple copies of OT:

$$\boldsymbol{\phi}^m_{\mathsf{OT}}((u_1, \ldots, u_m), (v_1, \ldots, v_m)) = \bigwedge_{i=1}^{m} \boldsymbol{\phi}_{\mathsf{OT}}(u_i, v_i)$$

where $\boldsymbol{\Phi}_{\mathsf{OT}}(u,v) = 1$ iff $\exists (x_0, x_1, b) \in \{0,1\}^3$ such that $u = (x_0, x_1)$ and $v = (b, x_b)$.

Definition 2. *The* balanced embedding complexity *of f, $|f|_{\text{EMB}}$ is the smallest m such that G_f has a balanced embedding into $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}}$.*

We remark that for our current result, the lower bound on the balanced embedding complexity of noisy coin-toss (sketched below) does not need to fully exploit all the conditions of a balanced embedding (e.g., in (3), $= 1$ can be replaced by > 0). However, for facilitating potential applications to other functions in the future, we retain the above version. For the sake of explicitness, we detail two constructions of balanced embedding of any Boolean function to OT predicate–from its truth table and from a Boolean circuit of the function–in Appendix A.

A Lower Bound for a Balanced Embedding of Noisy Coin-Tossing. Below we summarize the short argument to show that $|f|_{\text{EMB}} = \Omega(\log 1/p)$, where f is the p-noisy coin-toss functionality with $p < \frac{1}{2}$; i.e., if G_f has a balanced embedding into $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}}$, then $m = \Omega(\log 1/p)$.

Let $G_f = (\{0_A, 1_A\}, \{0_B, 1_B\}, \boldsymbol{\omega})$, where $\boldsymbol{\omega}(b_A, b_B) = (1-p)/2$ and $\boldsymbol{\omega}(b_A, (1-b)_B) = p/2$ for all $b \in \{0,1\}$. Let $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}} = (\mathcal{U}, \mathcal{V}, \boldsymbol{\Phi}^m_{\mathsf{OT}})$. Suppose (π, θ) is a balanced embedding of G_f to $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}}$.

Now, we choose $(u^*, \alpha^*) \in \mathcal{U} \times \{0_A, 1_A\}$ such that $\pi(u^*, \alpha^*) \geq \pi(u, \alpha)$ for all (u, α). W.l.o.g, let $\alpha^* = 0_A$ (as the other case is symmetric). Using (1)-(3) we can argue that for some $v^* \in \mathcal{V}$ such that $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u^*, v^*) = 1$, $\theta(v^*, 1_B) > 0$. Then, applying (2) to both $(\alpha, \beta) = (1_A, 1_B)$ and $(0_A, 1_B)$, and taking their ratio, we get

$$\frac{\sum_u \pi(u, 1_A) \cdot \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)}{\sum_u \pi(u, 0_A) \cdot \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)} = \frac{\boldsymbol{\omega}(1_A, 1_B)}{\boldsymbol{\omega}(0_A, 1_B)} = \frac{1-p}{p}.$$

Since $\pi(u, 1_A) \leq \pi(u^*, 0_A)$ for all u, and since $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u^*, v^*) = 1$,

$$|\{u : \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)\}| \geq \frac{\sum_u \pi(u, 1_A) \cdot \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)}{\sum_u \pi(u, 0_A) \cdot \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)} = \frac{1-p}{p}.$$

Since $|\{u : \boldsymbol{\Phi}^m_{\mathsf{OT}}(u, v^*)\}| = 2^m$, we have $m \geq \log(1/p) - 1$.

Virtually the same argument holds for the case of n noisy coin-flips, but with the ratio of probabilities used to obtain the bound being $\left(\frac{1-p}{p}\right)^n$, leading to a bound of $m = \Omega(n \log 1/p)$.

Recap of SZCR. We start with a quick recap of SZCR, as introduced in [23]. A μ-SZCR from a 2-party function f (which takes two inputs and produces two outputs, possibly randomized) to a predicate $\boldsymbol{\Phi}$, is a minimalistic computation model, in which Alice and Bob, on being given respective inputs x and y, produce respective outputs (a, u) and (b, v) *without any communication*, with the guarantee that (a, b) is distributed as $f(x, y)$ (or, in the case of deterministic functions, $(a, b) = f(x, y)$) conditioned on $\boldsymbol{\Phi}(u, v) = 1$. It is required that $\boldsymbol{\Phi}(u, v) = 1$ with

a fixed probability (irrespective of (x, y)) which is at least $2^{-\mu}$. The security condition captures the idea that Alice's view (which is considered to include the predicate's outcome $\mathbf{\Phi}(u, v)$, as well as her input x and output (a, u)) reveals nothing about Bob's input and output (y, b), beyond what is revealed by (x, a); similarly, Bob's view reveals nothing more about (x, a) than (y, b) itself reveals.

SZCR leads to a pair of natural complexity measures associated with a (possibly randomized) function f: smallest possible μ and m for which there is a μ-SZCR from f to $\mathbf{\Phi}_{\mathsf{OT}}^m$. In [23], the minimum such $\mu + m$ was suggested as a convenient complexity measure of a function f. In this work, for simplicity, we shall use the smallest m for which there is a μ-SZCR from f to $\mathbf{\Phi}_{\mathsf{OT}}^m$ for *any* (finite) μ as the complexity measure $|f|_{\text{SZCR}}$.[7]

Balanced Embedding and SZCR. Given an SZCR that reduces f to $\mathbf{\Phi}$, we obtain a balanced embedding of G_f into $H_{\mathbf{\Phi}}$. This amounts to assigning weights $\pi(u, \alpha)$ and $\theta(u, \alpha)$ for all $u \in \mathcal{U}$ and $\alpha \in \mathcal{X} \times \mathcal{A}$, and $\pi(v, \beta)$ and $\theta(v, \beta)$ for all $v \in \mathcal{V}$ and $\beta \in \mathcal{Y} \times \mathcal{B}$ in a way that satisfies (1), (2), and (3). Let $\Theta(\mathfrak{A}, \mathfrak{B})$ be an SZCR from f to $\mathbf{\Phi}$. For $\alpha = (x, a)$ and u, we choose $\pi(u, \alpha) \propto \mathsf{Pr}_{\mathfrak{A}}(u, a|x)$ and $\theta(u, \alpha)$ such that $\pi(u, \alpha) \cdot \theta(u, \alpha) = \mathsf{Pr}_{\hat{S}_A}(u|x, a, D = 1)$, where $\hat{S}_A$ is the simulator for Alice in the SZCR. For $\beta = (y, a)$ and v, $\pi(v, \beta)$ and $\theta(v, \beta)$ are chosen analogously. Having chosen the product of $\pi(u, \alpha)$ and $\theta(u, \alpha)$ in this manner,

$$\sum_u \pi(u, \alpha) \cdot \theta(u, \alpha) = \sum_u \mathsf{Pr}_{\hat{S}_A}(u|x, a, D = 1) = 1,$$

ensuring (3). Since $D = 1$ whenever $\mathfrak{A}$ and $\mathfrak{B}$ choose u and v, respectively, such that $\mathbf{\Phi}(u, v) = 1$, with π defined as above, and $\alpha = (x, a)$ and $\beta = (y, b)$,

$$\sum_v \pi(v, \beta)\mathbf{\Phi}(u, v) \propto \mathsf{Pr}_{\Theta}(D = 1, b|y, x, a, u) = \mathsf{Pr}_{\Theta}(b|x, y, a) \cdot \frac{\mathsf{Pr}_{\Theta}(D = 1, u|x, y, a, b)}{\mathsf{Pr}_{\Theta}(u|x, y, a, b)}.$$

Using the correctness of Θ conditioned on the event $D = 1$ and the fact that (u, a) and (v, b) are sampled depending only on x and y, respectively,

$$\mathsf{Pr}_{\Theta}(b|x, y, a) \cdot \frac{\mathsf{Pr}_{\Theta}(D = 1, u|x, y, a, b)}{\mathsf{Pr}_{\Theta}(u|x, y, a, b)} \propto \frac{\mathsf{Pr}_{\Theta}(u|x, y, a, b, D = 1)\mathsf{Pr}_f(a, b|x, y)}{\mathsf{Pr}_{\mathfrak{A}}(u, a|x)}.$$

At this point, noting that $\mathsf{Pr}_{\Theta}(u|x, y, a, b, D = 1) = \mathsf{Pr}_{\hat{S}_A}(u|x, a, D = 1)$ for all (y, b) and choosing the proportionality constant to be $\sqrt{\mathsf{Pr}_{\Theta}(D = 1|x, y)}$, we get (1). (2) is shown analogously.

We remark that in translating an SZCR to a balanced embedding, we ignore the SZCR security requirements related to the simulatability of views when the computation is *rejected* by the predicate.

OT Complexity and SZCR. In [23], it was shown that a 2-party secure function evaluation protocol Π^{OT} for a deterministic function f, using m OTs

[7] Our connection between SZCR and OT-based 2-PC does extend to both μ and m. But our formulation of balanced embedding complexity $|f|_{\text{EMB}}$ omits μ, and lower bounds on $|f|_{\text{EMB}}$ yield lower bounds on m rather than only on $m + \mu$.

can be transformed into a μ-SZCR from f to the predicate $\boldsymbol{\phi}_{\mathsf{OT}}$ corresponding to m instances of OT,[8] where $\mu = O(m)$. The high-level idea is for Alice and Bob to sample candidate pairs of views in Π^{OT} such that conditioned on $\boldsymbol{\phi}_{\mathsf{OT}}$ accepting the OTs in these views, these views are distributed correctly as in the protocol. Also, it would be ensured that the acceptance probability of the predicate is constant independent of x, y. Then the security guarantee of Π^{OT} translates to the security requirement of SZCR.

Being able to carry out the rejection sampling of views using $\boldsymbol{\phi}_{\mathsf{OT}}$ relies on the fact that protocols (secure or not) admit *transcript factorization*: i.e., the probability of a transcript q occurring in an execution of Π^{OT}, given inputs (x, y) and OT correlation (r, s) to the two parties respectively, can be written as

$$\mathsf{Pr}_{\Pi^{\mathsf{OT}}}(q|x, y, r, s) = \rho(x, r, q) \cdot \sigma(y, s, q),$$

for some functions ρ and σ. Given a particular transcript q (say, as a common reference string),[9] each of the two parties can locally sample its views from OTs (r or s, respectively), conditioned on its own input and q, with probability proportional to $\rho(x, r, q)$ or $\sigma(y, s, q)$, respectively with a proportionality constant independent of x or y; then, the probability that the parties end up with a valid joint view in the protocol (for which $\boldsymbol{\phi}_{\mathsf{OT}}(r, s) = 1$, and where all such (r, s) have the same probability) is proportional to that in the protocol, conditioned on (x, y, q).

Above, using proportionality constants that are independent of x and y runs into a problem since $\sum_r \rho(x, r, q)$ and $\sum_s \sigma(y, s, q)$ can depend on x, y. To resolve this, the parties are allowed to output an invalid r or s with some probability (implemented by setting $u = (u_0, r)$ and $v = (v_0, s)$, so that Alice or Bob can unilaterally force $\boldsymbol{\phi}_{\mathsf{OT}}(u, v) = 0$ by choosing a special value $\perp$ for u_0 or v_0, respectively).

To get a μ-SZCR, with $\mu = O(m)$, it is important to keep the probability with which the parties force aborting bounded. A key aspect in ensuring this turns out to be how the transcript q is chosen. As detailed in [23], if the function is "common-information-free,"–i.e., its characteristic bipartite graph is connected– then a single fixed transcript can be used. But otherwise, if the graph has multiple connected components, a transcript is chosen from among a small set of transcripts $q_1^*, \cdots, q_k^*$, indexed by the different values that the common information can take. An additional rejection step is introduced (see below), corresponding to rejecting a choice of this index that is not consistent with the input-output pair. A somewhat lengthy analysis shows that with appropriately chosen transcripts, the probability of the SZCR accepting is at least $2^{-O(m)}$.

[8] That is, $\boldsymbol{\phi}_{\mathsf{OT}}(u, v) = 1$ iff $u = (r_1, \cdots, r_m)$, $v = (s_1, \cdots, s_m)$ and each (r_i, s_i) is in the support of the OT correlation. Looking ahead, $\boldsymbol{\phi}_{\mathsf{OT}}$ in fact uses $m + 1$ instances of OT, where the extra instance is used as an "abort switch." Following the notation in [23], later, we denote $\boldsymbol{\phi}_{\mathsf{OT}}$ as $\boldsymbol{\phi}_{\mathsf{supp}(\mathsf{OT}^+)}$.

