One-Message Secure Reductions: On the Cost of Converting Correlations

Yuval Ishai[*] Mahimna Kelkar[†] Varun Narayanan[‡] Liav Zafar[*]

Abstract

Correlated secret randomness is a useful resource for secure computation protocols, often enabling dramatic speedups compared to protocols in the plain model. This has motivated a line of work on identifying and securely generating useful correlations.

Different kinds of correlations can vary greatly in terms of usefulness and ease of generation. While there has been major progress on efficiently generating *oblivious transfer* (OT) correlations, other useful kinds of correlations are much more costly to generate. Thus, it is highly desirable to develop efficient techniques for securely *converting* copies of a given source correlation into copies of a given target correlation, especially when the former are cheaper to generate than the latter.

In this work, we initiate a systematic study of such conversions that only involve a single uni-directional message. We refer to such a conversion as a *one-message secure reduction* (OMSR). Recent works (Agarwal et al, Eurocrypt 2022; Khorasgani et al, Eurocrypt 2022) studied a similar problem when no communication is allowed; this setting is quite restrictive, however, with few non-trivial conversions being feasible. The OMSR setting substantially expands the scope of feasible results, allowing for direct applications to existing MPC protocols.

We obtain the following positive and negative results.

- **OMSR constructions.** We present a general rejection-sampling based technique for OMSR with OT source correlations. We apply it to substantially improve in the communication complexity of optimized protocols for distributed symmetric cryptography (Dinur et al., Crypto 2021).
- **OMSR lower bounds.** We develop general techniques for proving lower bounds on the communication complexity of OMSR, matching our positive results up to small constant factors.

A preliminary version of this paper appears in the proceedings of CRYPTO 2023. This is the full version.

[*] Technion. {yuvali, liavz}@cs.technion.ac.il

[†] Cornell University. mahimna@cs.cornell.edu

[‡] UCLA. varunnkv@gmail.com

Contents

1 Introduction

Secure multiparty computation [30, 48] (MPC) is a fundamental cryptographic primitive that enables mutually distrusting parties to collaboratively compute a function over their combined data while keeping their local data secret. While MPC is a general tool, it can be quite heavyweight in terms of both computation and communication compared to a non-secure evaluation. To minimize this cost, a common paradigm is to use preprocessing in the following way. Before the inputs are known, the parties run an *offline protocol* to generate some input-independent local information. The latter then serves as a resource for speeding up the *online protocol*, which is executed once the inputs are known. A qualitative advantage of this paradigm is that expensive cryptographic operations can be pushed to the offline phase, resulting in a simple and lightweight online protocol.

Abstractly, in the offline phase, the parties securely generate instances of *correlated randomness* or *correlations* that are independent of the protocol inputs, and can therefore be processed in advance. Examples for standard correlations include *oblivious transfer* (OT) correlations, which serve as a natural basis for MPC protocols for Boolean circuits [30, 34, 38, 48], and *multiplication triples* [6, 7, 21], which serve as a basis for MPC for arithmetic circuits.

While the above standard correlations are universal, in the sense that they suffice for any online computation, in many cases it is more efficient to use a specially crafted correlation geared towards the particular function being evaluated. Moreover, while recent techniques support generation of n (pseudorandom) copies of any correlation with $o(n)$ communication cost [15, 16], they are concretely efficient only for a few standard correlations, and therefore are not practical for most other useful correlations.

This motivates the central theme of our work: the study of efficiently and securely *deriving one type of correlation (the target) from another (the source)*; of particular relevance is when the source correlation can be generated very cheaply using known techniques. As an upshot, this allows us to broaden the class of correlations which can be concretely efficiently generated. We will restrict our attention to two-party protocols in this paper, and focus mainly on target correlations that are useful for secure computation in the *semi-honest* model.

Converting between correlations. A recent line of work [1, 31, 36, 37, 39] introduces secure "non-interactive" reductions/simulations (SNIR/SNIS) for securely converting one type of correlation to another. Here "non-interactive" refers to the strict notion of having *no communication at all.* This can be viewed as a secure analog of (non-secure) non-interactive simulations of joint distributions (NIS), which have been extensively studied in the information theory literature (see [42] for a recent survey). The SNIR/SNIS model, however, is highly restrictive. As intuition perhaps suggests, very few conversions between correlations are feasible, and many conversions are provably impossible. In fact, the model remains highly restrictive even when the security condition is dropped. As a result, there are no examples for nontrivial applications of SNIR/SNIS towards generating useful correlations for MPC.

One-message secure reductions. Motivated by these limitations of the zero-interaction setting, we take the next natural step and study the same secure conversion problem in a setting where only one-way communication is allowed. In particular, denoting the two parties by the sender S and the receiver R, we allow for a single message to be sent from S to R. We refer to a conversion protocol in this model as a *one-message secure reduction* (OMSR). This relaxation of SNIR/SNIS dramatically changes the landscape, since in the OMSR setting many source correlations are *universal* in the sense that they can be converted into every target correlation [27]. Such source correlations include the string-OT correlation, which can be generated cheaply and "silently" (i.e., locally) using recent techniques [16]. This directly confirms the wide applicability of OMSR for generating correlations relevant for MPC, but leaves open the asymptotic and concrete *efficiency* of such reductions.

Limiting communication to be one-directional also comes with its own qualitative advantages which have motivated widely-studied models like non-interactive zero-knowledge (NIZK) [9], wire-tap coding [45] and one-way secure computation (OWSC) [27]. Our OMSR model inherits these advantages. Finally, the simplicity of the OMSR model makes it more tractable for analysis and lower bound proofs than the fully interactive setting, while still permitting conversions that can make existing MPC protocols more efficient.

Non-trivial OMSR constructions for generating specific types of correlations can be implicitly found in a recent work by Dinur et al. [24], although there was no concrete objective to restrict to only single message protocols. The OMSRs were used in the context of concretely efficient distributed protocols for MPC-friendly symmetric-key cryptography that mixed together linear functions over different small moduli. They also serve as a starting point for us; our work defines the formalism for OMSR, generalizes the constructions used by [24], as well as provides substantial improvements. This directly translates to concrete efficiency gains in a number of settings including oblivious pseudorandom function (OPRF) evaluation, the MPC-in-the-head paradigm for signatures, and the distributed generation of keys for function secret sharing (FSS) with applications to privacy-preserving machine learning.

1.1 Our Contributions

OMSR **formalism (Section 4).** We start by formalizing the notion of an OMSR, for securely converting from m copies of a source correlation (X, Y) to n copies of a target correlation (U, V). We also consider relaxed flavors of OMSR that are useful towards our positive and negative results: a "Las Vegas" variant, which allows rejection without leaking information, and OMR, which forgoes the security requirement. We primarily focus on two concrete efficiency metrics: the number of *bits of communication* from the sender to the receiver, and the number of source copies $m(n)$ required for the conversion. In practice, the latter also captures *computation* cost to generate the initial source copies.

Efficient OMSR**s from OT-correlations (Section 5).** We construct several OMSR protocols for converting from some type of OT source correlation to useful classes of target correlations; the use of OT as a source correlation is strongly motivated by a recent line of work on fast and "silent" generation of OT correlations [13, 15, 16, 19, 46].

We show OMSRs for generating two concrete correlations from OT: The first is the (t, q)-correlation (where $t < q$) which is the sharing of a random value r over both $\mathbb{Z}_t$ and $\mathbb{Z}_q$. This prepossessing is useful for the online conversion of a mod-t shared secret value to a mod-q sharing of the same value, and provides efficiency gains in protocols which work over different rings (e.g., [23]). Our second OMSR is for the $(3, 2)$-correlation which was used in [24] to convert between a mod-3 sharing of secret x to a mod-2 sharing of x mod 2.

The former generalizes $(2, 3)$-correlations which along with $(3, 2)$-correlations were instrumental within [24] in building concretely efficient distributed protocols for candidate MPC-friendly weak-PRF and PRG constructions that mixed linear operations over $\mathbb{Z}_2$ and $\mathbb{Z}_3$. Our new OMSR constructions concretely improve the communication cost by more than 2x over the (already heavily optimized) protocols from [24] (The improvement is orthogonal to the single-message feature, and applies even when comparing to protocols that use an arbitrary number of rounds.). The same kind of improvement is expected to apply to future designs of symmetric primitives based on the same alternating moduli paradigm.

Applications (Section 5.4). Our improved OMSRs translate to improvements in all of the application scenarios considered in [24]: post-quantum oblivious PRF, fully distributed MPC protocols for PRF evaluation, signatures based on the MPC-in-the-head paradigm, and distributed generation for function secret sharing (FSS) keys. The latter is particularly motivated by applications to privacy-preserving machine learning, where FSS is an increasingly popular building block for fast offline-online secure protocols for ReLU and other nonlinear activation functions [11, 17, 40, 41, 43]. The PRG candidates from [24] serve as an attractive choice for MPC-friendly PRGs in such contexts, and were recently used in the FssNN system to optimize the distributed generation of FSS keys [47]. Our new OMSR protocols for $(2, 3)$ and $(3, 2)$ correlations would lead to significant improvement in the concrete communication cost of the protocol from [47] and similar protocols. We leave an optimized implementation and benchmarking to future work.

Lower bounds (Section 6). We start by proving new lower bounds for (insecure) OMR. While the notion of OMR is meaningless in the presence of common randomness (which is cheap to generate in a cryptographic setting), we will later argue that OMR lower bounds without common randomness can be lifted to OMSR lower bounds that apply even in the presence of common randomness.

When a so-called S^* measure (see Definition 3.1) of the source correlation is strictly smaller than that of the target, Theorem 6.1 shows that the communication cost of OMR is necessarily linear. This result strictly strengthens an impossibility result for non-interactive simulation [5]. A more precise analysis yields concrete lower bounds (Corollaries 6.7.1 and 6.7.3) on the amortized communication cost of OMR between specific source and target correlations from our positive results. Theorem 6.8 shows that deriving n-bit unit vector correlations for large enough n from almost any correlation requires linear communication even *with interaction*. In fact, the result applies to other target correlation families too (See Section 6.2). This result can be thought of as a generalization of a result in [18] which is a similar result for deriving common randomness from noisy common randomness.

Role of common randomness (Section 7). Finally, in Theorem 7.1, we show that common randomness does not aid in OMSR; i.e., given an OMSR with common randomness we can derive an OMSR without common randomness with comparable error. Our proof relies on a convergence theorem for Markov chains. In contrast, with a sufficiently large amount of common randomness, any correlation can be non-securely derived without any communication.

A question that arises from the above discussion is whether OMSR is strictly harder to realize than OMR. The costs of OMSR and OMR trivially coincide for any pair of correlations that permit secure *non-interactive* reductions. In other non-trivial cases, our lower bound for both OMR and upper bound for OMSR are only tight up to a constant, hence, we do not know whether they match. However, security makes a big difference when augmenting nontrivial source correlations (such as

OT correlations) with public randomness. This trivializes the notion of OMR, but keeps our lower bounds for OMSR unchanged (Theorem 7.1). While our results imply separations between the two notions, we do not know of any explicit non-trivial instances where the two notions provably do not match, although intuition suggests that OMR is an easier primitive than OMSR to realize.

1.2 Related Work

As mentioned above, we are motivated by the question of efficiently and securely deriving one type of correlation from another. A recent line of work [1, 8, 36, 37] introduced secure non-interactive simulation/reduction which studied the problem of deriving a correlation from another without any communication and with information theoretic security. The non-secure variant of this problem, namely non-interactive simulation (NIS), has attracted a lot of attention from both computer science and information theory [5, 22, 29, 35, 42]. Generating correlations with *computational* security and low communication cost has been studied in a recent line of work on pseudorandom correlation generators (PCGs) [12, 16]. A (two-party) PCG is a local deterministic algorithm that stretches a pair of short, correlated seeds into many copies of a correlation, while ensuring that each seed does not reveal more than necessary about the other output. A large body of work in this area [13, 14, 16, 19, 46] has focused on generating OT-correlations extremely cheaply in practice; this makes OT very suitable as a source correlation. However, other useful correlations, such as OLE and multiplication triples [15], are much more expensive to generate, and many other useful correlations do not admit a concretely efficient PCG. This is a primary motivation for our work.

One-way secure computation (OWSC), introduced in [27] and subsequently studied in [2, 3], is closely related to OMSR. Here, a sender and a receiver securely implement a target channel with only one-way communication over a given source channel. Known results about OWSC can be used to realize limited forms of OMSR. In the other direction, OMSR implies OWSC whenever the induced channel of the target correlation allows a *random self reduction*. This is because after realizing OMSR, a random self reduction can be applied with one-directional communication to go from a random sample in the target correlation to the the given input to the sender. However, there is still a separation in terms of efficiency. The completeness of the string-OT correlation for OWSC, which in turn builds on information-theoretic analogs of garbled circuits [33, 49], implies an OMSR converting string-OT correlations to any target correlation. A similar result can be based on the simpler bit-OT correlation, though requiring a much bigger number of copies and inevitably introducing an inverse-polynomial security error in the number of copies [2]. These generic constructions are typically very inefficient. Another drawback is that they entangle communication cost with the number of copies of the source correlation used. Our objective in this paper is to reduce the number of bits transmitted while possibly burning up more copies of the source correlation. Owing to this, the lower bounds in the OWSC model do not provide interesting insights for our model.

All our lower bounds apply more generally to (the non-secure variant) OMR; however, in the presence of common randomness they are only meaningful in the more restrictive OMSR setting. The problem of deriving common randomness from correlations has been extensively studied. The zero-communication version of this problem was studied in [26, 44], which led way to the NIS model we previously described. Generalizing this, in [4], Ahlswede and Csiszar studied the rate of generating common randomness per use of correlation when communication is limited. Several works in computer science [10, 18, 28, 32] considered a related problem of agreement distillation, where parties have unlimited access to a source of noisy common randomness and want to derive common randomness while minimizing communication. Our lower bounds use techniques developed in both these areas of work. To the best of our knowledge, deriving a target correlation given unlimited access to a source correlation is not studied in information theory or theoretical computer science.

2 Technical Overview

We now present an overview of our main technical contributions. We split this into two parts: in the first (Section 2.1), we provide an overview for our concrete OMSR protocols; in the second (Sections 2.2 and 2.3), we present an overview of our lower bound results.

2.1 Concrete OMSR Protocols

OMSR **formalism.** Abstractly, in the OMSR model, we consider two parties: a sender S and a receiver R. The parties are given access to m copies of a *source* correlation (X, Y) jointly distributed according to p_{XY}, with the goal being for them to securely generate n copies of a *target* correlation (U, V) which is jointly distributed according to p_{UV}. To accomplish this task, S is only allowed to send a single message to R; no communication from R is allowed. Intuitively, the security of an OMSR is now defined as neither party learning anything about the other party. In other words, S's view u of the target correlation does not leak anything about R's view v and vice versa.

We focus on two concrete efficiency metrics: the primary one being the (expected) number of bits l communicated, and a secondary one being the number of source correlations $m(n)$ used to generate n target correlations (in the most general case, we allow access to an unlimited number of source correlations).

General OMSR **protocols from rejection sampling.** For our concrete constructions, we study OMSRs in the ρ-Las-Vegas model; here, the output must be correct whenever it is produced but the protocol is also allowed (with probability $\leq \rho$) to return a failure symbol in which case correctness is not guaranteed.

We build general protocols in this model using the following approach: Suppose that the sender's and receiver's views of the source correlation are $\hat{X}$ and $\hat{Y}$. Often, by conditioning this on some variable C (dependent only on the private randomness of the sender), both parties can locally convert to the required target correlation. Equivalently, the sender computes some function $f(\hat{X}; r)$ for some private randomness r such that whenever $f(\cdot) = 1$ is communicated to the receiver, both parties can locally produce the required target correlation. Note that here, the sender only needs to send a single *accept* symbol to indicate whether the conversion can be done. We refer to this as an *accept-reject* protocol with parameter ρ denoting the accept probability (i.e., the probability that the sender produces an accept message).