[9] The general definition in [23] allowed a CRS, or even more general correlations in an SZCR. For simplicity, we omit this from our adaptation, as we shall not need it for our specific result.

Extending to Randomized Functions. We first notice that the construction in [23] is no longer an SZCR when f is randomized and is not common-information free. To see this, we need to recall more details of the rejection step mentioned above for rejecting the wrong transcript index. Firstly, common information that Alice and Bob obtain when evaluating $f(x, y)$ to obtain outputs a and b respectively, corresponds to the connected component containing the edge $((x, a), (y, b))$ in a bipartite graph G_f representing f.[10] Now, in the SZCR of [23], given x and a common index ℓ, Alice checks if there is at least one a' such that the node (x, a') lies in the component specified by ℓ, and if so she samples r as described above, and computes an output a using the protocol Π on the view (x, r, q^*_ℓ). (Otherwise, she sets $u = (\perp, r)$ to force the predicate to fail.)

But when f is randomized, it is possible that the same transcript, and the same pair of inputs (x, y), could correspond to two different outputs (a_1, b_1) and (a_2, b_2), such that the edges $((x, a_1), (y, b_1))$ and $((x, a_2), (y, b_2))$ are in two different connected components of G_f.[11] So when the parties sample (r, s) conditioned on (x, y, q), it could correspond to either output. This breaks a crucial invariant in the analysis that when the predicate accepts, the outputs produced (a, b) will be such that $((x, a), (y, b))$ is in the connected component corresponding to the common index ℓ.

To fix this, we make a subtle change in the SZCR: Alice and Bob will first sample their respective outputs a and b (rather than computing them from r and s), and then check that the nodes (x, a) and (y, b) are in the connected component corresponding to the common index ℓ. This restores the invariant mentioned above, but necessitates a careful reanalysis. Our new analysis closely follows the original analysis, but needs to accommodate the above modification in the protocol, as well as the fact that G_f can have multiple edges (possibly in multiple connected components) of the form $((x, \cdot), (y, \cdot))$ for the same (x, y).

Our new proof incorporates an additional minor refinement. In [23], the SZCR constructed used a CRS, as this was a part of the model. Here, motivated by simplifying the (already minimalistic) SZCR model further, we restrict ourselves to a version which does not involve a CRS. Instead, the two parties guess a value of the CRS, and use the predicate $\boldsymbol{\Phi}_{\mathsf{OT}}$ to check if their guesses match. This does result in a slight quantitative degradation in the acceptance probability (when there are multiple connected components in G_f), but asymptotically, the result remains unchanged.

Finally, we remark that our result (as well as the one in [23]) is not only for an SZCR to OTs, but is shown for any "regular" complete correlation.

[10] For randomized functions, G_f is a weighted bipartite graph with the weight of an edge $((x, a), (y, b))$ being $\Pr[f(x, y) = (a, b)]$.

[11] Note that the common information that Alice and Bob obtain in an execution of the protocol Π^{OT} is not solely determined by the transcript, but also by their views of the OT correlation. Indeed, a protocol could use OTs to carry out an information-theoretically secure secret-key agreement protocol, and then use the key as a one-time pad for the rest of the transcript, so that the transcript by itself is distributed identically for all input-output pairs.

3 Preliminaries

Probability Notation. We adhere mostly to the notations used in [23]. In general, we denote a finite set by $\mathcal{X}, \mathcal{Y}, \ldots$ and so on. A member of $\mathcal{X}$ is denoted by x and a random variable taking values in $\mathcal{X}$ is denoted by X. The probability assigned by a distribution D (or a probabilistic process D) to a value x is denoted as $\mathsf{Pr}_D(x)$, or simply $\mathsf{Pr}(x)$, when the distribution is understood. Sampling x according to the distribution D is denoted as $x \leftarrow D$.

Functionalities and Correlations. A (potentially randomized) two party functionality $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ takes inputs x and y, respectively, from Alice and Bob and returns a and b, respectively, to them, where $(a, b) = f(x, y)$. We write $f_A : \mathcal{X} \times \mathcal{Y} \to \mathcal{A}$ to indicate the function obtained by projecting the output of f to the first coordinate (i.e., retaining only Alice's output). Similarly, $f_B : \mathcal{X} \times \mathcal{Y} \to \mathcal{B}$ denotes the function obtained from f by retaining only Bob's output.

A correlation $\boldsymbol{\psi}$ over a domain $\mathcal{R} \times \mathcal{S}$ is a 2-party functionality without inputs, i.e., $\boldsymbol{\psi} : \{\perp\} \times \{\perp\} \to \mathcal{R} \times \mathcal{S}$. The support of $\boldsymbol{\psi}$ is $\mathsf{supp}(\boldsymbol{\psi}) = \{(r, s) | \mathsf{Pr}_{\boldsymbol{\psi}}(r, s) > 0\}$. A correlation is said to be *regular* if (1) $\forall (r, s) \in \mathsf{supp}(\boldsymbol{\psi})$, $\mathsf{Pr}_{\boldsymbol{\psi}}(r, s) = \frac{1}{|\mathsf{supp}(\boldsymbol{\psi})|}$, (2) $\forall r \in \mathcal{R}$, $\sum_{s \in \mathcal{S}} \mathsf{Pr}_{\boldsymbol{\psi}}(r, s) = \frac{1}{|\mathcal{R}|}$, and (3) $\forall s \in \mathcal{S}$, $\sum_{r \in \mathcal{R}} \mathsf{Pr}_{\boldsymbol{\psi}}(r, s) = \frac{1}{|\mathcal{S}|}$. Common examples of regular correlations are those corresponding to Oblivious Transfer (OT) and Oblivious Linear Function Evaluation (OLE), and their n-fold repetitions. For $t \in \mathbb{N}$, t independent copies of a correlation $\boldsymbol{\psi}$ is denoted by $\boldsymbol{\psi}^t$.

Definition 3. *For a randomized function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ we define its* evaluation graph, *G_f as the bipartite graph on vertices $(\mathcal{X} \times \mathcal{A}) \cup (\mathcal{Y} \times \mathcal{B})$ such that the edge weight of an edge $((x, a), (y, b))$ is $\mathsf{Pr}_f(a, b | x, y)$.*

Two vertices u and v in G_f are said to be connected if there is a path from u to v consisting of edges with non-zero edge weight. Let $C \subseteq G_f$ be a connected component of G_f; we define:

$$\mathcal{X}_C = \{x : \exists a, y, b((x, a), (y, b)) \in C\} \quad \mathcal{Y}_C = \{y : \exists b, x, a((x, a), (y, b)) \in C\}$$

Predicates. A predicate is any deterministic function $\boldsymbol{\Phi} : \mathcal{U} \times \mathcal{V} \to \{0, 1\}$ with boolean output. The predicate $\boldsymbol{\Phi}_{(=^l)}$ takes a pair of l-bit strings u, v as input and accepts if $u = v$. Given a correlation $\boldsymbol{\psi}$ over $\mathcal{U} \times \mathcal{V}$, we define the predicate $\boldsymbol{\Phi}_{\mathsf{supp}(\boldsymbol{\psi})}$ so that $\boldsymbol{\Phi}_{\mathsf{supp}(\boldsymbol{\psi})}(u, v) = 1$ iff $(u, v) \in \mathsf{supp}(\boldsymbol{\psi})$. The predicate $\boldsymbol{\Phi}_{\mathsf{supp}^*(\boldsymbol{\psi})}$ is defined identically, except that we allow the domain of $\boldsymbol{\Phi}_{\mathsf{supp}^*(\boldsymbol{\psi})}$ to be $(\mathcal{U} \cup \{\perp\}) \times (\mathcal{V} \cup \{\perp\})$ where $\perp$ is a symbol not in $\mathcal{U} \cup \mathcal{V}$. Specifically, the predicate $\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^m)}$ allows a domain of $\{0, 1\}^{2m} \times \{0, 1\}^{2m}$ and accepts u, v if $\mathsf{Pr}_{\mathsf{OT}^m}(u, v) > 0$ and rejects otherwise; whereas, $\boldsymbol{\Phi}_{\mathsf{supp}^*(\mathsf{OT}^m)}$ behaves the same way but the input domain is now $(\{0, 1\}^{2m} \cup \{\perp\}) \times (\{0, 1\}^{2m} \cup \{\perp\})$.

Let $\boldsymbol{\Phi} : \mathcal{U} \times \mathcal{V} \to \{0, 1\}$ and $\boldsymbol{\Phi}' : \mathcal{U}' \times \mathcal{V}' \to \{0, 1\}$ be two predicates. Their product $\boldsymbol{\Phi} \cdot \boldsymbol{\Phi}'$ takes $(u, u') \in \mathcal{U} \times \mathcal{U}'$ and $(v, v') \in \mathcal{V} \times \mathcal{V}'$ as inputs and accepts if $\boldsymbol{\Phi}(u, v) = 1$ and $\boldsymbol{\Phi}'(u', v') = 1$.

Secure 2-party Communication Protocols A communication protocol between Alice and Bob using the correlation $\boldsymbol{\psi}$, denoted by Π^{Ψ}, proceeds as follows: Alice and Bob receive inputs x and y, respectively, and, additionally, they get r and s, respectively, where $(r, s) \leftarrow \boldsymbol{\psi}$. They exchange messages in rounds (message of a party in each round being a randomized function of their current view) to generate a transcript $q \in \mathcal{Q}$. Finally, Alice (resp. Bob) computes their output a (resp. b) by applying a (randomized) map Π_A^{out} (resp. Π_B^{out}) to their final view (x, r, q) (resp. (y, s, q)). Thus, the outcome of an execution of Π^{Ψ} on inputs (x, y) is the joint distribution described by

$$\begin{aligned}&\mathsf{Pr}_{\Pi^{\Psi}}(r, s, q, a, b|x, y)\\ &= \mathsf{Pr}_{\boldsymbol{\psi}}(r, s) \cdot \mathsf{Pr}_{\Pi^{\Psi}}(q|x, y, r, s) \cdot \mathsf{Pr}_{\Pi_A^{\text{out}}}(a|x, r, q) \cdot \mathsf{Pr}_{\Pi_B^{\text{out}}}(b|y, s, q), \forall r, s, q, a, b.\end{aligned} \tag{4}$$

The protocol Π^{Ψ} is said to compute the functionality $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ with perfect security if the distribution $\mathsf{Pr}_{\Pi^{\Psi}}(r, s, q, a, b|x, y)$ described above satisfies the following conditions:

Correctness: For all x, y,

$$\mathsf{Pr}_{\Pi^{\Psi}}(a, b|x, y) = \mathsf{Pr}_f(a, b|x, y), \forall a, b. \tag{5}$$

Privacy against Alice: There exists a randomized simulator $\hat{S}_A : \mathcal{X} \times \mathcal{A} \to \mathcal{R} \times \mathcal{Q}$ such that, for all a, x, y, such that $f_A(a|x, y) > 0$,

$$\mathsf{Pr}_{\Pi^{\Psi}}(r, q|x, y, a) = \mathsf{Pr}_{\hat{S}_A}(r, q|x, a), \forall r, q. \tag{6}$$

Privacy against Bob: There exists a randomized simulator $\hat{S}_B : \mathcal{Y} \times \mathcal{B} \to \mathcal{S} \times \mathcal{Q}$ such that, for all b, x, y, such that $f_B(b|x, y) > 0$,

$$\mathsf{Pr}_{\Pi^{\Psi}}(s, q|x, y, b) = \mathsf{Pr}_{\hat{S}_B}(s, q|y, b), \forall s, q. \tag{7}$$

Transcript Factorization. In any 2-party communication protocol Π^{Ψ}, the probability of generating the transcript, as a randomized function of the inputs (x, y) and the correlation (r, s), can be factorized into separate functions of (x, r) and (y, s). A transcript $q = (m_1, \ldots, m_N)$ is generated by the protocol if Alice produces the message m_1 given (x, r) in round 1, and then Bob produces m_2 given (y, s, m_1) in round 2, and so forth. That is,

$$\begin{aligned}\mathsf{Pr}_{\Pi^{\Psi}}(m_1, \ldots, m_N|x, y, r, s) = \mathsf{Pr}(m_1|x, r) \times \mathsf{Pr}(m_2|y, s, m_1) \times \ldots\\ \times \mathsf{Pr}(m_i|y, s, m_1, \ldots, m_{i-1}) \times \ldots.\end{aligned}$$

Hence, by collecting the products of odd factors as $\rho(x, r, m_1, \ldots, m_N)$ and even factors as $\sigma(y, s, m_1, \ldots, m_N)$, we can write the transcript as a product of separate functions of (x, r) and (y, s).