General OMSR protocols with efficient asymptotic communication can be built using accept-reject protocols through rejection sampling. The intuitive idea is to consider k source copies together instead of one and send an accept message only when all k copies can be successfully converted locally. While this improves communication cost conditioned on an "accept" message, it exponentially reduces the probability of accepting. To get around this, the sender can now instead look at *batches* of k copies and send the *index* of the first batch where all k copies are accepting. This results in an efficient OMSR in the Las Vegas model. We show the following informal result for our general transformation.

Theorem 2.1 (Informal)**.** *A secure accept-reject protocol with probability ρ can be turned into a secure Las Vegas* OMSR *with asymptotic communication* $\log(1/\rho)$*.*

Concrete OMSR **protocol for** (t, q) **and** $(3, 2)$**-correlations.** We use the above general transformation to construct efficient OMSR protocols for (t, q) and $(3, 2)$-correlations. In a (t, q)-correlation (where $t < q$), the two parties are given (x_0, r_0) and (x_1, r_1) respectively where $x_i \in \mathbb{Z}_t$ and $r_i \in \mathbb{Z}_q$ such that $x_0 + x_1 \bmod t = r_0 + r_1 \bmod q$. The (t, q)-correlation generalizes the $(2, 3)$-correlation defined in [24].

We show an accept-reject protocol to generate a (t, q)-correlation using as the source correlation, a 1-out-of-t OT correlation over $\mathbb{Z}_q$ (we formally define this as well as other correlations we use in Section 4.3). For this source correlation, the sender is given a vector $\mathbf{v} = (v_0, \ldots, v_{t-1})$ with each $v_i \in \mathbb{Z}_q$, while the receiver is given (b, v_b) for a random $b \in \mathbb{Z}_t$.

The protocol has accept probability $\frac{tq}{q^t}$ and intuitively works as follows: given $\mathbf{v}$, the sender S checks if there is some (x, r) such that $(x + i) \bmod t = (r + v_i) \bmod q$ for all i. If this is the case, S can output (x, r) and send an accept symbol to R who then can just output its original OT source (b, v_b); this results in a (t, q)-correlation since by construction, regardless of b, $x + b \mod t = r + v_b \mod q$. We can now use the earlier general transformation (Theorem 2.1) to get an OMSR for (t, q)-correlations with communication of $\log(q^t/tq)$.

For $(3, 2)$-correlations, we first show that they are isomorphic to non-zero OLE correlations (see Section 4.3 and Lemma 5.3), following which we can use the above approach to construct an OMSR for them.

2.2 Lower Bounds

We prove linear lower bounds on the communication cost of OMSR for a large family of conversions. These lower bounds also hold more generally for non-secure one message reductions (OMR). For the specific conversions considered in the previous section, the lower bounds we obtain by applying these techniques justify the cost of their OMSR protocols.

Linear lower bound for one-message reductions. Our first result in this section can be stated informally as follows:

Theorem 2.2 (Informal)**.** *The amortized communication cost of OMR converting a correlation (X, Y) to (U, V) is linear if $\mathrm{S}^*(X, Y) < \mathrm{S}^*(U, V)$.*

Here, S^* (see Definition 3.1) of a correlation (X, Y) is defined as $\sup_U \frac{I(U;Y)}{I(U;X)}$ where the supremum is taken over all U that is generated from X (conditionally independent of Y); The connection of this quantity with several other information theoretic measures is outlined in [5]. To prove Theorem 2.2, we first show that the amortized communication complexity of OMR converting (X, Y) to common randomness is exactly $1 - \mathrm{S}^*(X, Y)$. For this, we use a seminal result [4] from information theory which characterized the so called common randomness capacity of any correlation with limited communication. For communication rate $R \geq 0$, common randomness capacity $C(R)$ of a correlation (X, Y) is the asymptotic rate at which common randomness can be derived per use of

(X, Y) by parties (S and R) using only one-way communication (from S to R) with rate limited to R. OMR converting (X, Y) to common randomness differs from this model in that the usage of correlation (X, Y) is not limited. Intuitively, the optimal communication cost of OMR converting (X, Y) to common randomness should be the smallest ratio between R to $C(R)$ as R tends to zero. Although, this observation is referenced in several works [32, 42, 50], to the best of our knowledge this is not formally proved. Revisiting the proof of common randomness capacity region in [4] and using a careful analysis, we prove this fact.

Now, suppose agreeing on n bits of common randomness using (arbitrarily many copies of) (X, Y) requires at least $n \cdot c$ bits of communication. Whereas, only $c' \cdot n$ (where $c' < c$) bits of communication is sufficient to agree on n bits of common randomness using $k \cdot n$ copies of (U, V). Then, OMR converting (X, Y) to (U, V) ought to have a communication complexity of at least $(c - c')/k$. Otherwise, the parties can generate $k \cdot n$ copies of (U, V) correlation with $(c - c')n$ bits of communication and then convert it to n bits of common randomness using less than $c' \cdot n$ bits of communication leading to a contradiction. Hence, we prove the theorem by showing that when $\mathrm{S}^*(X, Y) < \mathrm{S}^*(U, V)$, there is a sufficiently large k and $c' < \mathrm{S}^*(U, V)$, such that two parties can agree on n bits of common randomness using $k \cdot n$ copies of (U, V) and $c' \cdot n$ bits of communication. We then exactly compute S^* of several correlations and use the above technique to obtain concrete linear lower bounds. In Corollary 6.7.1 we show a lower bound on the communication cost of converting 1-out-of-2 OT to $(2, 3)$ correlation that is half of what our construction achieves; in Corollary 6.7.3 we show a lower bound for converting 1-out-of-3 OT to $(3, 2)$ correlation that is a third of what our construction achieves.

Linear lower bound for interactive reductions. We demonstrate much stronger lower bounds on communication costs when the target correlation of interest is "close" to common randomness. Consider an n-bit unit vector correlation in which the parties receive an additive secret sharing of an n-bit unit vector (a string of Hamming weight 1). This correlation is close to n-bits of common randomness in that the two strings are uniformly distributed and differ on exactly one (random) index. Another correlation that is close to common randomness is a 1-out-of-k n-bit string OT, where one party's uncertainty about the other party's part of the correlation is just $1/k$. Our next result shows that, deriving such correlations using any correlation (other than correlations with inherent common randomness) requires linear communication even using an *interactive protocol*. We state the result specifically for unit vector correlations.

Theorem 2.3 (Informal)**.** *If a source correlation (X, Y) lacks common randomness, then any interactive protocol generating an n-bit unit vector correlation requires $\Omega(n)$ communication.*

In [18], Canonne et al. showed that a protocol in which parties with unlimited access to a source of noisy common randomness derive common randomness with ℓ bits of interactive communication can be converted into a zero communication protocol for deriving common randomness from the same source with $(2^{-O(\ell)})$ success probability. We observe that a similar approach can be used to convert an interactive protocol for deriving an n-bit unit vector correlation using a correlation (X, Y) with ℓ bits of interactive communication into a zero communication protocol for deriving the n-bits of common randomness from the same source with about $2^{-O(\ell)}$ success probability. For this, we only use the fact that n-bit unit vector correlation can be converted (with zero-communication) to n bits of common randomness with success probability $1/n$ by having one party simply flipping one of the n bits at random in their share of the correlation.

In the other direction, we show the success probability for agreeing on n bits is $2^{-\Omega(n)}$. Bogdanov and Mossel's result from [10] which showed that deriving n bits of common randomness using a source of noisy common randomness and zero communication succeeds only with $2^{-\Omega(n)}$ probability. The authors show this using hypercontractive inequality. We generalize this result to show that if correlation (X, Y) has no common randomness, then an analogous condition holds. The proof of the statement uses generalized hypercontractive inequality as defined in [35] and Holder's inequality. The lower bound follows from the above observations.

2.3 Role of Common Randomness in OMSR

In general, common randomness does not aid in OMSR. This is in line with the intuition that common randomness available to both parties cannot be used to achieve security. Note that this is in contrast with OMR where any correlation can be generated from common randomness with zero communication. The main result of this section is as follows:

Theorem 2.4 (Informal)**.** *Suppose correlation (U, V) lacks common randomness. Given an OMSR converting a correlation (X, Y) to (U, V) using common randomness we can construct an OMSR for the same conversion without common randomness.*

All (target) correlations considered in this work lack inherent common randomness; more generally, this holds for most correlations with cryptographic applications because, intuitively, common information does not enable cryptographic tasks. In the case of OMSR with perfect security, an OMSR without common randomness can be obtained by simply conditioning on any of its realizations. Such an approach is used to the ineffectiveness of common randomness in statistical NISR [1]; but this approach fails for statistical OMSR. This is because the conversion amplifies the security error by a factor that is inversely proportional to the smallest probability assigned by the correlation to any member in the support. In OMSR, the conversion is to several copies of the target correlation which makes the error in the conversion increase exponentially.

We use a different approach. Consider an ϵ-secure OMSR using common randomness for converting a given source to n copies of the target correlation. By a Markov bound, there exists a realization of common randomness conditioned on which privacy against both parties is guaranteed with at most $\sqrt{\epsilon}$ error. Hence, on average, conditioned on sender's output, the receiver's output is distributed as prescribed by the target distribution, and vice versa. Since the conditional distribution on outputs is correct on average, it is sufficient to show that the marginal distribution of, say, the sender's output is correct. To show this, we consider two experiments; in the first one, we sample the receiver's output conditioned on sender's output and then sample back the receiver's output conditioned on the receiver's output according to conditional distributions prescribed by the output distribution of the OMSR. In the second, we do the same sampling but according to conditional distributions prescribed by the target correlation. These are Markov processes with the stationary distributions being the sender's output distribution and marginal distribution at the sender in the target distribution, respectively. We then use the closeness of the two Markov processes to show that stationary processes are close in total variation distance. This proves the theorem.

3 Preliminaries

Notation. We use calligraphic letters (e.g., $\mathcal{X}$, $\mathcal{Y}$) to denote finite sets or alphabets; the corresponding small letter x is used for members in $\mathcal{X}$, while the capital letter X is used for a random variable with values in $\mathcal{X}$. The distribution induced by X is denoted by p_X, while $X \sim \mu$ means that X follows the distribution μ.

For random variables X, and X' over the same domain $\mathcal{X}$, the total variation (aka statistical) distance TVD between their distributions is defined as:

$$\mathsf{TVD}(p_X, p_{X'}) = \frac{1}{2} \sum_{x \in \mathcal{X}} \left| \Pr\left[X = x\right] - \Pr\left[X' = x\right] \right|$$

For ease of exposition, without loss of generality, we may also often use the notation $\mathsf{TVD}(X, X')$; this will be equivalent to using $\mathsf{TVD}(p_X, p_{X'})$. For any $\epsilon \geq 0$, $X \approx^{\epsilon} X'$ denotes that $\mathsf{TVD}(p_X, p_{X'}) \leq \epsilon$.

We write $(X_i)_{i \in [n]} \overset{\text{i.i.d.}}{\sim} p_X$ to mean that $X_1, \ldots, X_n$ are i.i.d. according to p_X. A sequence of random variables $(X_1, \ldots, X_n)$ will be succinctly represented as X^n; similarly, joint random variables (X_1, Y_1) to (X_n, Y_n) will be represented as (X^n, Y^n).

Random variables (X, Y, Z) satisfy the Markov chain $X \leftrightarrow Y \leftrightarrow Z$ if X and Z are conditionally independent conditioned on Y; i.e., for all x, y, z.

$$\Pr[X = x | Y = y, Z = z] = \Pr[X = x | Y = y].$$

The unit vector with 1 at the i^{th} position is denoted by e_i. The elements of the field $\mathbb{F}_4$ are represented by $\{0, 1, \alpha, \beta\}$ where $\alpha + 1 = \beta$.

Useful quantities. We recall some basic information-theoretic quantities (see [20] for a primer).

The Shannon entropy of X, denoted by $\mathsf{H}(X)$, is defined as $\sum_{x \in \mathcal{X}} \Pr[X = x] \log\left(\frac{1}{\Pr[X=x]}\right)$. The binary entropy function for parameter $\rho \in [0, 1]$ is defined as $\mathsf{H}_b(\rho) = -\rho \log(\rho) - (1 - \rho) \log(1 - \rho)$.

The mutual information of (X, Y), denoted by $I(X; Y)$ is defined as $\mathsf{H}(X, Y) - \mathsf{H}(X|Y) - \mathsf{H}(Y|X)$.

Correlations. The central objects of interest in this work are pairwise joint distributions or correlations. We will often write "correlation (X, Y)" to refer to the correlation p_{XY} induced by (X, Y). We write "parties A and B receive/possess correlation $(X, Y) \sim p_{XY}$" to mean that A and B receive/possess random variables X and Y, respectively, where (X, Y) are jointly distributed according to the distribution p_{XY}.

Definition 3.1 (S* value of a correlation [5])**.** For a correlation (X, Y), the quantity $\mathrm{S}^*(X, Y)$ is defined as

$$\sup_U \frac{I(U; Y)}{I(U; X)},$$

where the supremum is taken over all random variables U generated from X; i.e., (U, X, Y) satisfy the Markov chain $U \leftrightarrow X \leftrightarrow Y$ (See Notations in Section 3) and not independent of X; i.e., $I(X; U) > 0$. Observe that S* is not necessarily symmetric.

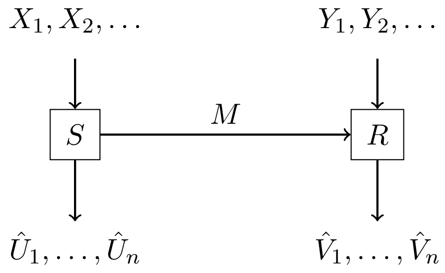

Figure 1: The random variables involved in an OMSR converting p_{XY} into p_{UV}. Parties receive, $(X_1, Y_1), (X_2, Y_2) \dots$, which are i.i.d. according to correlation p_{XY}. M denotes the single message sent from S to R. Parties output $(\hat{U}_1, \hat{V}_n), \dots, (\hat{U}^n, \hat{V}^n)$ which are (close to being) i.i.d. according to correlation (U, V).

Definition 3.2 (Geometric Random Variable)**.** A geometric random variable X with success probability $0 \leq \rho \leq 1$, denoted by $X \sim \mathsf{Geo}(\rho)$ is defined by $\Pr[X = k] = (1-\rho)^k \rho$, for every $k = 0, 1, 2, \dots$.

Fact 1. If $X \sim \mathsf{Geo}(\rho)$ then $\mathsf{H}(X) = \frac{\mathsf{H}_b(\rho)}{\rho}$.

4 One-Message Secure Reductions

We now formally introduce one-message secure reductions, or OMSRs. Abstractly, an OMSR is a secure protocol for converting copies of a source correlation (X, Y) to copies of a target correlation (U, V) by using only a single (uni-directional) message. Later, in Section 4.3, we also define several simple but useful correlations considered by our protocols.

4.1 OMR and OMSR Definitions

Basic model. In an OMSR protocol, there are two parties: a sender S and receiver R. Consider two distributions (correlations) p_{XY} and p_{UV} referred to as the *source correlation* and the *target correlation* respectively. S and R are given an unbounded number of independent copies of correlation p_{XY}; i.e., for $i = 1, 2, \dots$, S gets X_i and R gets Y_i, where (X_i, Y_i) are i.i.d. according to p_{XY}.