Formally, there exist *transcript factorization functions* $\rho : \mathcal{X} \times \mathcal{R} \times \mathcal{Q} \to [0,1]$ and $\sigma : \mathcal{Y} \times \mathcal{S} \times \mathcal{Q} \to [0,1]$, such that

$$\mathsf{Pr}_{\Pi^\Psi}(q|x,y,r,s) = \rho(x,r,q) \cdot \sigma(y,s,q). \tag{8}$$

It is worth noting that, for any x, y, r, s, functions ρ and σ by themselves are not probability mass functions. We shall use this important and well-known transcript factorization property (e.g., [8]) of a protocol in our constructions.

3.1 Zero-Communication Secure Reductions

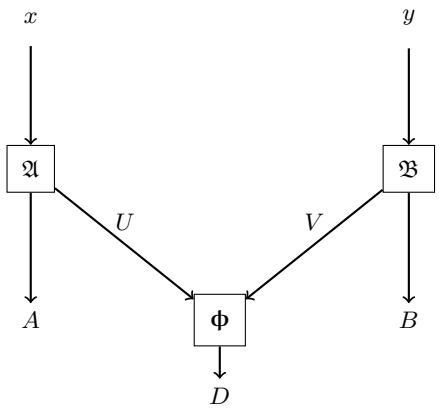

Fig. 1. The random variables involved in a SZCR.

The zero-communication reduction Θ from a functionality f to predicate $\boldsymbol{\Phi}$ is specified by a pair of randomized algorithms $(\mathfrak{A}, \mathfrak{B})$. The random variables involved in the reduction are illustrated in Figure 1. The reduction proceeds as follows: Alice and Bob receive inputs x, y to the functionality f, respectively. Alice samples $(a, u) \leftarrow \mathfrak{A}(x)$, where a is her proposed output for the functionality f, and u is her input to the predicate $\boldsymbol{\Phi}$. Similarly, Bob samples $(b, v) \leftarrow \mathfrak{B}(y)$. On receiving u, v from Alice and Bob, respectively, the predicate outputs $d = \boldsymbol{\Phi}(u, v)$. Thus, the outcome of an execution of Θ on inputs (x, y) is the joint distribution described by

$$\mathsf{Pr}_{\Theta}(u,v,a,b,d|x,y) = \mathsf{Pr}_{\mathfrak{A}}(u,a|x) \cdot \mathsf{Pr}_{\mathfrak{B}}(v,b|y) \cdot \mathsf{Pr}_{\boldsymbol{\Phi}}(d|u,v). \tag{9}$$

Definition 4. *Let $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ and $\boldsymbol{\Phi} : \mathcal{U} \times \mathcal{V} \to \{0,1\}$ be randomized functions. For any $\mu \geq 0$, a* μ-secure zero-communication reduction *(μ-SZCR) $\Theta(\mathfrak{A}, \mathfrak{B})$ from f to the predicate $\boldsymbol{\Phi}$ is a pair of probabilistic algorithms $\mathfrak{A} : \mathcal{X} \to \mathcal{U} \times \mathcal{A}$ and $\mathfrak{B} : \mathcal{Y} \to \mathcal{V} \times \mathcal{B}$ such that the following holds for the distribution described in (9).*

Non-Triviality and Weak Security $\exists \mu' \leq \mu$, $\forall (x,y) \in \mathcal{X} \times \mathcal{Y}$,

$$\mathsf{Pr}_{\Theta}(D = 1|x,y) = 2^{-\mu'}. \tag{10}$$

Correctness $\forall x,y,a,b \in \mathcal{X} \times \mathcal{Y} \times \mathcal{A} \times \mathcal{B}$,

$$\mathsf{Pr}_{\Theta}(a,b|x,y,D=1) = \mathsf{Pr}_{f}(a,b|x,y). \tag{11}$$

Security against Alice *There exists a randomized function* $S_A : \mathcal{X} \times \mathcal{A} \times \{0,1\} \to \mathcal{U}$ *such that* $\forall x,y,a \in \mathcal{X} \times \mathcal{Y} \times \mathcal{A}$ *such that* $\mathsf{Pr}_{f_A}(a|x,y) > 0$,

$$\mathsf{Pr}_{\Theta}(u|x,y,a,D=1) = \mathsf{Pr}_{S_A}(u|x,a,1). \tag{12}$$

$$\mathsf{Pr}_{\Theta}(u|x,y,D=0) = \sum_{a} \mathsf{Pr}_{f_A}(a|x,y) \cdot \mathsf{Pr}_{S_A}(u|x,a,0). \tag{13}$$

Security against Bob *There exists a randomized function* $S_B : \mathcal{Y} \times \mathcal{B} \times \{0,1\} \to \mathcal{V}$ *such that* $\forall x,y,b \in \mathcal{X} \times \mathcal{Y} \times \mathcal{B}$ *such that* $\mathsf{Pr}_{f_B}(b|x,y) > 0$,

$$\mathsf{Pr}_{\Theta}(v|x,y,b,D=1) = \mathsf{Pr}_{S_B}(v|y,b,1). \tag{14}$$

$$\mathsf{Pr}_{\Theta}(v|x,y,D=0) = \sum_{b} \mathsf{Pr}_{f_B}(b|x,y) \cdot \mathsf{Pr}_{S_B}(v|y,b,0). \tag{15}$$

In other words, in a SZCR, Alice and Bob compute "candidate outputs" a and b, as well as two messages u and v, respectively, such that correctness (i.e., $f(x,y) = (a,b)$) is required only when $\boldsymbol{\Phi}$ "accepts" (u,v). To be non-trivial, we require a lower bound $2^{-\mu}$ on the probability of $\boldsymbol{\Phi}$ accepting. Weak security requires that an "eavesdropper" who gets to observe whether the predicate $\boldsymbol{\Phi}$ accepts or not learns nothing about the inputs x,y. This is ensured by require the probability of accepting to remain the same as the inputs are changed. Note that as μ increases from 0 to ∞, the non-triviality and weak security constraint gets relaxed.

Finally, the security condition corresponds to security against passive corruption of one of Alice and Bob in a secure computation protocol (using $\boldsymbol{\Phi}$) that realizes the following functionality f_μ: After computing $(a,b) \leftarrow f(x,y)$, with probability $2^{-\mu}$ the functionality sends the respective outputs to the two parties ("accepting" case); with the remaining probability, it sends the output only to the corrupt party. In the above, (12) and (13) correspond to corrupting Alice, with the first one being the accepting case. Note that in these cases the adversary's view consists of U, in addition to the input x and the boolean variable D (accepting or not), which are given to the environment as well. In the accepting case, the environment also observes the outputs (a,b). In either case, $\hat{S}_A$ is given $(x, f_A(x,y), D)$ as inputs; in the accepting case, we naturally require that the simulated view has the same output a as $f_A(x,y)$ given to $\hat{S}_A$. Security conditions against Bob are interpreted analogously.

4 Balanced Embedding

For a randomized function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ and a deterministic predicate $\boldsymbol{\Phi} : \mathcal{U} \times \mathcal{V} \to \{0,1\}$, we study the *balanced embedding*, defined in Definition 1 in Section 2 of the evaluation graph G_f into the evaluation graph $G_{\boldsymbol{\Phi}}$.

Theorem 1. *If a randomized function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ has a* SZCR *to $\boldsymbol{\Phi} : \mathcal{U} \times \mathcal{V} \to \{0,1\}$, then there is a balanced embedding of the evaluation graph G_f into the predicate graph $G_{\boldsymbol{\Phi}}$.*

Proof. For each $\alpha = (x, a)$ and $\beta = (y, b)$, define

$$\pi(u,\alpha) = \frac{\mathsf{Pr}_{\mathfrak{A}}(u, a|x)}{\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \text{ and } \theta(u,\alpha) = \frac{\mathsf{Pr}_{S_A}(u|x,a,D=1)}{\pi(u,\alpha)} \forall u \in \mathcal{U}$$

$$\pi(v,\beta) = \frac{\mathsf{Pr}_{\mathfrak{B}}(v, b|y)}{\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \text{ and } \theta(v,\beta) = \frac{\mathsf{Pr}_{S_B}(v|y,b,D=1)}{\pi(v,\beta)} \forall v \in \mathcal{V}$$

Note that we set $\theta(u,\alpha) = 0$ whenever $\pi(u,\alpha) = 0$ and $\theta(v,\beta) = 0$ whenever $\pi(v,\beta) = 0$ since $\mathsf{Pr}(u|x,a)$ and $\mathsf{Pr}(v|y,b)$ are going to be 0 in these cases.
For each $u \in \mathcal{U}$, when $\beta = (y,b)$, and $\alpha = (x,a)$ for any (x,a) such that $\mathsf{Pr}_{\mathfrak{A}}(u,a|x) > 0$,

$$\begin{aligned}
\sum_{v\in\mathcal{V}} \pi(v,\beta)\cdot\boldsymbol{\Phi}(u,v) &\overset{(a)}{=} \frac{\sum_{v:\boldsymbol{\Phi}(u,v)=1}\mathsf{Pr}_{\mathfrak{B}}(v|y,b)\mathsf{Pr}_{\mathfrak{B}}(b|y)}{\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&\overset{(b)}{=} \frac{\mathsf{Pr}_{\Theta}(D=1|y,b,x,a,u)\mathsf{Pr}_{\mathfrak{B}}(b|y)}{\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&\overset{(c)}{=} \frac{\mathsf{Pr}_{\Theta}(u,D=1|y,b,x,a)\mathsf{Pr}_{\mathfrak{B}}(b|y)}{\mathsf{Pr}_{\mathfrak{A}}(u|x,a)\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&= \frac{\mathsf{Pr}_{\Theta}(u|x,y,a,b,D=1)\mathsf{Pr}_{\Theta}(D=1|x,a,y,b)\mathsf{Pr}_{\mathfrak{B}}(b|y)}{\mathsf{Pr}_{\mathfrak{A}}(u|x,a)\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&\overset{(d)}{=} \frac{\mathsf{Pr}_{S_A}(u|x,a,D=1)\mathsf{Pr}_{\Theta}(D=1,a,b|x,y)\mathsf{Pr}_{\mathfrak{B}}(b|y)}{\mathsf{Pr}_{\mathfrak{A}}(u|x,a)\mathsf{Pr}_{\Theta}(a,b|x,y)\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&\overset{(e)}{=} \frac{\mathsf{Pr}_{S_A}(u|x,a,D=1)\mathsf{Pr}_{\Theta}(D=1|x,y)\mathsf{Pr}_{\Theta}(a,b|x,y,D=1)}{\mathsf{Pr}_{\mathfrak{A}}(u|x,a)\mathsf{Pr}_{\mathfrak{A}}(a|x)\sqrt{\mathsf{Pr}_{\Theta}(D=1|x,y)}} \\
&\overset{(f)}{=} \frac{\mathsf{Pr}_{S_A}(u|x,a,D=1)\mathsf{Pr}_{f}(a,b|x,y)}{\pi(u,\alpha)} \\
&\overset{(g)}{=} \theta(u,\alpha)\cdot\boldsymbol{\omega}(\alpha,\beta). \qquad (16)
\end{aligned}$$

Here, (a) used the definition of $\pi(v,\beta)$; (b) used the fact that, for all (x,a) such that $\mathsf{Pr}_{\mathfrak{A}}(u,a|x) > 0$,

$$\mathsf{Pr}_{\Theta}(D=1|y,b,x,a,u) = \sum_{v:\boldsymbol{\Phi}(u,v)=1}\mathsf{Pr}_{\Theta}(v|y,b,x,a,u) = \sum_{v:\boldsymbol{\Phi}(u,v)=1}\mathsf{Pr}_{\mathfrak{B}}(v|y,b);$$

(c) used $\mathsf{Pr}_{\Theta}(u|x,a,y,b) = \mathsf{Pr}_{\mathfrak{A}}(u|x,a)$ as u is sampled locally by Alice in Θ; (d) follows from the privacy condition (12); (e) used $\mathsf{Pr}_{\Theta}(a,b|x,y) = \mathsf{Pr}_{\mathfrak{A}}(a|x)\cdot\mathsf{Pr}_{\mathfrak{B}}(b|y)$; (f) follows from (11) - the correctness of Θ, and the definition of $\pi(u,\alpha)$; finally, (g) follows from the definitions of $\theta(u,\alpha)$ and $\boldsymbol{\omega}(\alpha,\beta)$.

Similarly,

$$\sum_{u\in\mathcal{U}} \pi(u,\alpha)\cdot \boldsymbol{\Phi}(u,v) = \theta(v,\beta)\cdot \boldsymbol{\omega}(\alpha,\beta) \qquad \forall v\in\mathcal{V} \tag{17}$$

And finally, when $\boldsymbol{\omega}(\alpha,\beta) > 0$,

$$\begin{aligned}\sum_{u\in\mathcal{U}} \pi(u,\alpha)\cdot \theta(u,\alpha) &= \sum_{u\in\mathcal{U}} \mathsf{Pr}_{S_A}(u|x,a,D=1) = 1, \text{ and}\\ \sum_{v\in\mathcal{V}} \pi(v,\beta)\cdot \theta(v,\beta) &= \sum_{v\in\mathcal{V}} \mathsf{Pr}_{S_B}(v|y,b,D=1) = 1.\end{aligned} \tag{18}$$

Theorem follows from (16), (17) and (18). □

Theorem 2. *For the p-noisy coin-toss functionality f, the balanced embedding complexity $|f|_{\text{EMB}} = \Omega(\log 1/p)$, when $p < \frac{1}{2}$; i.e., if G_f has a balanced embedding into $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}}$, then $m = \Omega(\log 1/p)$.*

Proof. Let $G_f = (\mathcal{S},\mathcal{T},\boldsymbol{\omega})$, where $\mathcal{S} = \{0_A, 1_A\}$ and $\mathcal{T} = \{0_B, 1_B\}$, and $\boldsymbol{\omega}(b_A, b_B) = (1-p)/2$ and $\boldsymbol{\omega}(b_A,(1-b)_B) = p/2$ for all $b\in\{0,1\}$. Let $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}} = (\mathcal{U},\mathcal{V},\boldsymbol{\Phi}^m_{\mathsf{OT}})$. Suppose (π,θ) is a balanced embedding of G_f to $H_{\boldsymbol{\Phi}^m_{\mathsf{OT}}}$. Define $(u^*,\alpha^*)\in\mathcal{U}\times\{0_A,1_A\}$ as

$$(u^*,\alpha^*) = \underset{(u,\alpha):\theta(u,\alpha)>0}{\arg\max}\ \pi(u,\alpha). \tag{19}$$

Note that $\pi(u^*,\alpha^*) > 0$ since otherwise $\pi(u,\alpha)\cdot\theta(u,\alpha) = 0$ for all (u,α), which violates (3). W.l.o.g., let $\alpha^* = 0_A$ (the other case being symmetric).

Since $\theta(u^*,0_A) > 0$ and $\boldsymbol{\omega}(0_A,1_B) > 0$, by (1), $\pi(v^*,1_B) > 0$ for some $v^*\in\mathcal{V}$ such that $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u^*,v^*) = 1$. Further, by (2), $\theta(v^*,1_B) > 0$ since $\pi(u^*,0_A) > 0$ and $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u^*,v^*) = 1$. Applying (2) to both $(\alpha,\beta) = (1_A,1_B)$ and $(0_A,1_B)$, and taking their ratio, we get

$$\frac{\sum_u \pi(u,1_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)}{\sum_u \pi(u,0_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)} = \frac{\boldsymbol{\omega}(1_A,1_B)}{\boldsymbol{\omega}(0_A,1_B)} = \frac{1-p}{p}.$$

By (1), for all u such that $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*) = 1$, $\theta(u,1_A) > 0$ since $\pi(v^*,0_B) > 0$. But then, by (19), $\pi(u,1_A) \le \pi(u^*,0_A)$ for all u. Therefore, noting that $\boldsymbol{\Phi}^m_{\mathsf{OT}}(u^*,v^*) = 1$,

$$\begin{aligned}\sum_u \pi(u,1_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*) &\le \pi(u^*,0_A)|\{u:\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)\}|\\ &\le |\{u:\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)\}|\sum_u \pi(u,0_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*).\end{aligned}$$

Hence,

$$|\{u:\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)\}| \ge \frac{\sum_u \pi(u,1_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)}{\sum_u \pi(u,0_A)\cdot\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)} = \frac{1-p}{p}.$$

Since $|\{u:\boldsymbol{\Phi}^m_{\mathsf{OT}}(u,v^*)\}| = 2^m$, we have $m \ge \log(1/p) - 1$. □

Virtually the same argument holds for the case of n noisy coin-flips, but with the ratio of probabilities used to obtain the bound being $\left(\frac{1-p}{p}\right)^n$, leading to a bound of $m = \Omega(n \log 1/p)$.

5 SZCR from MPC protocols

In this section, we construct an SZCR from a potentially randomized function to OT check predicate from an MPC protocol for the function using OT; the complexity of the constructed SZCR coincides with the OT complexity of the MPC protocol. We will use this connection to obtain randomized functions that require super-linear OT complexity. The following theorem states more generally for all regular correlations. This is a generalization of one of the main results in [23] that proves this result for *deterministic* functions.

Theorem 3. *If a protocol Π^{Ψ} using a regular correlation $\boldsymbol{\psi}$ distributed over $\mathcal{R} \times \mathcal{S}$ computes a randomised function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ with perfect security, then there exists a μ-SZCR to $\boldsymbol{\Phi}_{(=\lceil \log k \rceil)} \cdot \boldsymbol{\Phi}_{\mathsf{supp}^*(\boldsymbol{\psi})}$, where k is the number of connected components in the evaluation graph G_f and $\mu \leq \log \frac{|\mathcal{R}||\mathcal{S}||\mathcal{X}|^2|\mathcal{Y}|^2|\mathcal{A}||\mathcal{B}|}{|\mathsf{supp}(\boldsymbol{\psi})|}$.*

This theorem is proved through Claim 1-Claim 6. In the following section, we make some observations and define some quantities that are used in the construction and analysis of the SZCR we construct.

Let $\mathcal{Q}$ be the set of all transcripts that can be produced in the protocol Π^{Ψ}. We observed that communication protocols admit transcript factorization; i.e., there exist functions $\rho : \mathcal{X} \times \mathcal{R} \times \mathcal{Q} \to [0,1]$ and $\sigma : \mathcal{Y} \times \mathcal{S} \times \mathcal{Q} \to [0,1]$ such that, when x, y are the inputs to Alice and Bob, and (r, s) is the realization of the correlation $\boldsymbol{\psi}$, for any transcript q,

$$\mathsf{Pr}_{\Pi^{\Psi}}(q|x, y, r, s) = \rho(x, r, q)\sigma(y, s, q).$$

The following set of observations are about a protocol Π^{Ψ} that uses a correlation $\boldsymbol{\psi}$ and computes a given function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{A} \times \mathcal{B}$ with perfect security. We will exploit the perfect security of the protocol to establish how the protocol behaves in each connected component of the evaluation graph G_f.

Lemma 1. *For each connected component C of the evaluation graph G_f, if $((x_1, a_1), (y_1, b_1))$ and $((x_2, a_2), (y_2, b_2))$ belong to C, then*

$$\mathsf{Pr}_{\Pi^{\Psi}}(q|x_1, y_1, a_1, b_1) = \mathsf{Pr}_{\Pi^{\Psi}}(q|x_2, y_2, a_2, b_2), \forall q \in \mathcal{Q}.$$

Proof. Since Π^{Ψ} is perfectly secure, there exists simulators $\hat{S}_A$ and $\hat{S}_B$ such that, for all x, y, a, b such that $\mathsf{Pr}_f(a, b|x, y) > 0$,

$$\mathsf{Pr}_{\Pi^{\Psi}}(q, r|x, y, a, b) = \mathsf{Pr}_{\hat{S}_A}(q, r|x, a) \quad \mathsf{Pr}_{\Pi^{\Psi}}(q, s|x, y, a, b) = \mathsf{Pr}_{\hat{S}_B}(q, s|y, b).$$

Hence, if edges $((x, a), (y, b))$ and $((x, a), (y', b'))$ belong to C, then, for all $q \in \mathcal{Q}$,

$$\mathsf{Pr}_{\Pi^{\Psi}}(q|x, y, a, b) = \sum_r \mathsf{Pr}_{\hat{S}_A}(q, r|x, a) = \mathsf{Pr}_{\Pi^{\Psi}}(q|x, y', a, b').$$

A similar condition holds for edges $((x,a),(y,b))$ and $((x',a'),(y,b))$ belonging to C. Hence, if edges $((x_1,a_1),(y_1,b_1))$, $((x_2,a_2),(y_2,b_2))$ belong to C, then applying the above two conditions alternatively along a path that begins with the edge $((x_1,a_1),(y_1,b_1))$ and ends with the edge $((x_2,a_2),(y_2,b_2))$, we get the statement of the lemma. □

Lemma 2. *If* $\mathsf{Pr}_{\Pi^\Psi}(a_1,b_1|x,y,r,s,q) > 0$ *and* $\mathsf{Pr}_{\Pi^\Psi}(a_2,b_2|x,y,r,s,q) > 0$ *then* $((x,a_1),(y,b_1))$ *and* $((x,a_2),(y,b_2))$ *belong to the same connected component of* G_f.

Proof. For all $(a,b) \in \{(a_1,b_1),(a_2,b_2)\}$, we have

$$\mathsf{Pr}_{\Pi_A^{\text{out}}}(a|x,r,q) \cdot \mathsf{Pr}_{\Pi_B^{\text{out}}}(b|y,s,q) = \mathsf{Pr}_{\Pi^\Psi}(a,b|x,y,r,s,q) > 0.$$

Hence, $\mathsf{Pr}_{\Pi_A^{\text{out}}}(a|x,r,q) > 0$ for $a \in \{a_1,a_2\}$ and $\mathsf{Pr}_{\Pi_B^{\text{out}}}(b|y,s,q) > 0$ for $b \in \{b_1,b_2\}$. Thus, by the perfect correctness of Π^Ψ,

$$\begin{aligned}\mathsf{Pr}(f(x,y) = (a_2,b_1)) &> \mathsf{Pr}_{\Pi^\Psi}(a_2,b_1|x,y,r,s,q)\\ &= \mathsf{Pr}_{\Pi_A^{\text{out}}}(a_2|x,r,q) \cdot \mathsf{Pr}_{\Pi_B^{\text{out}}}(b_1|y,s,q) > 0.\end{aligned}$$

This implies that $((x,a_2),(y,b_1))$ has non-zero weight in G_f, consequently, $(x,a_1)-(y,b_1)-(x,a_2)-(y,b_2)$ is a path in G_f, implying the statement of the lemma. □

In Lemma 1, we showed that for all connected component C of the evaluation graph G_f, and for all edges $((x_1,a_1),(y_1,b_1))$ and $((x_2,a_2),(y_2,b_2))$ belonging to C,

$$\mathsf{Pr}_{\Pi^\Psi}(q|x_1,y_1,a_1,b_1) = \mathsf{Pr}_{\Pi^\Psi}(q|x_2,y_2,a_2,b_2).$$

By an abuse of notation, we denote $\mathsf{Pr}_{\Pi^\Psi}(q|x,y,a,b)$ for all edges $((x,a),(y,b))$ belonging to the connected component C by $\mathsf{Pr}_{\Pi^\Psi}(q|C)$.

To present our SZCR protocol $\Theta(\mathfrak{A},\mathfrak{B})$ from f to $\boldsymbol{\phi}$, that is constructed from the secure computation protocol for f, we need the following quantities.

Definition 5. *For each connected component* C *in* G_f*, we define the following quantities:*

$$\rho_C^\dagger(q) = \max_{x \in \mathcal{X}_C} \sum_r \rho(x,r,q) \quad \sigma_C^\dagger(q) = \max_{y \in \mathcal{Y}_C} \sum_s \sigma(v,s,q)$$

Lemma 3. *For every connected component* C *in* G_f*, there exists* $q^* \in \mathcal{Q}$ *such that* $\mathsf{Pr}_{\Pi^\Psi}(q^*|C) > 0$ *and*

$$\rho_C^\dagger(q^*)\sigma_C^\dagger(q^*) \le |\mathcal{R}||\mathcal{S}||\mathcal{X}_C||\mathcal{Y}_C|\mathsf{Pr}_{\Pi^\Psi}(q^*|C)$$

Proof. Define $\widetilde{\boldsymbol{\psi}}$ to be the uniform distribution over $\mathcal{R}\times\mathcal{S}$. Consider the protocol $\Pi^{\widetilde{\Psi}}$ obtained by replacing the correlation $\boldsymbol{\psi}$ in Π^Ψ with $\widetilde{\boldsymbol{\psi}}$. Hence,

$$\mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q,r,s|x,y) = \mathsf{Pr}_{\widetilde{\boldsymbol{\psi}}}(r,s) \cdot \mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q|r,s,x,y) = \frac{\rho(x,r,q)\sigma(y,s,q)}{|\mathcal{R}||\mathcal{S}|} \tag{20}$$

Note that ρ, σ induced by Π^{Ψ} is well-defined for $(x, r, q) \in \mathcal{X} \times \mathcal{R} \times \mathcal{Q}$ and $(y, s, q) \in \mathcal{Y} \times \mathcal{S} \times \mathcal{Q}$, respectively.