The goal now is for S and R to generate n independent copies of the target correlation (U, V). Based on its copies of the source correlation, S will be allowed to send a single message to R. Following this message, S and R compute the required copies target correlations based on their local views. Of particular interest to us are two efficiency metrics: the (expected) number of instances $m = m(n)$ of X needed to generate n instances of Y, and the (expected) length $l = l(n)$ of the message from S to R. Hence, we define the expected amortized communication cost as $\limsup_n E[l(n)]/n$ and the worst case amortized communication cost as $\limsup_n \max(l(n))/n$.

Before detailing the security properties of OMSRs, we first introduce its non-secure counterpart—the one-message reduction (OMR).

Definition 4.1 (One-Message Reduction (OMR))**.** An ϵ-error *one-message reduction* (ϵ-OMR) over $(m, p_{XY}, n, p_{UV}, l)$ is a pair of randomized algorithms $\langle S, R \rangle$ for (non-securely) converting m

copies of a source correlation p_{XY} (over the domain $\mathcal{X} \times \mathcal{Y}$) to n copies of target correlation p_{UV} (over the domain $\mathcal{U} \times \mathcal{V}$) using l bits of communication.

Let the private randomness of the sender and receiver be uniformly distributed in arbitrary finite domains $\mathcal{Q}$ and $\mathcal{Q}'$, respectively. The algorithms are defined as $S : \mathcal{X}^m \times \mathcal{Q} \to \mathcal{U}^n \times \{0,1\}^l$ and $R : \{0,1\}^l \times \mathcal{Y}^m \times \mathcal{Q}' \to \mathcal{V}^n$ and they satisfy the following correctness condition:

Correctness. Let $(X_i, Y_i)_{i\in[m]} \overset{\text{i.i.d.}}{\sim} p_{XY}$, and $(U_i, V_i)_{i\in[n]} \overset{\text{i.i.d.}}{\sim} p_{UV}$. Let Q, Q' be uniformly distributed in $\mathcal{Q}, \mathcal{Q}'$, respectively. Then, $(\hat{U}^n, M) \leftarrow S(X^m, Q)$ and $\hat{V}^n \leftarrow R(M, Y^m, Q')$ are such that:

$$\left(\hat{U}^n, \hat{V}^n\right) \approx^{\epsilon} (U^n, V^n) \tag{1}$$

We say that an OMR is *perfect* if $\epsilon = 0$. When m is omitted (or $m = \infty$), the OMR will be given an unbounded number of copies of the source correlation. We also allow for m and l to be randomized functions of the n, in which case we will look at the expected number of source correlations used and the expected number of bits communicated as our efficiency metrics.

OMSR **security.** An OMSR is an OMR where the conversion is also done *securely.* Informally, we define security as neither party learning more about the output of the other party than it should. We formalize this in Definition 4.2.

Definition 4.2 (One-Message Secure Reduction (OMSR))**.** An ϵ-error one-message secure reduction (ϵ-OMSR) over $(m, p_{XY}, n, p_{UV}, l)$ is an ϵ-OMR $\langle S, R\rangle$ (for converting m copies of the source correlation p_{XY} to n copies of the target correlation p_{UV} using l bits of communication) which also satisfies the following security properties:
Privacy against S**.** Let $X^m, Y^m, Q, Q', M, \hat{U}^n, \hat{V}^n$ be as defined in Definition 4.1. Then,

$$\mathbb{E}_{X^m,Q}\left[\mathsf{TVD}\left(\left(\hat{V}^n\middle|X^m, Q\right), \left(\hat{V}^n\middle|\hat{U}^n\right)\right)\right] \leq \epsilon. \tag{2}$$

Privacy against R**.** For $X^m, Y^m, \hat{U}^n, \hat{V}^n$ and M as defined above,

$$\mathbb{E}_{Y^m,M,Q}\left[\mathsf{TVD}\left(\left(\hat{U}^n\middle|M, Y^m, Q'\right), \left(\hat{U}^n\middle|\hat{V}^n\right)\right)\right] \leq \epsilon. \tag{3}$$

We say that an OMSR is perfect if $\epsilon = 0$.

We provide an alternate definition for OMSR with simulation-based security in Appendix A and prove that both definitions are equivalent with a comparable error.

Statistical OM(S)R. For a function $\epsilon : \mathbb{N} \to \mathbb{R}_{\geq 0}$, we say there is an $\epsilon(n)$-statistical one-message (secure) reduction converting p_{XY} to p_{UV} if, for each n, there exists an $\epsilon(n)$-OM(S)R converting (arbitrarily many copies of) p_{XY} to n copies of p_{UV} using $l(n)$ bits of communication. The communication cost for the OM(S)R is computed as $\limsup_n l(n)/n$. If $\epsilon(n)$ is a negligible function, we call the reduction an OM(S)R with negligible error.

The lower bounds we develop in this paper apply to statistical OMSR, in fact, more generally to statistical OMR. Note that the lower bounds also apply to OM(S)R with expected communication cost with variable message length and perfect correctness (and privacy). This can be seen as follows:

suppose the OM(S)R has an expected communication cost of ℓ. Suppose the scheme is run n times independently, and let ℓ_i be the length of the message in the i-th execution of the scheme. The amortized length of the combined message is $(\ell_1 + \ldots + \ell_n)/n$; to combine messages we crucially use the fact that they are prefix-free. Fix $\epsilon > 0$; by the law of large numbers, for any $\delta > 0$, there exists a large enough n such that the amortized length is $\ell + \epsilon$ with probability $1 - \delta$. Hence, by aborting (sending $\perp$) whenever the amortized length is more than $\ell + \epsilon$, we obtain a statistical OM(S)R with a communication cost of $\ell + \epsilon$. Thus, a lower bound on the communication cost of statistical OM(S)R implies a lower bound on the expected amortized communication cost of OM(S)R.

OMSR **for distribution families.** In many cases, we are also interested in generating families of correlations starting from a given source correlation. Here, a *family* $\mathfrak{F}$ of correlations is a sequence of correlations parameterized by $n \in \mathbb{N}$, i.e., $\mathfrak{F} = \{(U_n, V_n)\}_{n \in \mathbb{N}}$. Generating correlation families are of practical interest. Examples include, the unit vector correlation family–a sequence of n-bit unit vector correlations for $n \in \mathbb{N}$, the string-OT correlation family–a sequence of n-bit 1-out-of-2 OT correlations for $n \in \mathbb{N}$, etc.

Note that, in Definition 4.1 and Definition 4.2, the correlations families of interest are $\mathfrak{F} = \{(U_n, V_n)\}_{n \in \mathbb{N}}$, where (U_n, V_n) is a sequence of n i.i.d. copies of a target distribution. Hence, the definitions for OMR and OMSR in Definition 4.1 and Definition 4.2 naturally extend to correlation families. We provide a formal definition of OMSR for distribution families in Appendix A.

4.2 OMSR in the Las Vegas Model

The Las Vegas model of computation requires algorithms or protocols to always output the correct result whenever some result is produced but allows for the output of a special failure symbol $\perp$, in which case no guarantees are made about correctness. In the Las Vegas model, the runtime may also depend on the input (and randomness). Many of our concrete constructions use this Las Vegas model; we formally define OMSR in this model below.

Definition 4.3 (OMSR in the Las Vegas Model)**.** A OMSR in the ρ-Las-Vegas model is the same as a perfect (i.e., with $\epsilon = 0$) OMSR from Definition 4.2 over the parameters $(m, p_{XY}, n, p_{UV}, l)$ except for the property that the parties are additionally allowed to output a failure symbol $\perp$ (with probability $\leq \rho$). The correctness and security properties are exactly the same as Definition 4.2 except that now they are conditioned on the output not being $\perp$.

Las-Vegas OMSR is a stronger notation than statistical OMSR. Hence, we use this notion in our constructions. The lower bounds we develop for statistical OMSR naturally applies to Las-Vegas OMSR as well.

The accept-reject paradigm. To build OMSR protocols in the Las Vegas model, we find it useful to define a simpler primitive where only a *single symbol* is sent from the sender to the receiver, after which both parties produce an output. Informally, this symbol represents an "accept" indication which signals that a target correlation can be realized based on the sender's view of the source correlation. Following this, the receiver can locally convert its view of the source correlation to the target correlation. If there is no communication, both parties output the failure symbol $\perp$ in which case no instance of the target correlation is produced. We refer to this as the accept-reject paradigm.

More formally, a probability ρ (secure) accept-reject protocol ρ-$\mathsf{Acc\text{-}Rej}$ from a source correlation (X, Y) to a target correlation (U, V) is an OMSR in the Las Vegas model over the parameters $(m, p_{XY}, n = 1, p_{UV}, l = 1)$ where S sends an "accept" symbol to R with probability ρ (taken over the the source correlations X^m). The goal is to generate just a single instance of the target correlation (as opposed to the general Las Vegas model). We often use the terminology that the sender's view is *accepting* if it results in an accept message being sent. While we denote $l = 1$, note that only a single symbol (rather than a bit) needs to be transmitted when accepting while nothing is communicated when rejecting.

4.3 Useful Correlations

We now define several simple but useful (2-party) correlations that are widely applicable for building efficient secure computation protocols. Several of our OMSR protocols will involve securely converting between these correlations.

Oblivious transfer (OT) correlation. A 1-out-of-k OT correlation over group $\mathbb{G}$ is a tuple $(\mathbf{r}, (b, r_b))$ where $\mathbf{r} = (r_0, \dots, r_{k-1})$ is uniform over $\mathbb{G}^k$ and b is uniform over $\mathbb{Z}_k$.

Oblivious linear evaluation (OLE) correlation. OLE can be viewed as an arithmetic extension of 1-out-of-2 OT. Specifically, an OLE correlation over a field $\mathbb{F}$ is a tuple $((a, s), (b, r))$ where a, b, s are uniform over $\mathbb{F}$ and $r = ab + s$.

Non-zero OLE correlation. An nzOLE correlation is simply an OLE correlation that is conditioned on the event that $a, b \neq 0$.

(t, q)-correlation for $t < q$. A (t, q)-correlation where $t < q$ is the tuple $((x_0, r_0), (x_1, r_1))$ where we choose $x_0, x_1 \in \mathbb{Z}_t$ and $r_0, r_1 \in \mathbb{Z}_q$ at random under the constraint that $x_0 + x_1 \pmod t = r_0 + r_1 \pmod q$. This generalizes the $(2, 3)$-correlation used by Dinur et al. [24] to securely convert an additively shared bit over $\mathbb{F}_2$ to an additive sharing of the same bit over $\mathbb{F}_3$.

$(3, 2)$-correlation. A $(3, 2)$-correlation is the tuple $((x_0, u_0, v_0), (x_1, u_1, v_1))$ where we choose $x_i \in \mathbb{Z}_3$, $u_i, v_i \in \mathbb{Z}_2$ at random under the following constraints: x_0, u_0, v_0, x_1 are uniformly random and independent. Define $x = x_0 + x_1 \bmod 3, u = u_0 + u_1 \bmod 2, v = v_0 + v_1 \bmod 2$. Then we require that $u = x \bmod 2$ and $v = (x + 1 \bmod 3) \bmod 2$. This correlation was used in [24] to securely convert a mod-3 sharing of x to a mod-2 sharing of $x \bmod 2$. Perhaps surprisingly, we show that the $(3, 2)$-correlation is also isomorphic to a non-zero OLE over $\mathbb{F}_4$.

n-bit unit vector (n-UV) correlation. An n-UV correlation is a tuple (u_0, u_1) where $u_0, u_1 \in \mathbb{F}_2^n$ and $u_0 + u_1$ is a random unit vector.

Additive correlation. We will say a correlation (X, Y) is an additive correlation if there exists a distribution ψ over an abelian group $\mathbb{G}$ such that $X + Y \sim \psi$ and X and Y are both uniform over the group. Note that this generalizes several correlations, including n-UV and the $(2, 3)$-correlation.

5 Concrete OMSR Protocols

In this section, we provide concrete one-message secure reductions for converting from a source correlation (X, Y) to a target correlation (U, V). As mentioned earlier, we find it useful to build accept-reject protocols as a stepping stone to building Las Vegas OMSRs. We start with a general transformation that enables us to, in many cases, work with simpler accept-reject protocols.

Theorem 5.1. *Suppose that there exists a secure accept-reject protocol π from source correlation (X, Y) to target correlation (U, V) with accept probability ρ. Then, for every $\epsilon > 0$, there exists a perfect* OMSR *converting from (X, Y) to (U, V) with an expected amortized communication cost smaller than $\log(1/\rho) + \epsilon$.*

Proof. Suppose that the protocol π has an accept probability ρ. We show how to use this in a black-box way to construct an OMSR; the key idea is to process source correlations in batches of size km (where m in the number of correlations used by π) before sending a message.

Concretely, suppose that π involves the sender S looking at m source correlations, and based on some (possibly randomized) function computation $f(X^m; r)$ with private randomness r resulting in 1, decides on whether to send an accept symbol to R. For the OMSR model, recall that we always want to produce the target correlation (instead of e.g., with probability ρ). To achieve this, the basic idea is to have S send the *index* of the first set of correlations for which π would result in an accept message.

It turns out that we can substantially bring down the asymptotic cost by considering batches of size km. Notice as a first step that we can directly get an accept-reject protocol to generate k target copies; for this, S will look at km copies of the source correlation (as k batches X_i^m each of size m) and send a single-bit accept message only when $f(X_i^m; r_i) = 1$ is true for all $i \in \{1, \ldots, k\}$. This however results in accept probability ρ^k, i.e., exponentially decreasing in k. Now, to convert this into an OMSR, we can let S look at its source copies in batches of size km, and now send the index of the first batch where all the k instances in the batch would result in an accept message in π. Upon receiving this index, the receiver R can use its corresponding set of source correlations; this generates k copies of the target correlation. We denote this protocol by π_k^*.

Observe that the message from S is now a geometric random variable with success probability ρ^k. We can compress this message down to its entropy and achieve an expected amortized communication cost of $\frac{(1/\rho)^k \cdot \mathsf{H}_b(\rho^k)}{k}$. Notice that by taking a larger k, the communication cost is reduced at the cost of consuming more source copies. Since the limit of the function as $k \to \infty$ is $\log(1/\rho)$, for any $\epsilon > 0$, we can choose an appropriate k such that the expected amortized communication cost is smaller than $\log(1/\rho) + \epsilon$.

To complete the proof, we now show the privacy of π_k^*. This is easy to see intuitively: all of the copies of the correlation are independent, the output is only dependent on the utilized batch of the km source correlations, and the message from S doesn't reveal any more information than the message in π. Therefore, since π is secure, π_k^* should also be secure.

More formally, let us prove the privacy against R. Let $M = i$ be the message sent, Y^{ki} that was used by R and $\hat{V}^n$ the output of the protocol. We will denote by $\pi(Y)$ the output of R when an accept message was sent.

$$\begin{aligned}(\hat{U}^n|\hat{V}^n, Y^{ki}, M = i) &\equiv ((\pi(Y_{k(i-1)+1}), \dots, \pi(Y_{k(i-1)+k}))|Y^{ki}, M = i)\\ &\equiv ((\pi(Y_1), \dots, \pi(Y_k))|Y^k, M = i)\\ &\equiv ((\pi(Y_1), \dots, \pi(Y_k))|Y^k)\\ &\equiv (U^n|V^n)\end{aligned}$$

The distributions are equivalent for every message and copies of the source correlation used, and thus in expectation over M and copies of the source, their total variation distance will be 0.