By imposing a distribution over the inputs, namely the uniform distribution over $\mathcal{X}_C \times \mathcal{Y}_C$, for all $q \in \mathcal{Q}$, define:

$$\mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q) = \sum_{(x,y)\in\mathcal{X}_C\times\mathcal{Y}_C} \sum_{(r,s)\in\mathcal{R}\times\mathcal{S}} \frac{\mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q|x,y,r,s)}{|\mathcal{X}_C||\mathcal{Y}_C||\mathcal{R}||\mathcal{S}|}.$$

Since $\mathsf{Pr}_{\Pi^{\Psi}}(q|C)$ and $\mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q)$ are distributions over $\mathcal{Q}$, there exists $q^* \in \mathcal{Q}$ such that

$$\mathsf{Pr}_{\Pi^{\Psi}}(q^*|C) \geq \mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q^*) > 0.$$

Hence,

$$\begin{aligned}\mathsf{Pr}_{\Pi^{\Psi}}(q^*|C) \geq \mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q^*) &= \sum_{(x,y)\in\mathcal{X}_C\times\mathcal{Y}_C} \sum_{(r,s)\in\mathcal{R}\times\mathcal{S}} \frac{\mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q^*|x,y,r,s)}{|\mathcal{X}_C||\mathcal{Y}_C||\mathcal{R}||\mathcal{S}|} \\ &\geq \frac{\sum_{(r,s)\in\mathcal{R}\times\mathcal{S}} \mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q^*|x,y,r,s)}{|\mathcal{X}_C||\mathcal{Y}_C||\mathcal{R}||\mathcal{S}|}, \forall (x,y) \in \mathcal{X}_C \times \mathcal{Y}_C. \end{aligned} \tag{21}$$

Choose $(x^*, y^*) \in \mathcal{X}_C \times \mathcal{Y}_C$ such that

$$x^* = \arg\max_{x\in\mathcal{X}_C} \sum_r \rho(x, r, q^*) \quad y^* = \arg\max_{y\in\mathcal{Y}_C} \sum_s \sigma(y, s, q^*).$$

Then, by Definition 5, $\rho^{\dagger}_C(q^*) = \sum_r \rho(x^*, r, q^*)$ and $\sigma^{\dagger}_C(q^*) = \sum_s \rho(y^*, s, q^*)$. Hence,

$$\begin{aligned}\rho^{\dagger}_C(q^*)\sigma^{\dagger}_C(q^*) &= \sum_{r,s} \rho(x^*, r, q^*)\sigma(y^*, s, q^*) \\ &\overset{(a)}{=} |\mathcal{R}||\mathcal{S}| \sum_{r,s} \mathsf{Pr}_{\Pi^{\widetilde{\Psi}}}(q^*, r, s|x^*, y^*) \\ &\overset{(b)}{\leq} |\mathcal{R}||\mathcal{S}||\mathcal{X}_C||\mathcal{Y}_C| \cdot \mathsf{Pr}_{\Pi^{\Psi}}(q^*|C),\end{aligned}$$

where (a) follows from (20) and (b) follows from (21). This concludes the proof. □

Definition 6. *Let $C_1, \ldots, C_k$ be the set of all connected components of the evaluation graph G_f. For each C_i, $i \in [k]$, Lemma 3 guarantees that there exists $q^\star_i \in \mathcal{Q}$ such that $\mathsf{Pr}_{\Pi^{\Psi}}(q^\star_i|C_i) > 0$ and*

$$\rho^{\dagger}_C(q^\star_i)\sigma^{\dagger}_C(q^\star_i) \leq |\mathcal{R}||\mathcal{S}||\mathcal{X}_C||\mathcal{Y}_C|\mathsf{Pr}_{\Pi^{\Psi}}(q^\star_i|C).$$

We define the distribution $\boldsymbol{\lambda}$ over $[k]$ as:

$$\mathsf{Pr}_{\boldsymbol{\lambda}}(i) = \frac{\sqrt{c_i}}{\sum_{t\in[k]} \sqrt{c_t}}, \text{ where } c_i = \frac{\rho^{\dagger}_{C_i}(q^\star_i) \cdot \sigma^{\dagger}_{C_i}(q^\star_i)}{\mathsf{Pr}_{\Pi^{\Psi}}(q^\star_i|C_i)}.$$

Let $C_1, \ldots, C_k$ be the connected components in the evaluation graph G_f. Let the inputs to f be $x \in \mathcal{X}$ and $y \in \mathcal{Y}$; for $i \in [k]$, choose $q_i^\star$ as defined in Definition 6 with respect to the protocol Π^Ψ; choose $\rho_i^\dagger = \rho_{C_i}^\dagger(q_i^\star)$ and $\sigma_i^\dagger = \sigma_{C_i}^\dagger(q_i^\star)$ as defined in Definition 5; finally, let the distribution $\boldsymbol{\lambda}$ over $[k]$ be as defined in Definition 6.

$\mathfrak{A}(x)$: Sample $i \leftarrow \boldsymbol{\lambda}$ and $r \in \mathcal{R}$ with probability $\frac{\rho(x,r,q_i^\star)}{\rho_i^\dagger}$, and with remaining probability set $r = \perp$. If $r \neq \perp$, sample $a \leftarrow \Pi_A^{\text{out}}(x, r, q_i^\star)$ and set $A = a$, otherwise $A = \perp$. If $r \neq \perp$ and there exist y', b' such that $((x,a),(y',b')) \in C_i$ then set $U = (I, R)$ to (i, r), else to $(i, \perp)$.

$\mathfrak{B}(y)$: Sample $j \leftarrow \boldsymbol{\lambda}$ and $s \in \mathcal{S}$ with probability $\frac{\sigma(y,s,q_j^\star)}{\sigma_j^\dagger}$, and with remaining probability set $s = \perp$. If $s \neq \perp$, sample $b \leftarrow \Pi_B^{\text{out}}(y, s, q_j^\star)$ and set $B = b$, otherwise set $B = \perp$. If $s \neq \perp$ and there exist x', a' such that $((x',a'),(y,b)) \in C_j$ then set $V = (J, S)$ to (j, s), else to $(j, \perp)$.

$\boldsymbol{\Phi}_{(=^{\lceil \log k \rceil})} \cdot \boldsymbol{\Phi}_{\mathsf{supp}^*(\Psi)}$: Returns $D = 1$ if $\boldsymbol{\Phi}_{(=^{\lceil \log k \rceil})}(i,j) = 1$ (i.e., $j = i$) and $\boldsymbol{\Phi}_{\mathsf{supp}^*(\Psi)}(r,s) = 1$ (i.e., $r \sim s$).

Fig. 2. An SZCR protocol $\Theta(\mathfrak{A}, \mathfrak{B})$ from f to $\boldsymbol{\Phi}$ constructed from the secure computation protocol Π^Ψ using the correlation $\boldsymbol{\Psi}$ that computes f with perfect security.

Now we present our SZCR protocol in Figure 2, which is analyzed below.

Proof of correctness. In the sequel, we will consider x, y, i, j, r, s as defined in Figure 2. R, S, I, J are the random variables corresponding to r, s, i, j, respectively. Recollect that, we shorten $\mathsf{Pr}(R = r, S = s, I = i, J = j)$ as $\mathsf{Pr}(r, s, i, j)$, whenever there is no scope for confusion. We first make the following claims that will be later used to prove the correctness in Claim 2.

Claim 1. *If $j \neq i$ or $((x,a),(y,b)) \notin C_i$, then $\mathsf{Pr}_\Theta(a, b, D = 1 | x, y, i, j) = 0$.*

Proof. Let E be the event $(D = 1, A = a, B = b)$. If $j \neq i$, then $\boldsymbol{\Phi}_{(=^{\lceil \log k \rceil})}(i,j) = 0$, hence $D = 0$, hence we consider the case where $j = i$. Towards a contradiction, suppose $j = i$ and $((x,a),(y,b)) \notin C_i$ and E occurs with non-zero probability. Event E occurs only if there exist r, s such that $r \sim s$, $\rho(x, r, q_i^\star) > 0$, $\sigma(y, s, q_i^\star) > 0$, $\mathsf{Pr}_{\Pi_A^{\text{out}}}(a | x, r, q_i^\star) > 0$, and $\mathsf{Pr}_{\Pi_B^{\text{out}}}(b | y, s, q_i^\star) > 0$.

$$
\begin{aligned}
&\mathsf{Pr}_{\Pi^\Psi}(a, b | x, y) \\
&\quad \geq \mathsf{Pr}_{\Pi^\Psi}(q_i^\star, a, b, r, s | x, y) \\
&\quad = \mathsf{Pr}_{\boldsymbol{\Psi}}(r, s) \cdot \mathsf{Pr}_{\Pi^\Psi}(q_i^\star | x, y, r, s) \cdot \mathsf{Pr}_{\Pi^\Psi}(a, b | x, y, r, s, q_i^\star) \\
&\quad = \frac{\rho(x, r, q_i^\star) \cdot \sigma(y, s, q_i^\star) \cdot \mathsf{Pr}_{\Pi_A^{\text{out}}}(a | x, r, q_i^\star) \cdot \mathsf{Pr}_{\Pi_B^{\text{out}}}(b | y, s, q_i^\star)}{|\mathsf{supp}(\boldsymbol{\Psi})|} > 0.
\end{aligned}
$$

Thus, by the perfect correctness of Π^Ψ, $\mathsf{Pr}_f(a, b | x, y) = \mathsf{Pr}_{\Pi^\Psi}(a, b | x, y) > 0$. Additionally, by the construction of Θ, E occurs only if there exist b', y' such that $((x,a),(y',b')) \in C_i$ since, otherwise, Alice would have aborted by sending $\perp$ (instead of sending some $u \in \mathcal{U}$). Hence, the edges $((x,a),(y,b))$ and

$((x,a),(y',b'))$ have non-zero weights in G_f and $((x,a),(y',b')) \in C_i$. But then, $((x,a),(y,b)) \in C_i$, a contradiction. This proves the claim. □

Claim 2. *The probability of acceptance for any inputs x, y is independent of the inputs, and is given by:*

$$\mathsf{Pr}_{\Theta}(D=1|x,y) = \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{l\in[k]}\sqrt{c_l}\right)^2} \tag{22}$$

Proof. Fix inputs x, y. We have,

$$\begin{aligned}\mathsf{Pr}_{\Theta}(D=1|x,y) &= \sum_{i,j,a,b} \mathsf{Pr}_{\Theta}(D=1,i,j,a,b|x,y)\\ &= \sum_{i,j,a,b} \mathsf{Pr}_{\Theta}(i,j|x,y)\cdot\mathsf{Pr}_{\Theta}(D=1,a,b|x,y,i,j).\end{aligned}$$

If $j \neq i$, then $D=0$, furthermore, $\mathsf{Pr}_{\Theta}(i,j|x,y) = \mathsf{Pr}_{\boldsymbol{\lambda}}(i)\cdot\mathsf{Pr}_{\boldsymbol{\lambda}}(j)$. Hence,

$$\begin{aligned}&\mathsf{Pr}_{\Theta}(D=1|x,y)\\ &= \sum_{i\in[k],j=i}\sum_{a,b}\mathsf{Pr}_{\boldsymbol{\lambda}}(i)\cdot\mathsf{Pr}_{\boldsymbol{\lambda}}(j)\cdot\mathsf{Pr}_{\Theta}(D=1,a,b|x,y,i,j)\\ &= \sum_{i\in[k],j=i}\mathsf{Pr}^2_{\boldsymbol{\lambda}}(i)\sum_{a,b}\sum_{r\sim s}\mathsf{Pr}_{\Theta}(r,s|x,y,i,j)\cdot\mathsf{Pr}_{\Theta}(a,b|x,y,i,j,r,s)\\ &= \frac{1}{\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}c_i\sum_{a,b}\sum_{r\sim s}\mathsf{Pr}_{\Pi^{\mathrm{out}}_A}(a|x,r,q_i^\star)\cdot\mathsf{Pr}_{\Pi^{\mathrm{out}}_B}(b|y,s,q_i^\star)\frac{\rho(x,r,q_i^\star)\cdot\sigma(y,s,q_i^\star)}{\rho_i^\dagger\sigma_i^\dagger}\end{aligned}$$