The proof of privacy against S proceeds in a similar manner; Let $M = i$ be the message sent, X^{ki} that was used by S and $\hat{U}^n$ the output of the protocol. Note that in this case we know that the first $(i-1)$ k-tuples will be rejected, and the last one will be accepted. We will denote by $\pi(X)$ the output of S when an accept message was sent.

$$\begin{aligned}(\hat{V}^n|X^{ki}, M = i) &\equiv ((\pi(X_{k(i-1)+1}), \dots, \pi(X_{k(i-1)+k}))|X^{ki}, M = i)\\ &\equiv ((\pi(X_1), \dots, \pi(X_k))|X^k, M = i)\\ &\equiv ((\pi(X_1), \dots, \pi(X_k))|X^k)\\ &\equiv (V^n|U^n)\end{aligned}$$

The distributions are equivalent for every message, copies of the source correlation used and the output and thus in expectation over them, the total variation distance between the distributions will be 0.

□

Remark (Extensions). In Section 5.3, we further optimize the above transformation for better concrete efficiency in a number of metrics. We also present a modified construction for which we can bound the worst-case communication (as opposed to the expected communication); this can be done within the Las Vegas model by outputting a failure symbol with only negligible probability.

5.1 Efficient OMSR for (t, q)-correlations

In this section, we show efficient OMSR protocols for generating (t, q)-correlations from OT-correlations. This provides concrete improvements as well as generalizes the protocol from [24, Protocol 5.2] for generating $(2, 3)$-correlations.

Theorem 5.2. *There exists a secure accept-reject protocol π for source 1-out-of-t OT correlation over $\mathbb{Z}_q$ and target (t, q)-correlation with accept probability $\frac{tq}{q^t}$. In particular, for $(2, 3)$-correlations, the accept probability is $\frac{2}{3}$.*

Proof. Recall that in a 1-out-of-t OT correlation over $\mathbb{Z}_q$, the sender S is given a tuple $\mathbf{r} = (r_0, \dots, r_{t-1})$ that is uniform over $\mathbb{Z}_q^t$, and the receiver R is given (b, r_b) where b is a uniformly random over $\mathbb{Z}_t$ and r_b is its corresponding element in $\mathbf{r}$.

Now, define a function $f_{r,s}(x) : \mathbb{Z}_t \to \mathbb{Z}_q$, parameterized by $r \in \mathbb{Z}_t$ and $s \in \mathbb{Z}_q$ as $f_{r,s}(x) = w$ where $w \in \mathbb{Z}_q$ is such that $(r + x) \bmod t = (s + w) \bmod q$. In other words, the output w is such that $((r, s), (x, w))$ is a valid (t, q)-correlation. Note that the function $f_{(r,s)}$ is distinct for distinct (r, s).

Now, for each (r, s), define the vector $\mathbf{a}_{(r,s)} = (f_{(r,s)}(0), \dots, f_{(r,s)}(t-1))$, i.e., defined by evaluating $f_{(r,s)}$ at each $i \in \mathbb{Z}_t$. Denote by Φ, the set of all possible vectors $\mathbf{r} \in \mathbb{Z}_q^t$ for which there exist some (r, s) such that $\mathbf{a}_{(r,s)} = \mathbf{r}$. Intuitively, Φ denotes the accept set for the conversion—when the OT source correlation $\mathbf{r}$ given to S is in Φ, then it will send an "accept" message after which both parties will compute the target (t, q) correlation by local computation; when $\mathbf{r} \notin \Phi$, then both parties will abort.

Now, given $\mathbf{r} \in \Phi$, to generate required the (t, q) correlation, S will find the (r, s) such that $\mathbf{r} = \mathbf{a}_{(r,s)}$ and output it; the receiver R will simply output (b, r_b).

Notice that this results in a valid (t, q)-correlation since $f_{(r,s)}(b) = r_b$, and therefore $r+b \bmod t = s + r_b \bmod q$.

Observe that there are tq valid tuples (r, s), each corresponding to a unique "accepting" $\mathbf{r}$ as given above. Therefore, over a random $\mathbf{r} \in \mathbb{Z}_q^t$, the probability that $\mathbf{r} \in \Phi$ will be $\frac{tq}{q^t}$; this is exactly the accept probability of the accept-reject protocol. Notice that this protocol works best when t is a small number.

It is easy to see that both parties don't learn any additional information about the other's output; The sender doesn't know which value of b the receiver has. The receiver's view consists of only b, r_b and he doesn't know which r, s were consistent with the sender's part of the correlation. □

Corollary 5.2.1. *For every $\epsilon > 0$, there exists an* OMSR *for converting from 1-out-of-2 OT over $\mathbb{Z}_3$ to $(2, 3)$-correlations with an expected amortized communication cost smaller than* $\log(3/2) + \epsilon$.

Proof. This is a direct consequence of Theorems 5.1 and 5.2. □

Remark. Our protocol provides concrete improvements in generating $(2, 3)$-correlations compared to the protocol from [24]. In particular, the protocol from [24] has an expected communication cost of $1.5 \cdot \mathsf{H}_b(1/3) \approx 1.377$ bits to generate one $(2, 3)$-correlation instance; our protocol brings this cost down to just $\log(3/2) + \epsilon \approx 0.585 + \epsilon$ for any $\epsilon > 0$—an over 2x improvement to an already optimized protocol.

Remark. In the lower bounds section, we show a lower bound of $\log(q/2)/2$ for any OMR converting OT to $(2, q)$ correlation, exactly a half of our upper bound (Corollary 6.7.1). In addition, we also show that an optimal OMR converting OT correlation to $(2, 3)$ correlation would be with 1-out-of-2 OT (Corollary 6.7.2).

5.2 Efficient OMSR for $(3, 2)$-correlations

We now show efficient OMSR protocols for generating $(3, 2)$-correlations from OT. An interesting result we show is that $(3, 2)$-correlations are isomorphic to non-zero OLE correlations. The full proofs are deferred to Appendix B.

Lemma 5.3. *The non-zero OLE (* nzOLE*) over $\mathbb{F}_4$ is isomorphic to the $(3, 2)$-correlation. In other words, there is a secure no-communication reduction (i.e., an SNIR) between the two correlations (in both directions).*

This lemma is proved in Appendix B.1.

Theorem 5.4. *There exists a secure accept-reject protocol π for source 1-out-of-3 OT over $\mathbb{F}_4$ and target $(3, 2)$-correlation with accept probability $\frac{3}{16}$.*

This protocol is quite similar to the one for (t, q) correlation. For completeness, the full proof is given in Appendix B.2.

Corollary 5.4.1. *For every $\epsilon > 0$ there exists an* OMSR *for converting 1-out-of-3 OT over $\mathbb{F}_4$ to $(3, 2)$-correlations with expected amortized communication cost of* $\log(\frac{16}{3}) + \epsilon$*.*

Proof. This is a direct consequence of Theorems 5.1 and 5.4. □

Remark. Our protocol provides concrete improvements in generating $(3, 2)$-correlations compared to the protocol from [24]. In particular, the protocol from [24] has a communication cost of 6 bits to generate one $(3, 2)$-correlation instance; our protocol brings this cost down to just $\log(16/3) + \epsilon \approx 2.415 + \epsilon$ for any $\epsilon > 0$—an over 2x improvement.

Remark. In the lower bounds section, we show a lower bound of $\log(16/3)/3$ for any OMR (i.e., even without security) for converting 1-out-of-3 OT to a $(3, 2)$ correlation—exactly a third of our upper bound (Corollary 6.7.3).

5.3 Efficiency Metrics and Optimizations

In this section, we present several optimizations for our generic transformation (Theorem 5.1) from accept-reject protocols to OMSRs in the Las Vegas model.

Number of source correlation copies used. In the context of Theorem 5.1, the number of copies of the source correlation used grows exponentially with k. We now show how to use fewer copies of an OT source correlation while keeping communication the same by using long *string* OT-correlations instead of OT correlations over a small group. This technique is highly useful in practice since string OT-correlations are equally easy to produce as regular OTs—a trivial use of a PRG can extend short random strings to long pseudo-random strings.

For the optimization, first notice that the decision for whether to accept or reject a k-tuple of the source correlation is only based on the view of the sender S. We exploit this observation in the following way: Consider the source correlation to be a long string-OT. S proceeds to "cut" the long string into small segments according the size of the original OT correlation required (for example, if the source correlation in the original protocol was OT over $\mathbb{F}_4$ then each segment will be of length of two bits). Now, S just needs to send the index of the first batch of k segments where the underling accept-reject protocol would send an accept message. This results in the same communication as before since the message here again is the same geometric random variable; note that it also does not decrease the computation required in terms of bits that the sender has to read. Still, the upshot of this technique now is that it only requires 1 string-OT correlation to generate k instances of the target correlation.

Optimizing computation. While the optimization using string-OT reduces the number of source correlations required, the computation required in terms of the number of bits read does not decrease; it is still exponential in k. More specifically, for an accept-reject protocol with probability ρ, the computation per instance of the target correlation generated is proportional to $(1/\rho)^k$. We will now show how to optimize the number of source copies in the original protocol, thereby also reducing the computation required by the parties.

Correlation	Efficiency	k value					
		1	2	5	10	15	∞
$(2,3)$-correlation	comm. (bits)	1.377	1.114	0.853	0.727	0.681	0.585
	computation	1.5	2.25	7.59	57.66	437.8	∞
$(3,2)$-correlation (optimized)	comm. (bits)	3.08	2.87	2.66	2.55	2.51	2.415
	computation	1.33	1.77	4.21	17.75	74.83	∞

Table 1: Trade-off between the communication (in bits) and the computation (in terms of the expected number of original source copies to be read) per instance of the target correlation generated as the value of k increases.

As a specific illustrative example, we focus on the protocol to generate $(3,2)$-correlations. Recall that for $k = 1$, using the notation from Theorem 5.4, the sender will first check whether its view $\mathbf{r} = (r_0, r_1, r_2)$ of the OT correlation is of the form $(f_{a,s}(1), f_{a,s}(\alpha), f_{a,s}(\beta))$ for some $a \in \mathbb{F}_4 \setminus \{0\}; s \in \mathbb{F}_4$ where $\mathbb{F}_4$ is written as $\{0, 1, \alpha, \beta = \alpha + 1\}$. The protocol accepts if the sender's OT correlation is of the correct form, following which both parties can locally compute the target $(3,2)$-correlation. It is easy to see that there are $3 \times 4 = 12$ distinct valid OT-correlations that will result in an accept message; equivalently, for a particular (r_0, r_1), there is exactly 1 r' such that (r_0, r_1, r') results in an accept message. This happens only with probability 1/4 over a random source correlation, which makes the probability of the accept-reject protocol $\frac{12}{16} \cdot \frac{1}{4} = \frac{3}{16}$. In turn, this affects the communication and the source copies used (i.e., computation) of the compiled Las Vegas OMSR protocol.

We use a simple optimization trick here—instead of hoping for r_2 in (r_0, r_1, r_2) to be of the correct form, the sender can simply force it to be so by sending the $\mathbb{F}_4$ element r^* such that $(r_0, r_1, r_2 + r^*)$ will result in an accepting execution. This results in an extra 2 bits of communication from the sender. But since it increases the accept probability by a factor of 4, there is no change in the communication cost. Still, the upshot is that it decreases the number of source correlation copies used from $(\frac{16}{3})^k$ to $(\frac{4}{3})^k$, i.e., an improvement by a factor of 4^k. Such optimization can therefore exponentially improve the computation cost of the protocol.

We point out that this optimization is quite general. Whenever the sender correlation is (u, v) and conditioned on u, there is only one accepting v, acceptance can be forced by sending an appropriate v' such that $(u, v + v')$ is now accepting. Such an optimization can drastically improve the computation cost.

Computation vs communication trade-off. Following our string-OT optimization, a trade-off between computation and communication is uncovered as we increase the value of k; while the use of string-OT allows reducing the number of instances of the source correlation used, the number of bits read (or in other words, the computation required) is still exponential in k. In particular, observe that as $k \to \infty$, while the amortized communication tends to $\log(1/\rho)$ where ρ is the accept probability of the underlying protocol, the amortized computation required grows exponentially. Despite this, we find that we can choose k to be fairly small and still achieve a substantial reduction in the communication cost. This is illustrated in Table 1 for the $(2,3)$ and $(3,2)$-correlations.

Expected to worst-case. So far, for the general protocol (Theorem 5.1), we only looked at the *expected* communication and computation cost; the worst case cost is unbounded, of course with exponentially decreasing probability. We now show that in the Las Vegas model, by allowing a negligible (in n) failure probability, we can bound even the *worst case* cost of the protocol; notably, this worst case bound is only slightly worse than the expected cost of the original protocol.

Consider t mk-tuples of the source correlation copies. Let B be the event that among the t tuples, the number of accepted ones is smaller than $\frac{n}{k}$. We now want to bound the probability of the event B.

Let $\{I_i\}_{i=1}^{t}$ be indicator variables for whether the i-th mk-tuple is accepted and μ denote the expected number of total tuples accepted. The indicators are all i.i.d. and binomially distributed with probability ρ^k. Observe also that $\mathbb{E}\left[\sum_{i=1}^{t} I_i\right] = \mu = \rho^k t$. Define $\delta = 1 - \frac{n/k-1}{\mu}$. Now,

$$\Pr[B] = \Pr\left[\sum_{i=1}^{t} I_i \leq n/k - 1\right] = \Pr\left[\sum_{i=1}^{t} I_i \leq (1-\delta)\cdot\mu\right] \leq 2^{-\Omega\left((1-\frac{n/k-1}{\mu})^2\mu\right)}$$

where the inequality is by using the Chernoff bound.

Therefore, for $\mu \geq (1+\epsilon)(n/k-1)$ (which can be achieved with $t = \lceil(1+\epsilon)\cdot\frac{n/k}{\rho^k}\rceil$), it holds that $\Pr[B] \leq 2^{-\Omega(n\epsilon^2/k)}$. Thus by using $\epsilon m(n/\rho^k)$ more copies of the source correlation and with all but negligible probability, the OMSR will succeed and we can bound the worst-case complexity for the number of source correlations used. In order to achieve a worst case bound on the amortized communication cost, we cannot use compression like in Theorem 5.1; note however, that we can encode the indices of the batches that we accepted using $\frac{n}{k}\log(t) = \frac{n}{k}(\log(1+\epsilon) + k\log(1/\rho) + \log(n/k))$ bits and achieve an amortized communication cost of $\log(1/\rho)$ as $k \to \infty$ and thus achieve the same asymptotic communication cost. An implicit point to be noted here is that in order for the probability of the event B to be negligible using the Chernoff bound, it must be that $k = o(n)$; for instance, we can use $k = \sqrt{n}$.

5.4 Concrete Improvements for Existing Applications

We now show how our OMSR protocols for $(2,3)$ and $(3,2)$-correlations lead to concrete efficiency improvements for existing applications. As mentioned earlier, a primary motivation for our study of $(2,3)$ and $(3,2)$-correlations was their importance in recent work by Dinur et al. [24]; we significantly improve over their optimized constructions for these types of correlations.

In particular, [24] proposes candidate constructions in the so called *alternating moduli* paradigm for symmetric-key primitives like weak-PRFs, OWFs, and PRGs; the key idea here being that by mixing linear functions over different moduli (e.g., 2 and 3), resistance to known cryptanalysis techniques can be argued. At the same time, this leads to highly efficient evaluations particularly in distributed settings since most parts of the construction involve linear operations which are cheap to perform. The only non-linear operations required in these constructions were conversions of secret shared values from $\mathbb{Z}_2$ to $\mathbb{Z}_3$ and vice versa. Both the $(2,3)$ and $(3,2)$-correlations were introduced in this context to enable more efficient online protocols; optimized ways to generate these correlations from OT were also constructed.

Our OMSR protocols show further improvements to the generation of these useful correlations; This directly translates to significant improvements in the distributed protocols, which were shown to already be competitive compared to prior work. Table 2 shows a 2x improvement across the

MPC-Friendly Constructions [24]	Parameters (η, υ, τ)	Offline comm.		Online comm. (for both)
		[24]	This work	
$(2, 3)$-wPRF	$(256, 256, 81)$	353	150	1536
LPN-wPRF	$(256, 256, 128)$	1889	768	2860
$(2, 3)$-OWF	$(128, 452, 81)$	623	265	904
LPN-PRG	$(128, 512, 256)$	3249	1312	1880

Table 2: Concrete comparison of our OMSRs vs [24] in the context of the cost for different MPC-friendly constructions in the distributed 2PC setting. All numbers in the table are in bits. The parameters (η, υ, τ) denote the length of the input, the length of the intermediate layer, and the length of the output in the constructions. Concrete parameters were chosen in [24] based on cryptanalysis.

board for the prepossessing cost associated with 2PC distributed protocols for all constructions studied in [24]. Concrete applications of our efficiency gains include:

- (Oblivious PRFs). The $(2, 3)$-weak-PRF was shown to have significantly better performance (4-5x faster) in the *oblivious* evaluation setting compared to existing algebraic OPRFs at the cost of requiring some prepossessing; our optimizations provide the same performance while reducing the client's offline communication by roughly 33%.

- (MPC-in-the-head signatures). The $(2, 3)$-OWF, when used to generate signatures through the MPC-in-the-head paradigm, resulted in roughly 10% smaller signatures sizes than using the LowMC blockcipher; our optimizations would result in even smaller signature sizes.

- (Function secret sharing and applications). The LPN-PRG is useful for several distributed applications that require length-doubling PRGs with the same input and output space. Of particular relevance are the distributed generation of function secret sharing (FSS) keys, distributed point functions (DPF), and distributed comparison functions (DCF); the core operation here is the distributed evaluation of the PRG, which is significantly more communication efficient through an MPC-friendly PRG than e.g., using AES. For all these applications, our optimizations provide further efficiency by reducing the preprocessing cost of the LPN-PRG by roughly 2.5x.

FSS has also found applications in privacy-preserving machine learning for dealing with non-linear functions, which arguably contribute to the bulk of the cost in the secure computation setting. In this vein, the usefulness of an MPC-friendly PRG was recently demonstrated by [47] to build a DCF for neural network training (each ReLU activation function involves a DCF evaluation which translates to n distributed PRG evaluations when the data values are over $\mathbb{Z}_{2^n}$) . Prior work either required distributed evaluation of a PRG that was not MPC-friendly which led to high communication cost, or used a technique by Doerner-shelat [25] which requires computation exponential in the input size and therefore is useful only for small input domains ($< 2^{16}$). In contrast, [47] used the LPN-PRG from [24] to efficiently support large input domains of size 2^{32}. Concretely, given inputs in $\mathbb{Z}_{2^{32}}$, each ReLU evaluation requires 32 PRG evaluations; therefore, for each non-linear ReLU layer, as earlier, our optimizations translate to roughly 2.5x smaller offline and 1.6x smaller total communication.

6 Lower Bounds

In this section, we prove lower bounds on the communication cost of (non-secure) one-message reduction.

6.1 Linear Lower Bound for One-Message Reductions

Theorem 6.1 proves a constant lower bound on amortized communication cost of statistical OMR converting a target correlation to another with larger S^* value. In Corollaries 6.7.1 to 6.7.3, we use a more fine-grained analysis to prove concrete lower bounds for the correlations we considered in the previous section.

Theorem 6.1. *Amortized communication cost of statistical OMR converting a correlation (X, Y) to (U, V) is strictly positive if $\mathrm{S}^*(X, Y) < \mathrm{S}^*(U, V)$.*

We will assume that the sender's algorithm is a deterministic function of their part of the correlation and that of the receiver is a deterministic function of the received message and their part of the correlation. This is without loss of generality because the protocol can keep aside sufficiently many copies of the source correlation for both sender and receiver to extract private randomness from. Our lower bounds allow the parties to use arbitrarily many copies of source correlation.

We first state Lemmas 6.2 to 6.4 which together imply the theorem.

Lemma 6.2. *Amortized communication cost of OMR converting a correlation (X, Y) to common randomness is lower bounded by $1 - \mathrm{S}^*(X, Y)$.*

We will show that any achievable amortized communication cost c of OMR converting a correlation (X, Y) to common randomness satisfies

$$c \geq 1 - \sup_{U:U \leftrightarrow X \leftrightarrow Y} \frac{I(U;Y)}{I(U;X)},$$

where U is any random variable that is generated from X, i.e., it satisfies Markov chain $U \leftrightarrow X \leftrightarrow Y$. Since $\mathrm{S}^*(X, Y)$ is defined as the expression in the RHS, the lemma follows. In order to prove the above inequality, we revisit the problem of common randomness capacity of a correlation which was studied by Ahlswede and Csiszar in [4]. For communication rate $R \geq 0$, common randomness capacity $C(R)$ of a correlation (X, Y) is the asymptotic rate at which common randomness can be derived *per use* of (X, Y) using only one-way communication (from S to R) with rate limited to R. By retracing the proof of converse for common randomness capacity, we show that the amortized communication cost of statistical OMR converting (X, Y) to common randomness cannot be lower than the smallest ratio between R to $C(R)$ as R tends to zero. Our proof closely follows the aforementioned converse; this is provided in Appendix C.1.

Lemma 6.3. *For any correlation (U, V) and $c > 1 - \mathrm{S}^*(U, V)$, there exists a constant $\lambda > 0$ such that, for each $\epsilon > 0$ and for all sufficiently large n, there are functions $S : \mathcal{U}^n \to [\lfloor 2^{n/\lambda} \rfloor] \times [\lceil 2^{c \cdot n/\lambda} \rceil]$ and $R : \mathcal{V}^n \times [\lceil 2^{c \cdot n/\lambda} \rceil] \to [\lfloor 2^{n/\lambda} \rfloor]$ such that, when (U_i, V_i) is i.i.d. according to p_{UV} for all $i \in [n]$,*

$(K, M) \leftarrow S(U^n)$ and $L \leftarrow R(V^n, M)$,

$$\Pr[K \neq L] \leq \epsilon, \tag{4}$$

$$\sum_{k \in [\lfloor 2^{n/\lambda} \rfloor]} \left| \Pr[K = k] - \frac{1}{\lfloor 2^{n/\lambda} \rfloor} \right| \leq \epsilon. \tag{5}$$

Proof. To prove this lemma, we once again invoke a result from [4]. The common randomness capacity $C(r)$ of correlation (U, V) for communication rate of $r \geq 0$ from sender to receiver is defined as the supremum of all $s \geq 0$ such that for each $\epsilon > 0$ and all sufficiently large n, there exist functions $S : \mathcal{U}^n \to [\lfloor 2^{s \cdot n} \rfloor] \times [\lceil 2^{r \cdot n} \rceil]$ and $R : \mathcal{V}^n \times [2^{r \cdot n}] \to [\lfloor 2^{s \cdot n} \rfloor]$ such that, Eqs. (4) to (5) are satisfied for $\lambda = \frac{1}{s}$. [50, Theorem 3] showed that

$$\lim_{r \downarrow 0} \frac{C(r)}{r} = \inf_{p(u|x)} \frac{I(U; X)}{I(U; X) - I(U; Y)} = \frac{1}{1 - \mathrm{S}^*(U, V)}. \tag{6}$$

Hence, by basic real analysis, for any $c > 1 - \mathrm{S}^*(U, V)$, there exists $r > 0$ such that $\frac{C(r)}{r} > \frac{1}{c}$.

Define $s = \frac{r}{c}$; since $C(r) > s$, by definition of $C(r)$, for each $\epsilon > 0$ and all sufficiently large n, there exist functions $S : \mathcal{U}^n \to [\lfloor 2^{s \cdot n} \rfloor] \times [\lceil 2^{c \cdot s \cdot n} \rceil]$ and $R : \mathcal{V}^n \times [2^{c \cdot s \cdot n}] \to [\lfloor 2^{s \cdot n} \rfloor]$ such that, Eqs. (4) to (5) are satisfied for $\lambda = \frac{1}{s}$. The proof follows by taking $\lambda = \frac{1}{s}$. □

Lemma 6.4. *For any correlation (U, V), suppose $\lambda > 0$ and $c > 0$ are such that, for each $\epsilon > 0$ and for all sufficiently large n, there are functions $S : \mathcal{U}^n \to [\lfloor 2^{n/\lambda} \rfloor] \times [\lceil 2^{c \cdot n/\lambda} \rceil]$ and $R : \mathcal{V}^n \times [\lceil 2^{c \cdot n/\lambda} \rceil] \to [\lfloor 2^{n/\lambda} \rfloor]$ such that, when (U_i, V_i) is i.i.d according to p_{UV} for all $i \in [n]$, $(K, M) \leftarrow S(U^n)$ and $L \leftarrow R(V^n, M)$,*

$$\Pr[K \neq L] \leq \epsilon, \tag{7}$$

$$\sum_{k \in [\lfloor 2^{n/\lambda} \rfloor]} \left| \Pr[K = k] - \frac{1}{\lfloor 2^{n/\lambda} \rfloor} \right| \leq \epsilon. \tag{8}$$

If $c < 1 - \mathrm{S}^*(X, Y)$, then the amortized communication cost of OMR converting (X, Y) to (U, V) is at least $\frac{1}{\lambda}(1 - \mathrm{S}^*(X, Y) - c)$.

Proof. Suppose an amortized communication cost $c' > 0$ is achievable for OMR converting (X, Y) to (U, V). Then, for each $i \in \mathbb{N}$, there exist functions $S'_i : \mathcal{X}^{m_i} \to \mathcal{U}^{n_i} \times [2^{c \cdot n_i}]$ and $R'_i : \mathcal{Y}^{m_i} \times [2^{c \cdot n_i}] \to \mathcal{V}^{n_i}$ such that, for each i, when (X^{m_i}, Y^{m_i}) is i.i.d. according to p_{XY}, (U^{n_i}, V^{n_i}) is i.i.d. according to p_{UV}, $(\hat{U}^{n_i}, M) \leftarrow S'_i(X^{m_i})$ and $\hat{V}^{n_i} \leftarrow R'_i(V^{n_i}, M)$,

$$\left(\hat{U}^{n_i}, \hat{V}^{n_i}\right) \approx^{\epsilon_i} (U^{n_i}, V^{n_i}), \tag{9}$$

where, $\epsilon_i \to 0$ as $i \to \infty$.

We can choose i such that 1) S'_i, R'_i in the aforementioned sequence satisfies Eq. (9) with $\epsilon_i \leq \epsilon$, and 2) there exist $S_i : \mathcal{U}^{n_i} \to [\lfloor 2^{n_i/\lambda} \rfloor] \times [\lceil 2^{c \cdot n_i/\lambda} \rceil]$ and $R_i : \mathcal{V}^{n_i} \times [\lceil 2^{c \cdot n_i/\lambda} \rceil] \to [\lfloor 2^{n_i/\lambda} \rfloor]$ such that, Eqs. (7) to (8) are satisfied for $n = n_i$.

We construct an OMR converting (X, Y) to common randomness as follows:

1. Using (X^{m_i}, Y^{m_i}) and $c' \cdot n_i$ bits of communication, generate $(\hat{U}^{n_i}, \hat{V}^{n_i})$, where $(\hat{U}^{n_i}, M) \leftarrow S'_i(X^{m_i})$ and $\hat{V}^{n_i} \leftarrow R'_i(V^{n_i}, M)$.

2. Using $(\hat{U}^{n_i}, \hat{V}^{n_i})$ and $c \cdot n_i/\lambda$ bits of communication, generate (K, L), where $(K, M) \leftarrow S(U^n)$ and $L \leftarrow R(V^n, M)$.

By Eqs. (7) to (9) and the data processing inequality of total variation distance,

$$\Pr[K \neq L] \leq \epsilon + \epsilon_i \leq 2\epsilon,$$
$$\sum_{k \in [\lfloor 2^{n_i/\lambda} \rfloor]} \left| \Pr[K = k] - \frac{1}{\lfloor 2^{n_i/\lambda} \rfloor} \right| \leq \epsilon + \epsilon_i \leq 2\epsilon.$$

Thus, we have established that, there is a 2ϵ-OMR converting m_i copies of (X, Y) to $n_i/\lambda - 1$ bits of common randomness using $n_i(c'/\lambda + c) + 1$ bits of communication. Hence, an amortized communication cost of $c + \lambda c'$ is achievable for OMR converting (X, Y) to common randomness. But then, by Lemma 6.2,

$$c + \lambda c' \geq 1 - \mathrm{S}^*(X, Y) \implies c' \geq \frac{1}{\lambda}(1 - \mathrm{S}^*(X, Y) - c).$$

This proves the lemma. □

Proof of Theorem 6.1. Choose c such that $1 - \mathrm{S}^*(U, V) < c < 1 - \mathrm{S}^*(X, Y)$. Invoking Lemma 6.3, we get λ that satisfies the conditions in Lemma 6.4. Hence, by Lemma 6.4, since $c < 1 - \mathrm{S}^*(X, Y)$, the amortized communication cost of OMR converting (X, Y) to (U, V) is at least $\frac{1}{\lambda}(1 - \mathrm{S}^*(X, Y) - c)$. This proves the theorem.

The above approach for lower bound can be used to get concrete lower bounds for OMR between specific correlations. For this, we first calculate S^* of some correlations of interest.

Lemma 6.5. *For any finite group $\mathbb{G}$ and $k \in \mathbb{N}$, S^* of 1-out-of-k OT correlation over $\mathbb{G}$ is $\frac{1}{k}$.*

Using an information theoretic argument, we prove that $I(U; Y)/I(U; X) \leq \frac{1}{k}$ for any U generated from X. Taking $U = X$ we achieve this upper bound implying the lemma. A full proof of this lemma is provided in Appendix C.2.

In a similar manner, we prove in Appendix C.3 the next lemma.

Lemma 6.6. *For any finite field $\mathbb{F}$, S^* of OLE correlation over $\mathbb{F}$ is $\frac{1}{2}$.*

We prove the following lemma in Appendix C.4.

Lemma 6.7. *For any group $\mathbb{G}$, S^* of the additive correlation over $\mathbb{G}$ is strictly larger than $\frac{1}{2}$, if $H(\psi) < \frac{\log|\mathbb{G}|}{2}$ (where ψ is as described in the definition of additive correlation in Section 4.3),*

Using the above characterizations of S^* and Lemma 6.4, we get the following lower bounds on statistical OMR which justify our constructions in Section 5.

Corollary 6.7.1 (Corollary of Lemma 6.4 and Lemma 6.5). *The amortized communication cost of statistical OMR converting 1-out-of-k OT correlation over any group $\mathbb{G}$ to $(2, q)$ correlation is at least $\log(\frac{q}{2})/2$.*

We can construct a simple OMR converting $(2, q)$ correlation to CR, by simply sending x_0, and both outputting $x_0||r_0$. Thus we achieve the following functions $S : \mathcal{U}^n \to [\lfloor 2^{n/\lambda} \rfloor] \times [\lceil 2^{c \cdot n/\lambda} \rceil]$ and $R : \mathcal{V}^n \times [\lceil 2^{c \cdot n/\lambda} \rceil] \to [\lfloor 2^{n/\lambda} \rfloor]$ with $\lambda = c = \frac{1}{\log(2q)}$. Now, Lemma 6.4 with $\mathrm{S}^*(X, Y) = \frac{1}{k}$ gives us the lower bound of $\frac{k-1}{k} \log(2q) - 1$. This is an increasing function of k. Plugging $k = 2$, will give us the desirable $\log(\frac{q}{2})/2$ lower bound.