But, $\mathsf{Pr}_{\Pi^{\mathrm{out}}_A}(a|x,r,q_i^\star)\cdot\mathsf{Pr}_{\Pi^{\mathrm{out}}_B}(b|y,s,q_i^\star) = \mathsf{Pr}_{\Pi_\Psi}(a,b|x,y,q_i^\star,r,s)$ and, by transcript factorization property, $\rho(x,r,q_i^\star)\cdot\sigma(y,s,q_i^\star) = \mathsf{Pr}_{\Pi_\Psi}(q_i^\star|x,y,r,s)$. Furthermore, $c_i = \frac{\rho_i^\dagger\sigma_i^\dagger}{\mathsf{Pr}_{\Pi_\Psi}(q_i^\star|C_i)}$. Applying these observations to the RHS,

$$\begin{aligned}&\mathsf{Pr}_{\Theta}(D=1|x,y)\\ &= \frac{1}{\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\frac{\rho_i^\dagger\sigma_i^\dagger}{\mathsf{Pr}_{\Pi_\Psi}(q_i^\star|C_i)}\sum_{a,b}\sum_{r\sim s}\frac{\mathsf{Pr}_{\Pi_\Psi}(a,b|x,y,q_i^\star,r,s)\cdot\mathsf{Pr}_{\Pi_\Psi}(q_i^\star|x,y,r,s)}{\rho_i^\dagger\sigma_i^\dagger}.\end{aligned}$$

By Claim 1, $\mathsf{Pr}_{\Theta}(D=1,a,b|x,y,i,j) = 0$ if $((x,a),(y,b))$ is not an edge in C_i (or $j \neq i$). Furthermore, by definition,

$$0 < \mathsf{Pr}_{\Pi_\Psi}(q_i^\star|C_i) = \mathsf{Pr}_{\Pi_\Psi}(q_i^\star|x,y,a,b), \text{ for all } ((x,a),(y,b)) \in C_i.$$

Applying both these facts to the RHS,

$$\begin{aligned}
&\mathsf{Pr}_{\Theta}(D=1|x,y)\\
&\quad= \frac{1}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\sum_{(a,b):((x,a),(y,b))\in C_i}\sum_{r\sim s}\frac{\mathsf{Pr}_{\Pi^\Psi}(a,b,q_i^\star|x,y,r,s)}{\mathsf{Pr}_{\Pi^\Psi}(q_i^\star|x,y,a,b)}\\
&\quad= \frac{1}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\sum_{(a,b):((x,a),(y,b))\in C_i}\sum_{r\sim s}\frac{\mathsf{Pr}_{\Pi^\Psi}(a,b,q_i^\star,r,s|x,y)}{\mathsf{Pr}_{\Pi^\Psi}(q_i^\star|x,y,a,b)\mathsf{Pr}_{\Pi^\Psi}(r,s|x,y)}
\end{aligned}$$

For all r,s such that $r\sim s$, $\mathsf{Pr}_{\boldsymbol{\Psi}}(r,s)=\frac{1}{|\mathsf{supp}(\boldsymbol{\Psi})|}$. Applying this to the RHS,

$$\begin{aligned}
\mathsf{Pr}_{\Theta}(D=1|x,y) &= \frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\sum_{(a,b):((x,a),(y,b))\in C_i}\sum_{r\sim s}\frac{\mathsf{Pr}_{\Pi^\Psi}(a,b,q_i^\star,r,s|x,y)}{\mathsf{Pr}_{\Pi^\Psi}(q_i^\star|x,y,a,b)}\\
&= \frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\sum_{(a,b):((x,a),(y,b))\in C_i}\frac{\mathsf{Pr}_{\Pi^\Psi}(a,b,q_i^\star|x,y)}{\mathsf{Pr}_{\Pi^\Psi}(q_i^\star|x,y,a,b)}\\
&= \frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{i\in[k]}\sum_{(a,b):((x,a),(y,b))\in C_i}\mathsf{Pr}_{\Pi^\Psi}(a,b|x,y).
\end{aligned}$$

Since $\mathsf{Pr}_{\Pi^\Psi}(a,b|x,y)=\mathsf{Pr}_f(a,b|x,y)$ by perfect correctness, and

$$\bigcup_{i\in[k]}\{(a,b):((x,a),(y,b))\in C_i\}=\{(a,b):\mathsf{Pr}_f(a,b|x,y)>0\},$$

we get,

$$\mathsf{Pr}_{\Theta}(D=1|x,y)=\frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}\sum_{(a,b)}\mathsf{Pr}_f(a,b|x,y)=\frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum\limits_{t\in[k]}\sqrt{c_t}\right)^2}.$$

This proves the claim. □

Claim 3. *The reduction Θ is perfectly correct; i.e.,*

$$\mathsf{Pr}_{\Theta}(a,b|D=1,x,y)=\mathsf{Pr}_f(a,b|x,y)$$

Proof. Consider (x, y, a, b) such that $\mathsf{Pr}_f(a, b|x, y) > 0$. If $((x, a), (y, b)) \in C_\ell$, by Claim 1, if $i \neq \ell$ or $j \neq \ell$, $\mathsf{Pr}_\Theta(a, b, D = 1|x, y, i, j) = 0$. Hence,

$$\begin{aligned}\mathsf{Pr}_\Theta(a, b, D = 1|x, y) &= \sum_{i,j\in[k]} \mathsf{Pr}_\Theta(a, b, D = 1, i, j|x, y)\\ &= \mathsf{Pr}_\Theta(I = J = \ell|x, y) \cdot \mathsf{Pr}_\Theta(a, b, D = 1|x, y, I = J = \ell)\\ &= \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \cdot \mathsf{Pr}_\Theta(a, b, D = 1|x, y, I = J = \ell).\end{aligned}$$

Expanding this, we get

$$\begin{aligned}\mathsf{Pr}_\Theta(a, b, D = 1|x, y) &= \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \sum_{r\sim s} \mathsf{Pr}_\Theta(r, s|x, y, I = J = \ell) \cdot \mathsf{Pr}_\Theta(a, b|x, y, r, s, I = J = \ell)\\ &= \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \sum_{r\sim s} \frac{\rho(x, r, q^\star_\ell)\sigma(y, s, q^\star_\ell)}{\rho^\dagger_\ell \sigma^\dagger_\ell} \cdot \mathsf{Pr}_{\Pi\Psi}(a, b|x, y, q^\star_\ell, r, s)\\ &= \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \sum_{r\sim s} \frac{\mathsf{Pr}_{\Pi\Psi}(q^\star_\ell|x, y, r, s)}{\rho^\dagger_\ell \sigma^\dagger_\ell} \cdot \mathsf{Pr}_{\Pi\Psi}(a, b|x, y, q^\star_\ell, r, s)\\ &= \sum_{r\sim s} \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \frac{\mathsf{Pr}_{\Pi\Psi}(a, b, q^\star_\ell|x, y, r, s)}{\rho^\dagger_\ell \sigma^\dagger_\ell}.\end{aligned}$$

Since $\mathsf{Pr}_{\boldsymbol{\Psi}}(r, s) = \frac{1}{|\mathsf{supp}(\boldsymbol{\Psi})|}$, multiplying and dividing each term with $\mathsf{Pr}_{\boldsymbol{\Psi}}(r, s)$, and expanding $\mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell)$,

$$\begin{aligned}\mathsf{Pr}_\Theta(a, b, D = 1|x, y) &= |\mathsf{supp}(\boldsymbol{\Psi})| \cdot \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \cdot \frac{\sum_{r\sim s} \mathsf{Pr}_{\Pi\Psi}(a, b, r, s, q^\star_\ell|x, y)}{\rho^\dagger_\ell \sigma^\dagger_\ell}\\ &= \frac{|\mathsf{supp}(\boldsymbol{\Psi})| \cdot \rho^\dagger_\ell \sigma^\dagger_\ell}{\mathsf{Pr}_{\Pi\Psi}(q^\star_\ell|C_i)\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \cdot \frac{\mathsf{Pr}_{\Pi\Psi}(a, b, q^\star_\ell|x, y)}{\rho^\dagger_\ell \sigma^\dagger_\ell}\\ &= \frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \frac{\mathsf{Pr}_{\Pi\Psi}(a, b|x, y) \cdot \mathsf{Pr}_{\Pi\Psi}(q^\star_\ell|x, y, a, b)}{\mathsf{Pr}_{\Pi\Psi}(q^\star_\ell|C_i)}\\ &= \frac{|\mathsf{supp}(\boldsymbol{\Psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \mathsf{Pr}_f(a, b|x, y). \qquad (23)\end{aligned}$$

The final equality follows from the fact that $\mathsf{Pr}_{\Pi^\Psi}(q_\ell^\star|C_\ell) = \mathsf{Pr}_{\Pi^\Psi}(q_\ell^\star|x,y,a,b)$ since $((x,y),(y,b)) \in C_\ell$. Hence,

$$\mathsf{Pr}_\Theta(a,b|D=1,x,y) = \frac{\mathsf{Pr}_\Theta(a,b,D=1|x,y)}{\mathsf{Pr}_\Theta(D=1|x,y)} \overset{(a)}{=} \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2}\mathsf{Pr}_f(a,b|x,y)\cdot\frac{\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2}{|\mathsf{supp}(\boldsymbol{\psi})|} = \mathsf{Pr}_f(a,b|x,y), \tag{24}$$

where (a) follows from Claim 2. If $\mathsf{Pr}_f(a,b|x,y)=0$, then, by Claim 1, $\mathsf{Pr}_{\Pi^\Psi}(a,b,D=1|x,y)=0$, and hence $\mathsf{Pr}_\Theta(a,b|D=1,x,y)=0$. This concludes the proof. □

Proof of security. To prove the security of Θ, we need to show that there exists simulators $S'_A : \mathcal{X}\times\mathcal{A}\times\{0,1\} \to (\mathcal{R}\cup\{\perp\})\times[k]$ and $S'_B : \mathcal{Y}\times\mathcal{B}\times\{0,1\} \to (\mathcal{S}\cup\{\perp\})\times[k]$ such that if $\mathsf{Pr}_f(a,b|x,y)>0$,

$$\mathsf{Pr}_\Theta(r,i|x,y,a,b,D=1) = \mathsf{Pr}_{S'_A}(r,i|x,a,D=1),$$
$$\mathsf{Pr}_\Theta(s,j|x,y,a,b,D=1) = \mathsf{Pr}_{S'_B}(s,j|y,b,D=1),$$

and,

$$\mathsf{Pr}_\Theta(r,i|x,y,D=0) = \sum_a \mathsf{Pr}_{f_A}(a|x,y)\cdot\mathsf{Pr}_{S'_A}(r,i|x,a,D=0),$$
$$\mathsf{Pr}_\Theta(s,j|x,y,D=0) = \sum_b \mathsf{Pr}_{f_B}(b|x,y)\cdot\mathsf{Pr}_{S'_B}(s,j|y,b,D=0).$$

We prove the first two statements in Claim 4 and the last two in Claim 5.