Remark. The same lower bound can be given for source correlation OLE due to Lemma 6.6.

Corollary 6.7.2. *The optimal (in terms of amortized communication cost) protocol for* $(2, 3)$ *correlation using OT will be with* 1*-out-of-*2 *OT.*

From the proof of Corollary 6.7.1, we know that the amortized communication cost using source correlation 1-out-of-k OT, with $k \geq 3$, we get a lower bound of $\frac{2}{3} \log(6) - 1 \approx 0.72$. Since in Corollary 5.2.1 we achieve an OMSR with amortized communication cost $\log(3/2) < 0.72$ using 1-out-of-2 OT, we can infer that the optimal OMSR will also need to have the source correlation 1-out-of-2 OT.

Corollary 6.7.3 (Corollary of Lemma 6.4 and Lemma 6.5)**.** *The amortized communication cost of OMR converting* 1*-out-of-*3 *OT correlation over any group* $\mathbb{G}$ *to* $(3, 2)$ *correlation is at least* $\log(\frac{16}{3})/3$*.*

We can construct a simple OMR converting (3,2) correlation to CR, by simply sending x_0, and both outputting $x_0||u_0||v_0$. In a similar manner as in Corollary 6.7.1 we will achieve a lower bound of $\log(16/3)/3$.

Corollary 6.7.4 (Corollary of Theorem 6.1, Lemma 6.5 and Lemma 6.7)**.** *The amortized communication cost of OMR converting* 1*-out-of-*2 *OT (or OLE) correlation over any group* $\mathbb{G}_1$ *to an additive correlation over a group* $\mathbb{G}_2$ *where* $H(\psi) < \frac{\log |\mathbb{G}_2|}{2}$ *(where* ψ *is as described in the definition of additive correlation in Section 4.3) is strictly greater than* 0*.*

This simple corollary is due to the fact that S^* of the additive correlation is strictly greater than $\frac{1}{2}$.

6.2 Linear lower bound for interactive reductions

Our next result establishes a much more general lower bound albeit for a more restricted class of correlations. We show that generating an n-bit vector correlation using any correlation (X, Y) such that $\mathrm{S}^*(X, Y) < 1$ requires $\Omega(n)$ bits of *interactive communication.* We note that, as intuition suggests, $\mathrm{S}^*(X, Y) < 1$ if and only if the correlation lacks common randomness. We refer to Appendix E for a detailed discussion. Observe that we are interested in generating a single copy of a correlation of length n from a class of correlations (for increasing value of n). As n approaches infinity, S^* of n-bit vector correlation approaches 1; hence, in the case of OMR, the linear lower bound on communication cost is intuitively implied by our previous result. The following theorem makes a strictly stronger claim.

Theorem 6.8. *Let* π *be an interactive protocol between* S *and* R *using (arbitrarily many copies of) correlation* (X, Y) *and* ℓ *bits of communication which computes* n*-bit unit vector correlation with* $\epsilon \leq \frac{1}{6}$ *error. That is,* S *and* R *output* $\hat{U}$ *and* $\hat{V}$*, respectively, where* $(\hat{U}, \hat{V})$ *is* ϵ *far from being an* n*-bit unit vector correlation in total variation distance. If* $\mathrm{S}^*(X, Y) < 1$*, there exists* $1 < q < p$ *that depend only on the description of distribution* p_{XY} *such that* $\ell \geq \frac{n}{2}(\frac{p-q}{pq}) - \frac{1}{2}\log(n) - 1$*.*

We state a couple of lemmas that will be used to prove the theorem.

Lemma 6.9. *Let π be an interactive protocol between S and R using correlation (X, Y) and ℓ bits of communication in which S and R output $\hat{U}$ and $\hat{V}$, respectively, where $\hat{U}$ and $\hat{V}$ are uniformly distributed over $\{0,1\}^n$ and $(\hat{U}, \hat{V}) \approx^{\epsilon} (U, V)$, where (U, V) is an n-bit unit vector correlation.*

Using π, we can construct $f : \mathcal{X} \to \{0,1\}^n$ and $g : \mathcal{Y} \to \{0,1\}^n$ such that, $f(X)$ and $g(Y)$ are uniformly distributed over $\{0,1\}^n$, and

$$\Pr\left[f(X) = g(Y)\right] \geq \frac{4^{-\ell}}{32n}(1-\epsilon)^3.$$

Lemma 6.10. *Let (X_i, Y_i) be i.i.d. p_{XY} for each $i \in [m]$. Let $f : \mathcal{X} \to \{0,1\}^n$ and $g : \mathcal{Y} \to \{0,1\}^n$ be any pair of functions such that $f(X)$ and $g(Y)$ are uniformly distributed over $\{0,1\}^n$. If $\mathrm{S}^*(X, Y) < 1$, there exist $1 < q < p$ that depend only on the description of p_{XY} such that*

$$\Pr\left[f(X) = g(Y)\right] \leq 2^{-n\left(\frac{p-q}{pq}\right)}.$$

Proof of Lemma 6.9 follows the same blueprint as the proof of [18, Theorem 2.6]. We provide the proof in Appendix C.5. Before proving Lemma 6.10, we show how they imply the theorem.

Proof of Theorem 6.8 If $\hat{U}$ and $\hat{V}$ are not uniformly distributed over $\{0,1\}^n$, we transform π into π' which outputs correlation $(\tilde{U}, \tilde{V})$ with uniform marginals as follows: Let $Sup \subset \{0,1\}^n$, where $u \in Sup$ if $\Pr\left[\hat{U} = u\right] < \frac{1}{2^n}$. Let W be a random variable over the domain Sup s.t. $\Pr\left[W = u\right]$ is proportional to $\frac{1}{2^n} - \Pr\left[\hat{U} = u\right]$. Now, S on receiving $u \notin Sup$, w.p. $\frac{1}{\Pr\left[\hat{U}=u\right]}\frac{1}{2^n}$, S outputs u, and with the remainder probability, he outputs a sample from W. R outputs $\tilde{V}$ which is sampled analogously. It is easy to see that $\tilde{U}$ and $\tilde{V}$ are distributed uniformly over $\{0,1\}^n$. Since $(\hat{U}, \hat{V})$ is ϵ far from n-bit unit vector correlation in total variation distance,

$$\sum_{u \in \{0,1\}^n, \Pr\left[\hat{U}=u\right] > \frac{1}{2^n}} \Pr\left[\hat{U} = u\right] - \frac{1}{2^n} \leq \epsilon.$$

A similar condition holds for $\hat{V}$. By a union bound, $\Pr\left[\tilde{U} = \hat{U}, \tilde{V} = \hat{V}\right] \geq 1 - 2\epsilon$. From this it follows that, when (U, V) is distributed according to n-bit unit vector correlation,

$$\mathsf{TVD}\left((\tilde{U}, \tilde{V}), (U, V)\right) \leq (1 - 2\epsilon)\mathsf{TVD}\left((\hat{U}, \hat{V}), (U, V)\right) + 2\epsilon \leq 3\epsilon.$$

Since $(\tilde{U}, \tilde{V})$ has uniform marginals, invoking Lemma 6.9 with π', we obtain f, g such that $f(X^m)$ and $g(Y^m)$ are uniform in $\{0,1\}^n$ and

$$\Pr\left[f(X^m) = g(Y^m)\right] \geq \frac{4^{-\ell}}{32n}(1 - 3\epsilon)^3.$$

However, by Lemma 6.10,

$$\Pr\left[f(X^m) = g(Y^m)\right] \leq 2^{-n\left(\frac{p-q}{pq}\right)}.$$

The above inequalities together imply that $2(\ell + 1) \geq n(\frac{p-q}{pq}) - \log(n)$ when $\epsilon < \frac{1}{6}$, proving the theorem.

We conclude by proving Lemma 6.10.

Proof of Lemma 6.10. This lemma is proved along the lines of [10, Theorem 1]. Bogdanov and Mossel's showed in [10, Theorem 1] that deriving n bits of common randomness using a source of noisy common randomness and zero communication succeeds only with $2^{-\Omega(n)}$ probability. A noisy common randomness correlation with ϵ crossover probability is one in which the sender gets a random bit and the receiver gets the same bit with probability $1-\epsilon$ and its complement with the remaining probability. Their approach used a version of hypercontractive inequality for noisy common randomness distribution. The following argument is a generalization of their argument. For every $z \in \{0,1\}^n$, define $f_z : \mathcal{X}^m \to \{0,1\}$ and $g_z : \mathcal{Y}^m \to \{0,1\}$ as

$$f_z(x^m) = \begin{cases} 1, \text{if } f(x^m) = z \\ 0, \text{otherwise,} \end{cases} \qquad g_z(y^m) = \begin{cases} 1, \text{if } g(y^m) = z \\ 0, \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned} \Pr\left[f(X^m) = g(Y^m)\right] &= \sum_{z \in \{0,1\}^n} \Pr\left[f(X^m) = z \wedge g(Y^m) = z\right] \\ &= \sum_{z \in \{0,1\}^n} \mathbb{E}_{X^m,Y^m}[f_z(X^m) \cdot g_z(Y^m)]. \end{aligned} \tag{10}$$

If $\mathrm{S}^*(X,Y) < 1$, there exist $1 < q < p$ such that, defining $p' = \frac{p}{p-1}$, for any pair of functions $f : \mathcal{X}^m \to \mathbb{R}$ and $g : \mathcal{Y}^m \to \mathbb{R}$,

$$\mathbb{E}_{X^m,Y^m}[f(X^m) \cdot g(Y^m)] \le \left(\mathbb{E}_{X^m}[f^{p'}(X^m)]\right)^{\frac{1}{p'}} \left(\mathbb{E}_{Y^m}[g^q(Y^m)]\right)^{\frac{1}{q}}.$$

This follows almost immediately from known results; a formal proof is provided in Appendix F. Using this in Eq. (10), and noting that f_z and g_z are Boolean functions,

$$\begin{aligned} &\sum_{z \in \{0,1\}^n} \mathbb{E}_{X^m,Y^m}[f_z(X^m) \cdot g_z(Y^m)] \\ &\le \sum_{z \in \{0,1\}^n} \left(\mathbb{E}_{X^m}[f_z(X^m)]\right)^{\frac{1}{p'}} \left(\mathbb{E}_{Y^m}[g_z(Y^m)]\right)^{\frac{1}{q}} \\ &\le \left(\sum_{z \in \{0,1\}^n} \left(\mathbb{E}_{X^m}[f_z(X^m)]\right)^{\frac{p'}{p'}}\right)^{\frac{1}{p'}} \left(\sum_{z \in \{0,1\}^n} \left(\mathbb{E}_{Y^m}[g_z(Y^m)]\right)^{\frac{p}{q}}\right)^{\frac{1}{p}}. \end{aligned} \tag{11}$$

Here, the last inequality used Holder's inequality:

$$\sum_{i=1}^k |a_i| \cdot |b_i| \le \left(\sum_{i=1}^k |a_i|^{p'}\right)^{\frac{1}{p'}} \left(\sum_{i=1}^k |b_i|^p\right)^{\frac{1}{p}}.$$

Since $f(X^m)$ and $g(Y^m)$ are uniform over $\{0,1\}^n$, for all $z \in \{0,1\}^n$,

$$\mathbb{E}_{X^m}[f_z(X^m)] = 2^{-n} \qquad \mathbb{E}_{Y^m}[g(Y^m)] = 2^{-n}.$$

Hence, $\sum_{z\in\{0,1\}^n} \left(\mathbb{E}_{X^m}[f_z(X^m)]\right) = 1$, and since $p > q$,

$$\sum_{z\in\{0,1\}^n} \left(\mathbb{E}_{Y^m}[g_z(Y^m)]\right)^{\frac{p}{q}} = \sum_{z\in\{0,1\}^n} 2^{-n} \left(\mathbb{E}_{Y^m}[g_z(Y^m)]\right)^{\frac{p}{q}-1} = 2^{-n\left(\frac{p}{q}-1\right)}.$$

Using these observations in Eq. (11), we get $\Pr\left[f(X^m) = g(Y^m)\right] \leq 2^{-n\left(\frac{p-q}{pq}\right)}$. □

Remark. It is clear from the proof of Lemma 6.9 that n-bit unit vector correlation can be replaced by any correlation that can be converted to common randomness with n-bits of common randomness with success probability that is inverse polynomial in n. Hence, the result in Theorem 6.8 holds more generally for families of correlations with this property; this includes 1-out-of-$k(n)$ OT correlations of string length n when $k(n)$ is polynomial in n.

7 Role of Common Randomness in OMSR

In general, common randomness does not aid in OMSR whenever the target correlation does not have inherent common randomness. Correlation (U, V) is said to have non-trivial common randomness if it can be represented as $((W, U'), (W, V'))$ where W has non-zero entropy. In other words, there are deterministic functions f and g such that $f(U) = g(V)$ with probability 1 and $H(f(U)) > 0$. All (target) correlations considered in this work lack common information; more generally, this holds for most correlations with cryptographic applications because, intuitively, common information does not enable cryptographic tasks. In the case of perfect OMSR, by simply conditioning on any realization of common randomness we get a perfect OMSR without common randomness setup. However, for statistical OMSR, such a restriction need not necessarily be secure. The following theorem is proved by showing the existence of a realization of common randomness such that, restricted to this realization the OMSR still guarantees comparable security. A concrete consequence of the following theorem is that common randomness does not aid in statistical OMSR with negligible error for target correlations without common randomness. A proof of the theorem is provided in Appendix D.

Theorem 7.1. *Suppose there exists an ϵ-OMSR $\langle S, R\rangle$ converting a correlation (X, Y) to n copies of a target correlation (U, V) using common randomness Q. If (U, V) lacks common information, there exists an $O(n^2\sqrt{\epsilon})$-OMSR converting correlation (X, Y) to n copies of a target correlation (U, V) with the same cost without using common randomness.*

Acknowledgements

We thank the anonymous reviewers for helpful comments. M. Kelkar was partially supported by a Technion research scholarship. Y. Ishai and V. Narayanan were supported by ERC Project NTSC (742754), and ISF grant 2774/20. Y. Ishai was additionally supported by BSF grant 2018393.

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, pp. 797–827.

[2] Shweta Agrawal, Yuval Ishai, Eyal Kushilevitz, Varun Narayanan, Manoj Prabhakaran, Vinod M. Prabhakaran, and Alon Rosen. “Cryptography from One-Way Communication: On Completeness of Finite Channels”. In: *ASIACRYPT*. 2020, pp. 653–685.

[3] Shweta Agrawal, Yuval Ishai, Eyal Kushilevitz, Varun Narayanan, Manoj Prabhakaran, Vinod M. Prabhakaran, and Alon Rosen. “Secure Computation from One-Way Noisy Communication, or: Anti-correlation via Anti-concentration”. In: *CRYPTO*. 2021, pp. 124–154.

[4] R. Ahlswede and I. Csiszar. “Common randomness in information theory and cryptography. II. CR capacity”. In: *IEEE Transactions on Information Theory* 44.1 (1998), pp. 225–240.

[5] Venkat Anantharam, Amin Aminzadeh Gohari, Sudeep Kamath, and Chandra Nair. *On Maximal Correlation, Hypercontractivity, and the Data Processing Inequality studied by Erkip and Cover*. 2013. arXiv: `1304.6133`.

[6] Donald Beaver. “Efficient multiparty protocols using circuit randomization”. In: *CRYPTO*. 1991, pp. 420–432.

[7] Rikke Bendlin, Ivan Damgård, Claudio Orlandi, and Sarah Zakarias. *Semi-Homomorphic Encryption and Multiparty Computation*. Cryptology ePrint Archive, Paper 2010/514. 2010.

[8] Kaartik Bhushan, Ankit Kumar Misra, Varun Narayanan, and Manoj Prabhakaran. “Secure Non-Interactive Reducibility is Decidable”. In: *TCC*. 2022, pp. 408–437.

[9] Manuel Blum, Paul Feldman, and Silvio Micali. “Non-Interactive Zero-Knowledge and Its Applications (Extended Abstract)”. In: *STOC*. 1988, pp. 103–112.

[10] Andrej Bogdanov and Elchanan Mossel. “On Extracting Common Random Bits From Correlated Sources”. In: *IEEE Transactions on Information Theory* 57.10 (2011), pp. 6351–6355.

[11] Elette Boyle, Nishanth Chandran, Niv Gilboa, Divya Gupta, Yuval Ishai, Nishant Kumar, and Mayank Rathee. “Function Secret Sharing for Mixed-Mode and Fixed-Point Secure Computation”. In: *EUROCRYPT*. 2021, pp. 871–900.

[12] Elette Boyle, Geoffroy Couteau, Niv Gilboa, and Yuval Ishai. “Compressing Vector OLE”. In: *ACM CCS*. 2018, pp. 896–912.

[13] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, Nicolas Resch, and Peter Scholl. “Correlated Pseudorandomness from Expand-Accumulate Codes”. In: *CRYPTO*. 2022, pp. 603–633.

[14] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, Peter Rindal, and Peter Scholl. “Efficient Two-Round OT Extension and Silent Non-Interactive Secure Computation”. In: *ACM CCS*. 2019, pp. 291–308.

[15] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. “Efficient Pseudorandom Correlation Generators from Ring-LPN”. In: *CRYPTO*. 2020, pp. 387–416.

[16] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. “Efficient Pseudorandom Correlation Generators: Silent OT Extension and More”. In: *CRYPTO*. 2019, pp. 489–518.

[17] Elette Boyle, Niv Gilboa, and Yuval Ishai. “Secure Computation with Preprocessing via Function Secret Sharing”. In: *TCC*. 2019, pp. 341–371.

[18] Clement Louis Canonne, Venkatesan Guruswami, Raghu Meka, and Madhu Sudan. “Communication with Imperfectly Shared Randomness”. In: *ITCS*. 2015, 257–262.

[19] Geoffroy Couteau, Peter Rindal, and Srinivasan Raghuraman. “Silver: Silent VOLE and Oblivious Transfer from Hardness of Decoding Structured LDPC Codes”. In: *CRYPTO*. 2021, pp. 502–534.

[20] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2006.

[21] Ivan Damgård, Valerio Pastro, Nigel Smart, and Sarah Zakarias. “Multiparty computation from somewhat homomorphic encryption”. In: *CRYPTO*. 2012, pp. 643–662.

[22] Anindya De, Elchanan Mossel, and Joe Neeman. “Non interactive simulation of correlated distributions is decidable”. In: *SODA*. 2018, pp. 2728–2746.

[23] Daniel Demmler, Thomas Schneider, and Michael Zohner. “ABY - A Framework for Efficient Mixed-Protocol Secure Two-Party Computation”. In: *NDSS*. 2015.

[24] Itai Dinur, Steven Goldfeder, Tzipora Halevi, Yuval Ishai, Mahimna Kelkar, Vivek Sharma, and Greg Zaverucha. “MPC-Friendly Symmetric Cryptography from Alternating Moduli: Candidates, Protocols, and Applications”. In: *CRYPTO*. 2021, pp. 517–547.

[25] Jack Doerner and abhi shelat. “Scaling ORAM for Secure Computation”. In: *ACM CCS*. 2017, pp. 523–535.

[26] P. Gács and J. Körner. “Common information is far less than mutual information”. In: *Problems of Control and Information Theory* 2.2 (1973), pp. 149–162.

[27] Sanjam Garg, Yuval Ishai, Eyal Kushilevitz, Rafail Ostrovsky, and Amit Sahai. “Cryptography with one-way communication”. In: *CRYPTO*. 2015, pp. 191–208.

[28] Badih Ghazi and T. S. Jayram. “Resource-Efficient Common Randomness and Secret-Key Schemes”. In: *SODA*. 2018, pp. 1834–1853.

[29] Badih Ghazi, Pritish Kamath, and Madhu Sudan. “Decidability of Non-interactive Simulation of Joint Distributions”. In: *FOCS*. 2016, pp. 545–554.

[30] Oded Goldreich, Silvio Micali, and Avi Wigderson. “How to play ANY mental game”. In: *STOC*. 1987, pp. 218–229.

[31] Saumya Goyal, Varun Narayanan, and Manoj Prabhakaran. “Oblivious-Transfer Complexity of Noisy Coin-Toss via Secure Zero Communication Reductions”. In: *TCC*. 2022, pp. 89–118.

[32] Venkatesan Guruswami and Jaikumar Radhakrishnan. “Tight bounds for communication-assisted agreement distillation”. In: *CCC*. 2016, pp. 1–17.

[33] Yuval Ishai and Eyal Kushilevitz. “Randomizing polynomials: A new representation with applications to round-efficient secure computation”. In: *FOCS*. 2000, pp. 294–304.

[34] Yuval Ishai, Manoj Prabhakaran, and Amit Sahai. "Founding cryptography on oblivious transfer–efficiently". In: *CRYPTO*. 2008, pp. 572–591.

[35] Sudeep Kamath and Venkat Anantharam. "On Non-Interactive Simulation of Joint Distributions". In: *IEEE Trans. Inf. Theory* 62.6 (2016), pp. 3419–3435.

[36] Hamidreza Amini Khorasgani, Hemanta K. Maji, and Hai H. Nguyen. "Secure Non-interactive Simulation: Feasibility and Rate". In: *EUROCRYPT*. 2022, pp. 767–796.

[37] Hamidreza Amini Khorasgani, Hemanta K. Maji, and Hai H. Nguyen. "Secure Non-interactive Simulation from Arbitrary Joint Distributions". In: *TCC*. 2022, pp. 378–407.

[38] Joe Kilian. "Founding crytpography on oblivious transfer". In: *STOC*. 1988, pp. 20–31.

[39] Varun Narayanan, Manoj Prabhakaran, and Vinod M. Prabhakaran. "Zero-Communication Reductions". In: *TCC*. 2020, pp. 274–304.

[40] Théo Ryffel, Pierre Tholoniat, David Pointcheval, and Francis R. Bach. "AriaNN: Low-Interaction Privacy-Preserving Deep Learning via Function Secret Sharing". In: *Proc. Priv. Enhancing Technol.* 2022.1 (2022), pp. 291–316.

[41] Kyle Storrier, Adithya Vadapalli, Allan Lyons, and Ryan Henry. *Grotto: Screaming fast* $(2+1)$*-PC for* $\mathbb{Z}_{2^n}$ *via (2, 2)-DPFs*. Cryptology ePrint Archive, Paper 2023/108. 2023.

[42] Madhu Sudan, Himanshu Tyagi, and Shun Watanabe. "Communication for Generating Correlation: A Unifying Survey". In: *IEEE Trans. Inf. Theory* 66.1 (2020), pp. 5–37.

[43] Sameer Wagh. "Pika: Secure Computation using Function Secret Sharing over Rings". In: *Proc. Priv. Enhancing Technol.* 2022.4 (2022), pp. 351–377.

[44] H. S. Witsenhausen. "On Sequences of Pairs of Dependent Random Variables". In: *SIAM Journal on Applied Mathematics* 28.1 (1975), pp. 100–113.

[45] Aaron D Wyner. "The wire-tap channel". In: *Bell system technical journal* 54.8 (1975), pp. 1355–1387.

[46] Kang Yang, Chenkai Weng, Xiao Lan, Jiang Zhang, and Xiao Wang. "Ferret: Fast Extension for Correlated OT with Small Communication". In: *CCS*. 2020, 1607–1626.

[47] Peng Yang, Zoe L. Jiang, Shiqi Gao, Jiehang Zhuang, Hongxiao Wang, Junbin Fang, Siuming Yiu, and Yulin Wu. *FssNN: Communication-Efficient Secure Neural Network Training via Function Secret Sharing*. Cryptology ePrint Archive, Paper 2023/073. 2023.

[48] Andrew C. Yao. "Protocols for secure computations". In: *SFCS*. 1982, pp. 160–164.

[49] Andrew Chi Yao. "How to generate and exchange secrets". In: *SFCS*. 1986, pp. 162–167.

[50] Lei Zhao and Yeow-Kiang Chia. "The efficiency of common randomness generation". In: *Allerton*. 2011, pp. 944–950.

A Details omitted from Section 4

Definition A.1 (Simulator based security for OMSR)**.** An ϵ-error one-message secure reduction (ϵ-OMSR) over $(m, p_{XY}, n, p_{UV}, l)$ is an ϵ-OMR $\langle S, R\rangle$ (for converting m copies of the source correlation p_{XY} to n copies of the target correlation p_{UV} using l bits of communication) which also satisfies the following security properties:

Privacy against S**.** Let $(X_i, Y_i)_{i\in[m]} \overset{\text{i.i.d.}}{\sim} p_{XY}$ and Q, Q' distributed uniformly over $\mathcal{Q}$ and $\mathcal{Q}'$ are the private randomness of S and R, respectively. Let $(\hat{U}^n, M) \leftarrow S(X^m, Q)$, $\hat{V}^n \leftarrow R(M, Y^m, Q')$ be the outputs for the sender and the receiver respectively. There exists a randomized simulator $\mathcal{S}_S : \mathcal{U}^n \to \mathcal{X}^m \times \mathcal{Q}$ such that,

$$\left(X^m, Q, \hat{V}^n\right) \approx^\epsilon (\mathcal{S}_S(U^n), V^n)\,. \tag{12}$$

Privacy against R**.** There exists a randomized simulator $\mathcal{S}_R : \mathcal{V}^n \to \mathcal{Y}^m \times \mathcal{M} \times \mathcal{Q}'$ such that,

$$\left(\hat{U}^n, Y^m, M, Q'\right) \approx^\epsilon (U^n, \mathcal{S}_R(V^n))\,. \tag{13}$$

We say that an OMSR is *perfect* if $\epsilon = 0$.

Security conditions in Definition 4.2 imply simulators with comparable security. Consider a simulator $\mathcal{S}_S$ such that for all $u^n \in U^n$,

$$\Pr\left[\mathcal{S}_S(u^n) = (x^m, q)\right] = \Pr\left[X^m = x^m, Q = q | \hat{U}^n = u^n\right]$$

From correctness, we have

$$\left(\mathcal{S}_S(\hat{U}^n), \hat{V}^n\right) \approx^\epsilon (\mathcal{S}_S(U^n), V^n)\,.$$

Consider the joint distribution $(X^m, Y^m, Q, Q', \hat{U}^n, M, \hat{V}^n)$ induced by OMSR, and the joint distribution $(\hat{U}^n, \hat{V}^n, \mathcal{S}_S(U^n)$ induced by the simulation (when fed $\hat{U}^n$). Since x^m and q completely determine $\hat{u}^n$, both in the real execution and the simulation,

$$\begin{aligned}
&\mathsf{TVD}\left[\left(X^m, Q, \hat{V}^n\right), \left(\mathcal{S}_S(\hat{U}^n), \hat{V}^n\right)\right] \\
&= \sum_{x^m, q, \hat{u}^n, \hat{v}^n} \left| \Pr_{X^m, Q, \hat{U}^n, \hat{V}^n}(x^m, q, \hat{u}^n, \hat{v}^n) - \Pr_{\mathcal{S}_S(\hat{U}^n), \hat{U}^n, \hat{V}^n}(x^m, q, \hat{u}^n, \hat{v}^n)\right| \\
&= \sum_{x^m, q, \hat{u}^n, \hat{v}^n} \Big| \Pr_{\hat{U}^n}(\hat{u}^n) \Pr_{X^m Q|\hat{U}^n}(x^m, q|\hat{u}^n) \Pr_{\hat{V}^n|X^m Q\hat{U}^n}(\hat{v}^n|x^m, q, \hat{u}^n) \\
&\qquad\qquad - \Pr_{\hat{U}^n}(\hat{u}^n) \Pr_{X^m Q|\hat{U}^n}(x^m, q|\hat{u}^n) \Pr_{\hat{V}^n|\hat{U}^n}(\hat{v}^n|\hat{u}^n)\Big| \\
&= \sum_{x^m, q, \hat{u}^n, \hat{v}^n} \Pr_{\hat{U}^n}(\hat{u}^n) \Pr_{X^m Q|\hat{U}^n}(x^m, q|\hat{u}^n) \left| \Pr_{\hat{V}^n|X^m Q\hat{U}^n}(\hat{v}^n|x^m, q, \hat{u}^n) - \Pr_{\hat{V}^n|\hat{U}^n}(\hat{v}^n|\hat{u}^n)\right| \\
&= \mathbb{E}_{X^m Q}\mathsf{TVD}[\hat{V}^n|X^m Q\hat{U}^n, \hat{V}^n|\hat{U}^n] < \epsilon.
\end{aligned}$$

The other direction is true as well; simulators imply the security conditions in Definition 4.2 with comparable security. For simplicity, we restrict the analysis to senders that do not use private randomness. The case of randomized senders can be proved similarly by introducing private randomness as a random variable in the below argument.

Note that, since $(\tilde{X}^m, \tilde{U}^n, V^n) \approx^{\epsilon} (X^m, \hat{U}^n, \hat{V}^n)$, there exists a distribution $P_{X^m,\hat{U}^n,\hat{V}^n,\tilde{X}^m,\tilde{U}^n,V^n}$ such that $Pr[X^m\hat{U}^n\hat{V}^n \neq \tilde{X}^m\tilde{U}^m V^m] \leq \epsilon$. Now, let A be an indicator of the event of $X^m\hat{U}^n\hat{V}^n = \tilde{X}^m\tilde{U}^n V^n$.

$$\begin{aligned}
&\mathbb{E}_{X^m}\mathsf{TVD}[\hat{V}^n|X^m\hat{U}^n, \hat{V}^n|\hat{U}^n] \\
&= \sum_{x^m} \Pr_{X^m}(x^m) \sum_{v^n} \Big| \Pr_{\hat{V}^n|X^m}(v^n|x^m) - \Pr_{\hat{V}^n|\hat{U}^n}(v^n|u^n) \Big| \\
&= \sum_{x^m} \Pr_{X^m}(x^m) \sum_{v^n} \Big| \sum_{i} \Pr_{\hat{V}^n,A|X^m}(v^n, i|x^m) - \Pr_{\hat{V}^n,A|\hat{U}^n}(v^n, i|u^n) \Big| \\
&\leq \sum_{x^m} \Pr_{X^m}(x^m) \sum_{v^n} \sum_{i} Pr_A(i) \Big| \Pr_{\hat{V}^n|A,X^m}(v^n|i, x^m) - \Pr_{\hat{V}^n|A,\hat{U}^n}(v^n|i, u^n) \Big| \\
&\leq \epsilon + \sum_{x^m} \Pr_{X^m}(x^m) \sum_{v^n} \Big| \Pr_{\hat{V}^n|A,X^m}(v^n|1, x^m) - \Pr_{\hat{V}^n|A,\hat{U}^n}(v^n|1, u^n) \Big| \\
&= \epsilon + \sum_{i} \Pr_{A}(i) \sum_{x^m} \Pr_{X^m|A}(x^m|i) \sum_{v^n} \Big| \Pr_{\hat{V}^n|A,X^m}(v^n|1, x^m) - \Pr_{\hat{V}^n|A,\hat{U}^n}(v^n|1, u^n) \Big| \\
&\leq 2\epsilon + \sum_{x^m} \Pr_{X^m|A}(x^m|1) \sum_{v^n} \Big| \Pr_{\hat{V}^n|A,X^m}(v^n|1, x^m) - \Pr_{\hat{V}^n|A,\hat{U}^n}(v^n|1, u^n) \Big| \\
&= 2\epsilon + \sum_{x^m} \Pr_{\tilde{X}^m|A}(x^m|1) \sum_{v^n} \Big| \Pr_{V^n|A,\tilde{X}^m}(v^n|1, x^m) - \Pr_{V^n|A,\tilde{U}^n}(v^n|1, u^n) \Big|.
\end{aligned}$$

The last line follows from the fact that we are conditioning on the event of $X^m\hat{U}^n\hat{V}^n = \tilde{X}^m\tilde{U}^n V^n$. Note also, that for $\epsilon < 1/2$, and using the fact that the probability of the event $A = 1$ to occur is at least $1 - \epsilon$, we can upper bound the last line by

$$2\epsilon + 2\sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \Big| \Pr_{V^n|\tilde{X}^m}(v^n|x^m) - \Pr_{V^n|\tilde{U}^n}(v^n|u^n) \Big|.$$

Before wrapping up the proof, let us prove the following claim.