Claim 4. *There exists a randomized function $S'_A : \mathcal{X}\times\mathcal{A}\times\{0,1\}\to\mathcal{U}\times[k]$ such that, if $\mathsf{Pr}_f(a,b|x,y)>0$,*

$$\mathsf{Pr}_\Theta(r,i|x,y,a,b,D=1) = \mathsf{Pr}_{S'_A}(r,i|x,a,D=1).$$

Similarly, there exists a randomized function $S'_B : \mathcal{Y}\times\mathcal{B}\times\{0,1\}\to\mathcal{V}\times[k]$ such that, if $\mathsf{Pr}_f(a,b|x,y)>0$,

$$\mathsf{Pr}_\Theta(s,j|x,y,a,b,D=1) = \mathsf{Pr}_{S'_B}(s,j|y,b,D=1).$$

Proof. Consider (x,y,a,b) such that $\mathsf{Pr}_f(a,b|x,y)>0$; let $((x,a),(y,b))\in C_\ell$. By Claim 1,

$$\mathsf{Pr}_\Theta(r,i,a,b,D=1|x,y)=0 \text{ if } i\neq\ell \text{ or } r=\perp. \tag{25}$$

We focus on $\mathsf{Pr}_{\Theta}(r, i, a, b, D = 1|x, y)$, when $r \neq \perp$ and $i = \ell$. Noting that $\mathsf{Pr}_{\Theta}(D = 1, I \neq J|x, y) = 0$,

$$\begin{aligned}
&\mathsf{Pr}_{\Theta}(r, I = \ell, a, b, D = 1|x, y) \\
&\quad = \sum_{j\in[k]} \sum_{s:r\sim s} \mathsf{Pr}_{\Theta}(r, s, I = \ell, j, a, b|x, y) \\
&\quad = \mathsf{Pr}_{\Theta}(I = J = \ell|x, y) \sum_{s:r\sim s} \mathsf{Pr}_{\Theta}(r, s|x, y, I = J = \ell) \cdot \mathsf{Pr}_{\Theta}(a, b|x, y, r, s, I = J = \ell) \\
&\quad = \mathsf{Pr}^2_{\boldsymbol{\lambda}}(\ell) \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|x, y, r, s)}{\rho^{\dagger}_i \sigma^{\dagger}_i} \mathsf{Pr}_{\Theta}(a, b|x, y, r, s, I = J = \ell).
\end{aligned}$$

We have,

$$\mathsf{Pr}_{\Theta}(a, b|x, y, r, s, I = J = \ell) = \mathsf{Pr}_{\Pi^{\mathrm{out}}_A}(a|x, r, q^{\star}_{\ell}) \cdot \mathsf{Pr}_{\Pi^{\mathrm{out}}_B}(b|y, s, q^{\star}_{\ell}) = \mathsf{Pr}_{\Pi^{\Psi}}(a, b|x, y, r, s, q^{\star}_{\ell}).$$

Substituting for $\mathsf{Pr}_{\boldsymbol{\lambda}}(\ell)$ from Definition 6 and noting that $\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|C_{\ell}) = \mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|x, y, a, b)$ since $((x, a), (y, b)) \in C_{\ell}$,

$$\begin{aligned}
\mathsf{Pr}_{\Theta}(r, I = \ell, a, b, D = 1|x, y) &= \sum_{s:r\sim s} \frac{\rho^{\dagger}_i \sigma^{\dagger}_i}{\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|C_{\ell}) \left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(a, b, q^{\star}_{\ell}|x, y, r, s)}{\rho^{\dagger}_i \sigma^{\dagger}_i} \\
&= \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(a, b, q^{\star}_{\ell}|x, y, r, s)}{\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|x, y, a, b) \left(\sum_{t\in[k]} \sqrt{c_t}\right)^2}.
\end{aligned}$$

Since $\mathsf{Pr}_{\boldsymbol{\psi}}(r, s) = \frac{1}{|\mathsf{supp}(\boldsymbol{\psi})|}$, multiplying and dividing each term with $\mathsf{Pr}_{\boldsymbol{\psi}}(r, s)$,

$$\begin{aligned}
\mathsf{Pr}_{\Theta}(r, I = \ell, a, b, D = 1|x, y) &= \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\boldsymbol{\psi}}(r, s) \cdot \mathsf{Pr}_{\Pi^{\Psi}}(a, b, q^{\star}_{\ell}|x, y, r, s)}{\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|x, y, a, b)} \\
&= \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(a, b, q^{\star}_{\ell}, r|x, y)}{\mathsf{Pr}_{\Pi^{\Psi}}(q^{\star}_{\ell}|x, y, a, b)} \\
&= \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \mathsf{Pr}_{\Pi^{\Psi}}(a, b|x, y) \mathsf{Pr}_{\Pi^{\Psi}}(r|x, y, a, b, q^{\star}_{\ell}) \\
&= \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]} \sqrt{c_t}\right)^2} \mathsf{Pr}_f(a, b|x, y) \mathsf{Pr}_{\Pi^{\Psi}}(r|x, y, a, b, q^{\star}_{\ell}).
\end{aligned}$$

Hence, by (23),

$$\mathsf{Pr}_\Theta(r, I = \ell | a, b, x, y, D = 1) = \frac{\mathsf{Pr}_\Theta(r, I = \ell, a, b, D = 1 | x, y)}{\mathsf{Pr}_\Theta(a, b, D = 1 | x, y)} = \mathsf{Pr}_{\Pi^\Psi}(r | x, y, a, b, q_\ell^\star).$$

By perfect privacy of Π^Ψ, there exists a simulator $\hat{S}_A$ such that

$$\mathsf{Pr}_\Theta(r, I = \ell | a, b, x, y, D = 1) = \mathsf{Pr}_{\Pi^\Psi}(r | x, y, a, b, q_\ell^\star) = \mathsf{Pr}_{\hat{S}_A}(r | x, a, q_\ell^\star).$$

Since C_ℓ is determined by (x, a), we can set $\mathsf{Pr}_{S'_A}(x, a, 1) = \mathsf{Pr}_{\hat{S}_A}(r | x, a, q_\ell^\star)$. The first statement in the claim follows from this observation and (25). The second statement can be proved analogously. □

Claim 5. *There exists a randomized function* $S'_A : \mathcal{X} \times \mathcal{A} \times \{0, 1\} \to (\mathcal{R} \times \{\perp\}) \times [k]$ *such that*

$$\mathsf{Pr}_\Theta(r, i | x, y, D = 0) = \sum_a \mathsf{Pr}_{f_A}(a | x, y) \cdot \mathsf{Pr}_{S'_A}(r, i | x, a, D = 0).$$

Similarly, there exists a randomized function $S'_B : \mathcal{Y} \times \mathcal{B} \times \{0, 1\} \to \mathcal{S} \times [k]$ *such that*

$$\mathsf{Pr}_\Theta(s, j | x, y, D = 0) = \sum_b \mathsf{Pr}_{f_B}(b | x, y) \cdot \mathsf{Pr}_{S'_B}(s, j | y, b, D = 0).$$

Proof. When $r = \perp$, the predicate always rejects $(D = 0)$, hence, for all i,

$$\mathsf{Pr}_\Theta(R = \perp, i, D = 0 | x, y) = \mathsf{Pr}_\Theta(R = \perp, i | x, y) = \mathsf{Pr}_\Theta(R = \perp, i | x).$$

The predicate accepts $(D = 1)$ if and only if Alice and Bob choose i, j and r, s, respectively, such that $i = j$ and $r \sim s$. Hence,

$$\begin{aligned}
&\mathsf{Pr}_\Theta(r, i, D = 0 | x, y) \\
&\quad = \mathsf{Pr}_\Theta(r, i | x, y) - \mathsf{Pr}_\Theta(i, J = i | x, y) \sum_{s : r \sim s} \mathsf{Pr}_\Theta(r, s | x, y, i, J = i) \\
&\quad = \mathsf{Pr}_\Theta(r, i | x, y) - \mathsf{Pr}_\Theta(i, J = i | x, y) \sum_{s : r \sim s} \frac{\rho(x, r, q_i^\star) \cdot \sigma(y, s, q_i^\star)}{\rho_i^\dagger \sigma_i^\dagger} \\
&\quad = \mathsf{Pr}_\Theta(r, i | x, y) - \mathsf{Pr}_{\boldsymbol{\lambda}}^2(i) \sum_{s : r \sim s} \frac{\mathsf{Pr}_{\Pi^\Psi}(q_i^\star | x, y, r, s)}{\rho_i^\dagger \sigma_i^\dagger}.
\end{aligned}$$

We focus on the second term in the RHS. Expanding $\mathsf{Pr}^2_{\boldsymbol{\lambda}}(i)$ using Definition 6,

$$\begin{aligned}
\mathsf{Pr}^2_{\boldsymbol{\lambda}}(i) \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|x,y,r,s)}{\rho_i^\dagger \sigma_i^\dagger} &= \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|x,y,r,s)}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \\
&= \frac{1}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\cdot\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \sum_{s:r\sim s} \frac{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star,r,s|x,y)}{\mathsf{Pr}_{\Pi^{\Psi}}(r,s|x,y)} \\
&= \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\cdot\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star,r|x,y)
\end{aligned}$$

The last equality used the fact that $\mathsf{Pr}_{\boldsymbol{\psi}}(r,s) = \frac{1}{|\mathsf{supp}(\boldsymbol{\psi})|}$ for all $r \sim s$. Thus, when $\hat{S}_A$ is the simulator for Alice that witnesses the perfect security of Π^{Ψ},

$$\begin{aligned}
&\mathsf{Pr}_{\Theta}(r,i,D=0|x,y) \\
&= \mathsf{Pr}_{\Theta}(r,i|x) - \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\cdot\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \sum_a \mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star,r|x,y,a)\cdot\mathsf{Pr}_{\Pi^{\Psi}}(a|x,y) \\
&\overset{(a)}{=} \mathsf{Pr}_{\Theta}(r,i|x) - \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\cdot\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \sum_a \mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star,r|x,y,a)\cdot\mathsf{Pr}_{f_A}(a|x,y) \\
&\overset{(b)}{=} \mathsf{Pr}_{\Theta}(r,i|x) - \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\mathsf{Pr}_{\Pi^{\Psi}}(q_i^\star|C_i)\cdot\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2} \sum_a \mathsf{Pr}_{\hat{S}_A}(q_i^\star,r|x,a)\cdot\mathsf{Pr}_{f_A}(a|x,y).
\end{aligned}$$

Here, (a) and (b) follow from the perfect correctness and perfect security against Alice, respectively. The first statement of the claim now follows from the fact that $\mathsf{Pr}_{\Theta}(D=0|x,y)$ is the same non-zero value for all x,y as established in Claim 2 The corresponding statement for Bob (second statement) can be shown analogously. □

We conclude the proof of security by noting that the properties in Claim 4 and Claim 5 can be satisfied by the same S'_A and S'_B.

Bound on Accept Probability. It remains to upper bound the probability with which the predicate accepts $(D=1)$ for all inputs x,y.

Claim 6. *The protocol Θ accepts with probability $2^{-\mu}$ where $\mu \leq \log \frac{|\mathcal{R}||\mathcal{S}||\mathcal{X}|^2|\mathcal{Y}|^2|\mathcal{A}||\mathcal{B}|}{|\mathsf{supp}(\boldsymbol{\psi})|}$.*

Proof. By Claim 2, $\mathsf{Pr}_\Theta(D = 1|x,y) = \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{\left(\sum_{t\in[k]}\sqrt{c_t}\right)^2}$. By Definition 6, $c_t = \frac{\rho_t^\dagger \sigma_t^\dagger}{\mathsf{Pr}_\Pi(q_t^\star|C_t)}$ and $q_t^\star$ is chosen such that $\rho_t^\dagger\sigma_t^\dagger \le |\mathcal{R}||\mathcal{S}||\mathcal{X}_{C_t}||\mathcal{Y}_{C_t}|\mathsf{Pr}_{\Pi^\Psi}(q_t^\star|C_t)$ and $\mathsf{Pr}_{\Pi^\Psi}(q_t^\star) > 0$, hence, $c_t \le |\mathcal{R}||\mathcal{S}||\mathcal{X}_C||\mathcal{Y}_C|$. Using Cauchy-Schwartz,

$$\sum_{t\in[k]}\sqrt{c_t} \le \sqrt{\sum_{t\in[k]} c_t}\cdot\sqrt{k} \le \sqrt{k\cdot|\mathcal{R}||\mathcal{S}|\sum_{t\in[k]}|\mathcal{X}_{C_t}||\mathcal{Y}_{C_t}|} \le \sqrt{|\mathcal{R}||\mathcal{S}||\mathcal{A}||\mathcal{B}|}|\mathcal{X}||\mathcal{Y}|$$

The final inequality used the fact that each (x,a) shows up in at most one of the connected components; hence, $k \le \sqrt{|\mathcal{X}||\mathcal{Y}||\mathcal{A}||\mathcal{B}|}$ and $\sum_{t\in[k]}|\mathcal{X}_{C_t}||\mathcal{Y}_{C_t}| \le k\cdot|\mathcal{X}_C||\mathcal{Y}_C| \le |\mathcal{X}||\mathcal{Y}|\sqrt{|\mathcal{A}||\mathcal{B}|}$.

$$\mathsf{Pr}_\Theta(D=1|x,y) \ge \frac{|\mathsf{supp}(\boldsymbol{\psi})|}{|\mathcal{R}||\mathcal{S}||\mathcal{X}|^2|\mathcal{Y}|^2|\mathcal{A}||\mathcal{B}|} \Rightarrow \mu \le \log\frac{|\mathcal{R}||\mathcal{S}||\mathcal{X}|^2|\mathcal{Y}|^2|\mathcal{A}||\mathcal{B}|}{|\mathsf{supp}(\boldsymbol{\psi})|}.$$

□

Corollary 1. *Consider a randomised function $f : \mathcal{X}\times\mathcal{Y} \to \mathcal{A}\times\mathcal{B}$ with k connected components in its evaluation graph G_f. If a protocol Π^{OT^ℓ} using ℓ copies of OT correlation computes f with perfect security, then there exists a μ-SZCR to $\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{\ell+\lceil\log k\rceil+1})}$ such that $\mu \le \log\frac{|\mathcal{R}||\mathcal{S}||\mathcal{X}|^2|\mathcal{Y}|^2|\mathcal{A}||\mathcal{B}|}{|\mathsf{supp}(\boldsymbol{\psi})|}$.*

Proof. By Theorem 3, f has a μ-SZCR to $\boldsymbol{\Phi}_{(=^{\lceil\log k\rceil})}\cdot\boldsymbol{\Phi}_{\mathsf{supp}^*(\mathsf{OT}^m)}$. But, $\boldsymbol{\Phi}_{(=^{\lceil\log k\rceil})}$ can be realized using $\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{\lceil\log k\rceil})}$ (since 1-bit equality can be checked with 1 OT) and $\boldsymbol{\Phi}_{\mathsf{supp}^*(\mathsf{OT}^m)}$ can be realized using $\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{m+1})}$ (by encoding the input symbol $\perp$ in $\boldsymbol{\Phi}_{\mathsf{supp}^*(\mathsf{OT}^m)}$ using an extra OT). Consequently, the predicate $\boldsymbol{\Phi}_{(=^{\lceil\log k\rceil})}\cdot\boldsymbol{\Phi}_{\mathsf{supp}^*(\mathsf{OT}^m)}$ can be realized using $\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{\lceil\log k\rceil})}\cdot\boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{m+1})} = \boldsymbol{\Phi}_{\mathsf{supp}(\mathsf{OT}^{\lceil\log k\rceil+m+1})}$. This implies the corollary. □

References

1. Pratyush Agarwal, Varun Narayanan, Shreya Pathak, Manoj Prabhakaran, Vinod M. Prabhakaran, and Mohammad Ali Rehan. Secure non-interactive reduction and spectral analysis of correlations. In *EUROCRYPT 2022*, pages 797–827, 2022.
2. Hamidreza Amini Khorasgani, Hemanta K. Maji, and Hai H. Nguyen. Secure non-interactive simulation: Feasibility and rate. In *EUROCRYPT 2022*, pages 767–796, 2022.
3. Donald Beaver. Perfect privacy for two-party protocols. In Joan Feigenbaum and Michael Merritt, editors, *Proceedings of DIMACS Workshop on Distributed Computing and Cryptography*, volume 2, pages 65–77. American Mathematical Society, 1989.
4. Donald Beaver. Correlated pseudorandomness and the complexity of private computations. In *STOC*, pages 479–488, 1996.
5. Amos Beimel, Yuval Ishai, Ranjit Kumaresan, and Eyal Kushilevitz. On the cryptographic complexity of the worst functions. In *Theory of Cryptography*, pages 317–342, 2014.

6. Amos Beimel and Tal Malkin. A quantitative approach to reductions in secure computation. In *TCC*, pages 238–257, 2004.
7. Kaartik Bhushan, Ankit Kumar Misra, Varun Narayanan, and Manoj Prabhakaran. Secure non-interactive reducibility is decidable. In these proceedings, 2022.
8. Benny Chor and Eyal Kushilevitz. A zero-one law for boolean privacy. *SIAM J. Discrete Math.*, 4(1):36–47, 1991.
9. Imre Csiszár and Rudolf Ahlswede. On oblivious transfer capacity. In *International Symposium on Information Theory (ISIT)*, pages 2061–2064, 2007.
10. Yevgeniy Dodis and Silvio Micali. Lower bounds for oblivious transfer reductions. In *EUROCRYPT*, pages 42–55, 1999.
11. Zeev Dvir and Sivakanth Gopi. 2-server PIR with subpolynomial communication. *J. ACM*, 63(4):39:1–39:15, 2016.
12. Oded Goldreich. *Foundations of Cryptography: Basic Applications*. Cambridge University Press, 2004.
13. Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play ANY mental game. In ACM, editor, *STOC*, pages 218–229, 1987. See [12, Chap. 7] for more details.
14. Oded Goldreich and Ronen Vainish. How to solve any protocol problem - an efficiency improvement. In *CRYPTO*, pages 73–86, 1987.
15. Stuart Haber and Silvio Micali. Unpublished manuscript cited by [12], 1986.
16. Hideki Imai, Kirill Morozov, and Anderson C. A. Nascimento. On the oblivious transfer capacity of the erasure channel. In *International Symposium on Information Theory (ISIT)*, pages 1428–1431, 2006.
17. Hideki Imai, Kirill Morozov, and Anderson C. A. Nascimento. Efficient oblivious transfer protocols achieving a non-zero rate from any non-trivial noisy correlation. In *International Conference on Information Theoretic Security (ICITS)*, 2007.
18. Hideki Imai, Kirill Morozov, Anderson C. A. Nascimento, and Andreas Winter. Efficient protocols achieving the commitment capacity of noisy correlations. In *International Symposium on Information Theory (ISIT)*, pages 1432–1436, 2006.
19. Hideki Imai, Jörn Müller-Quade, Anderson C. A. Nascimento, and Andreas Winter. Rates for bit commitment and coin tossing from noisy correlation. In *International Symposium on Information Theory (ISIT)*, pages 45–, 2004.
20. Joe Kilian. Founding cryptography on oblivious transfer. In *STOC*, pages 20–31, 1988.
21. Eyal Kushilevitz. Privacy and communication complexity. In *FOCS*, pages 416–421, 1989.
22. Hemanta Maji, Manoj Prabhakaran, and Mike Rosulek. *Complexity of Multi-Party Computation Functionalities*, volume 10 of *Cryptology and Information Security Series*, pages 249 – 283. IOS Press, Amsterdam, 2013.
23. Varun Narayanan, Manoj Prabhakaran, and Vinod M. Prabhakaran. Zero-communication reductions. In Rafael Pass and Krzysztof Pietrzak, editors, *Theory of Cryptography - 18th International Conference, TCC 2020, Durham, NC, USA, November 16-19, 2020, Proceedings, Part III*, volume 12552 of *Lecture Notes in Computer Science*, pages 274–304. Springer, 2020.
24. Vinod Prabhakaran and Manoj Prabhakaran. Assisted common information with an application to secure two-party sampling. *IEEE Transactions on Information Theory*, 60(6):3413–3434, 2014. doi:10.1109/TIT.2014.2316011.
25. Severin Winkler and Jürg Wullschleger. Statistical impossibility results for oblivious transfer reductions. Cryptology ePrint Archive, Report 2009/508, 2009. http://eprint.iacr.org/.

26. Andreas Winter, Anderson C. A. Nascimento, and Hideki Imai. Commitment capacity of discrete memoryless channels. In *IMA Int. Conf.*, pages 35–51, 2003.
27. Stefan Wolf and Jürg Wullschleger. New monotones and lower bounds in unconditional two-party computation. In *CRYPTO*, pages 467–477, 2005.

A Basic Constructions

In this section, for the sake of explicitness, we detail two basic constructions of Balanced Embedding from any function to the OT predicate – from a truth table and from a boolean circuit of the function. The first construction is implied by the second one, which in turn is implied by the general construction of balanced embedding from SZCR.

A.1 Balanced Embedding from Truth Table

Theorem 4. *For any deterministic function* $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\} \times \{0,1\}$*, there exists a balanced embedding to* $\boldsymbol{\Phi}_{\mathsf{OT}}^k$ *for* $k = 2^{n+1}$.

Proof. To define the balanced embedding (π, θ) we will define inputs u_α and v_β to $\boldsymbol{\Phi}_{\mathsf{OT}}^k$ such that $\pi(u,\alpha) = \theta(u,\alpha) = 1$ for $u = u_\alpha$ and 0 for rest; and similarly $\pi(v,\beta) = \theta(v,\beta) = 1$ for $v = v_\beta$ and 0 for rest. u_α and v_β where $\alpha = (x,a)$ and $\beta = (y,b)$ are defined as follows:

- For $0 \le i \le 2^n - 1$, $u_i = (1,a)$, if $i = x$ and $u_i = (0,0)$ otherwise, whereas $v_i = (0, f_A(i,y))$.
- For $2^n \le i \le 2^{n+1} - 1$, $v_i = (1,b)$, if $i = 2^n + y$ and $v_i = (0,0)$ otherwise, whereas $u_i = (0, f_B(x,i))$.

It is straight forward to see that this definition satisfies the conditions of a balanced embedding as the only compatible u, v pairs correspond to correct outputs being sampled at both the ends. □

A.2 Constructing balanced embedding from circuit

Theorem 5. *Given a circuit* C *with NAND gates that computes a function* f*, we can construct a balanced embedding to* $\boldsymbol{\Phi}_{\mathsf{OT}}^{2|C|}$.

Proof. Let x and y be the inputs of Alice and Bob, respectively. For each wire w in C, Alice and Bob sample w_A and w_B, respectively, as follows:

(i). If w is an input wire that reads x_i, then $w_A = x_i$ and $w_B = 0$, and if w is an input wire that reads y_i, then $w_A = 0$ and $w_B = y_i$
(ii). If w is the output wire computing $f_A(x,y)$, then $w_A \leftarrow \{0,1\}$ and $w_B = 0$, and if w is the output wire computing $f_B(x,y)$, then $w_A = 0$, and $w_B \leftarrow \{0,1\}$.
(iii). Otherwise, $w_A \leftarrow \{0,1\}$ and $w_B \leftarrow \{0,1\}$.

For each gate g in C, we denote the two input wires by $\mathrm{In1}^g, \mathrm{In2}^g$ and the output wire by Out^g.

We define sets U_x and V_y corresponding to inputs x, y. Elements of these sets $(u_i \in \{0,1\}^2 : 1 \le i \le 2|C|)$ and $(v_i \in \{0,1\}^2 : 1 \le i \le 2|C|)$ are be sampled as follows:
Enumerate the gates in C as $g_1, g_2, \ldots, g_{|C|}$; for $1 \le i \le 2|C|$:

- Set $u_{2i-1} = (\alpha_A^{g_i}, \mathrm{In1}_A^{g_i} \oplus \alpha_A^{g_i})$ and $u_{2i} = (\beta_A^{g_i}, \mathrm{In2}_A^{g_i} \oplus \beta_A^{g_i})$, where $\alpha_A^{g_i}, \beta_A^{g_i}$ are sampled uniformly at random subject to:

$$\alpha_A^{g_i} \oplus \beta_A^{g_i} = \mathrm{Out}_A^{g_i} \oplus (\mathrm{In1}_A^{g_i} \cdot \mathrm{In1}_A^{g_i}) \oplus 1. \tag{26}$$

- Sets $v_{2i-1} = (\mathrm{In2}_B^{g_i}, \alpha_B^{g_i})$ and $u_{2i} = (\mathrm{In1}_B^{g_i}, \beta^{g_i})$, where $\alpha_B^{g_i}, \beta_B^{g_i}$ are sampled uniformly at random subject to:

$$\alpha_B^{g_i} \oplus \beta_B^{g_i} = \mathrm{Out}_B^{g_i} \oplus (\mathrm{In1}_B^{g_i} \cdot \mathrm{In1}_B^{g_i}). \tag{27}$$

Finally, set candidate outputs $a = \hat{w}_B$, where $\hat{w}$ is the wire that outputs $f_A(x, y)$ in C, and $b = \tilde{w}_A$, where $\tilde{w}$ is the wire that outputs $f_B(x, y)$ in C. We use functions $O_A : \mathcal{U} \times X \to \{0,1\}$ and $O_B : \mathcal{V} \times Y \to \{0,1\}$ to denote the a and b values generated for specific u, x and y, b pairs respectively.

We then define the embedding (π, θ) for $\alpha = (x, a)$ and $\beta = (y, b)$ as $\pi(u, \alpha) = \theta(u, \alpha) = 2^{-|C|}$ if $u \in U_x$ and $a = O_A(u, x)$ and 0 otherwise. Similarly, $\pi(v, \beta) = \theta(v, \beta) = 2^{-|C|}$ if $v \in V_y$ and $b = O_B(v, y)$ and 0 otherwise. It is easy to check that this construction is indeed correct, owing to the correctness of the circuit C. □