Claim. *$U^n = \tilde{U}^n$ with a probability of at least $1 - \epsilon_D$, where $\epsilon_D = n \cdot k \cdot \epsilon$ and k is a constant that depends on the description of the correlation (U, V).*

Proof. We will show that, for each $1 \leq i \leq n$, $U_i \neq \tilde{U}_i$ with probability at most $k\epsilon$. The claim then follows by a union bound.

We will prove that $U_i \neq \tilde{U}_i$ with probability at most $k\epsilon$ when $i = 1$. The argument is similar for any $i \neq 1$. Let T be $|\mathcal{U}| \times |\mathcal{U}|$ dimensional matrix representing the conditional distribution of $\tilde{U}_1$ conditioned on U_1. We will refer to any rows and columns of T by the corresponding element in $\mathcal{U}$. That is, the value in row $u \in \mathcal{U}$ and column $\tilde{u} \in \mathcal{U}$ is denoted by $T_{u,\tilde{u}}$. Clearly, T is a stochastic matrix since each entry is non-negative and each row adds to 1. We have, $T_{u,\tilde{u}} = \Pr_{\tilde{U}_1|U_1}(\tilde{u}|u)$ for all $u, \tilde{u} \in \mathcal{U}$.

$$\begin{aligned}
\Pr[U_1 \neq \tilde{U}_1] &= \sum_{u \in \mathcal{U}} \sum_{\tilde{u} \in \mathcal{U} \setminus \{u\}} \Pr\left[U_1 = u, \tilde{U}_1 = \tilde{u}\right] \\
&= \sum_{u \in \mathcal{U}} \Pr_U(u) \sum_{\tilde{u} \in \mathcal{U} \setminus \{u\}} \Pr\left[\tilde{U}_1 = u | \tilde{U}_1 = \tilde{u}\right] \\
&\leq \sum_{u \in \mathcal{U}} \sum_{\tilde{u} \in \mathcal{U} \setminus \{u\}} T_{u,\tilde{u}}. \qquad (14)
\end{aligned}$$

By the security definition, $(V^n, U^n) \approx_{2\epsilon} (V^n, S(\mathcal{S}_S(U^n)) = (V^n, \tilde{U}^n)$. Consequently,

$$(V_1, U_1) \approx_{2\epsilon} (V_1, \tilde{U}_1). \qquad (15)$$

Let the $|\mathcal{V}| \times |\mathcal{U}|$ dimensional matrices D and $\tilde{D}$ denote the joint distributions (V_1, U_1) and $(V_1, \tilde{U}_1)$, respectively. That is, $D_{v,u} = \Pr_{V_1,U_1}(v, u)$ and $D_{v,\tilde{u}} = \Pr_{V_1,\tilde{U}_1}(v, \tilde{u})$. By the definition of matrices T, D and $\tilde{D}$, it holds that $DT = \tilde{D}$. Furthermore, let $E = |D - \tilde{D}|$, where LHS is the absolute value of the difference between D and $\tilde{D}$. By Eq. (15), the sum of all entries of E is at most 4ϵ. It is shown in [1, Lemma 6] that this implies that the stochastic matrix T is $k\epsilon$ close to I, the identity matrix, for a constant k that depends only on the description of the distribution (U_1, V_1). In other words, the sum of all entries of $|I - T|$ is at most $k\epsilon$. Along with Eq. (14) and the fact that I is diagonal, this implies that $\Pr[U_1 \neq \tilde{U}_1] \leq k\epsilon$.

We note a caveat that the result from [1] holds only when correlation (U_1, V_1) is *non-redundant*, wherein non-redundancy requires that there exist no u, u' such that the conditional distribution of V_1 conditioned on $U_1 = u$ is identical to that conditioned on $U_1 = u'$. However, [1] also showed that there is a perfectly secure non-interactive reduction (perfect NISR) from any correlation to its so-called core which is obtained by collapsing the redundant symbols together and back. Thus, using any source correlation, OMSR of n copies of a target correlation is equivalent to OMSR of n copies of the core of the target correlation. This concludes the proof. □

Now, let us denote by B an indicator of the event $U^n = \tilde{U}^n$. Since $\tilde{U}^n = S(\tilde{X}^m)$, $\Pr_{V^n|\tilde{X}^m}(v^n|x^m) = \Pr_{V^n|\tilde{X}^m\tilde{U}^n}(v^n|x^m, u^n)$. Thus, we have,

$$\begin{aligned}
&\sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \left| \Pr_{V^n|\tilde{X}^m}(v^n|x^m) - \Pr_{V^n|\tilde{U}^n}(v^n|u^n) \right| \\
&= \sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \left| \Pr_{V^n|\tilde{X}^m\tilde{U}^n}(v^n|x^m, u^n) - \Pr_{V^n|\tilde{U}^n}(v^n|u^n) \right| \\
&= \sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \sum_{i} \Pr_B(i) \left| \Pr_{V^n|B\tilde{X}^m\tilde{U}^n}(v^n|i, x^m, u^n) - \Pr_{V^n|B\tilde{U}^n}(v^n|i, u^n) \right| \\
&\leq \epsilon_D + \sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \left| \Pr_{V^n|B\tilde{X}^m\tilde{U}^n}(v^n|1, x^m, u^n) - \Pr_{V^n|B\tilde{U}^n}(v^n|1, u^n) \right| \\
&= \epsilon_D + \sum_{x^m} \Pr_{\tilde{X}^m}(x^m) \sum_{v^n} \left| \Pr_{V^n|B\tilde{X}^mU^n}(v^n|1, x^m, u^n) - \Pr_{V^n|BU^n}(v^n|1, u^n) \right| \\
&= \epsilon_D + 0.
\end{aligned}$$

Overall, we have $\mathbb{E}_{X^m}\mathsf{TVD}[\hat{V}^n|X^m\hat{U}^n, \hat{V}^n|\hat{U}^n] \leq 2(\epsilon + \epsilon_D)$.

Definition A.2 (OM(S)R for distribution families)**.** Let $p_{(U_n,V_n)}, n \in \mathbb{N}$ be a family of correlations over the domains $\mathcal{U}_n \times \mathcal{V}_n$ for each $n \in \mathbb{N}$. For $\epsilon : \mathbb{N} \to \mathbb{R}_{\geq 0}$, an ϵ-error *one-message (secure) reduction* (ϵ-OM(S)R) for converting the source correlation p_{XY} over the domain $\mathcal{X} \times \mathcal{Y}$ to correlation family $p_{(U_n,V_n)}$, $n \in \mathbb{N}$ is a sequence of of randomized algorithms $\langle S_n, R_n \rangle$, for each $n \in \mathbb{N}$, where $S_n : \mathcal{X}^{m(n)} \to \mathcal{U}_n \times \{0,1\}^{\ell(n)}$ and $R : \{0,1\}^{\ell(n)} \times \mathcal{Y}^{m(n)} \to \mathcal{V}_n$ such that $\langle S_n, R_n \rangle$ is an ϵ-OM(S)R converting $m(n)$ copies of p_{XY} into a copy of $p_{U_nV_n}$.

Correctness. For i.i.d. $(X_i, Y_i) \sim p_{XY}$, $i \in [m]$, and $(U_n, V_n) \sim p_{(U_n,V_n)}$, when $(\hat{U}_n, M) \leftarrow S(X^m)$ and $\hat{V}_n \leftarrow R(M, Y^m)$,

$$\left(\hat{U}_n, \hat{V}_n\right) \approx^{\epsilon} (U_n, V_n) \tag{16}$$

An ϵ-OMR $\langle S, R \rangle$ is an ϵ-error *one-way secure message reduction* (ϵ-OMSR) if the following security conditions are met:

Privacy against S**.** There exists a randomized simulator $\mathcal{S}_S : \mathcal{U}_n \to \mathcal{X}^m$ such that,

$$\left(X^m, \hat{V}_n\right) \approx^{\epsilon} (\mathcal{S}_S(U_n), V_n)\,. \tag{17}$$

Privacy against R**.** There exists a randomized simulator $\mathcal{S}_R : \mathcal{V}_n \to \mathcal{Y}_n \times \mathcal{L}$ such that,

$$\left(\hat{U}_n, M, Y^m\right) \approx^{\epsilon} (U_n, \mathcal{S}_R(V_n))\,. \tag{18}$$

We say that an OMR (resp. OMSR) is *perfect* if $\epsilon(n) = 0$ for all n. The communication cost for the OM(S)R is computed as $\limsup \ell(n)/n$. If $\epsilon(n)$ is a negligible function, we say we can the sequence of reductions a statistical one-message (secure) reduction with negligible error.

B Details omitted from Section 5

B.1 Proof of Lemma 5.3

Proof. We will use the $\mathbb{F}_2^2$ representation for $\mathbb{F}_4$, i.e., we represent $\mathbb{F}_4$ as $\{0 = 00, 1 = 01, \alpha = 10, \beta = 11\}$. We will show how to locally convert an nzOLE to a $(3,2)$-correlation; since this conversion is a bijection, the conversion in the other direction will hold as well. This will allow us to generate nzOLE instead of $(3,2)$-correlations in our protocols.

Consider an nzOLE correlation $((a, s), (b, r))$ where $ab = s + r$ (and $a, b \neq 0$). Denote the bit decomposition of r and s by $r = r_1 \parallel r_0$ and $s = s_1 \parallel s_0$. Define x_0 such that $\alpha^{x_0} = a$ (in the multiplicative subgroup $\mathbb{F}_4^*$). Similarly define x_1 such that $\alpha^{x_1} = b$. Note that both α and β are generators of $\mathbb{F}_4^* \cong \mathbb{Z}_3$ and that $x_0, x_1 \in \mathbb{Z}_3$. Let $x = x_0 + x_1$ (in $\mathbb{Z}_3$).

Now, given (a, s), the first party will output $w = (x_0, u_0, v_0) = (x_0, s_0 + 1, s_1 + 1)$; similarly, given (b, r), the second party will output $y = (x_1, u_1, v_1) = (x_1, r_0, r_1)$. It is easy to see that (w, y) forms a $(3,2)$-correlation, as illustrated below:

- When $x = 0$, we have $ab = \alpha^0 = 01 = s + r$ which makes $s_0 + r_0 = 1$ and $s_1 + r_1 = 0$. Therefore, $(u, v) = (0, 1)$ since $u = u_0 + u_1 = (s_0 + 1) + r_0 = 0$, and $v = v_0 + v_1 = (s_1 + 1) + r_1 = 1$.

- When $x = 1$, we have $ab = \alpha^1 = \alpha = (10) = s + r$ which makes $s_0 + r_0 = 0$ and $s_1 + r_1 = 1$. Therefore, $(u, v) = (1, 0)$ since $u = (s_0 + 1) + r_0 = 1$ and $v = (s_1 + 1) + r_1 = 0$.
- When $x = 2$, we have $ab = \alpha^2 = \beta = (11) = s + r$ which makes $s_0 + r_0 = 1$ and $s_1 + r_1 = 1$. Therefore, $(u, v) = (0, 0)$ since $u = (s_0 + 1) + r_0 = 0$ and $v = (s_1 + 1) + r_1 = 0$.

Observe that $u = u_0 + u_1 = x \bmod 2$, and $v = v_0 + v_1 = (x + 1 \bmod 3) \bmod 2$ and that both u and v are shared randomly. Therefore, we can conclude that the parties indeed output a $(3, 2)$-correlation. Moreover, since this transformation was a bijection, we can also reverse the process to locally convert from a $(3, 2)$-correlation to a nzOLE correlation. □

B.2 Proof of Theorem 5.4

Proof. It suffices to show a protocol for non-zero OLE over $\mathbb{F}_4$ since using Lemma 5.3, the parties can locally convert it into a $(3, 2)$-correlation.

We will first note that it is possible for the following protocol to be generalized to non-zero OLE over any field, but its efficiency scales poorly with the field size. The protocol will be quite similar to the one in Theorem 5.2.

Recall that in a 1-out-of-3 OT correlation over $\mathbb{F}_4$, the sender S is given a tuple $\mathbf{r} = (r_0, r_1, r_2)$ that is uniform over $\mathbb{F}_4^3$. For ease of presentation the receiver R is given (b, r_b) where b is a uniformly random over $\mathbb{F}_4\backslash\{0\}$ and r_b is its corresponding element in $\mathbf{r}$.

Now, define a function $f_{a,s}(b) : \mathbb{F}_4\backslash\{0\} \to \mathbb{F}_4$, parameterized by a, s as $f_{a,s}(b) = r$ where $r = ab + s$. In other words, the output r is such that $((a, s), (b, r))$ is a valid nzOLE. Note that the function $f_{(a,s)}$ is distinct for distinct (a, s).

Now, for each (a, s), define the vector $\mathbf{r}_{(a,s)} = (f_{(a,s)}(1), f_{(a,s)}(\alpha), f_{(a,s)}(\beta))$, i.e., defined by evaluating $f_{(a,s)}$ at each $b \in \mathbb{F}_4\backslash\{0\}$. Denote by Φ, the set of all possible vectors $\mathbf{r} \in \mathbb{F}_4^3$ for which there exist some (a, s) such that $\mathbf{r}_{(a,s)} = \mathbf{r}$. Intuitively, Φ denotes the accept set for the conversion—when the OT source correlation $\mathbf{r}$ given to S is in Φ, then it will send an "accept" message after which both parties will compute the target nzOLE correlation by local computation; when $\mathbf{r} \notin \Phi$, then both parties will abort.

Now, given $\mathbf{r} \in \Phi$, to generate required the nzOLE correlation, S will find the (a, s) such that $\mathbf{r} = \mathbf{r}_{(a,s)}$ and output it; the receiver R will simply output (b, r_b).

Notice that this results in a valid nzOLE correlation since $f_{(a,s)}(b) = r_b$, and therefore $r = ab+s$.

Over a random $\mathbf{r}$, the probability that $\mathbf{r} \in \Phi$ can be given by $\frac{12}{64} = \frac{3}{16}$; this is exactly the accept probability of the protocol.

It is easy to see that both parties don't learn any additional information about the other's output; The sender doesn't know which value of b the receiver has. The receiver's view consists of only b, r_b and he doesn't know which a, s were consistent with the sender's part of the correlation. □

C Details omitted from Section 6

C.1 Proof of Lemma 6.2

If the amortized communication cost of c is achievable for OMR converting p_{XY} to common randomness, for an increasing sequence $n_1, n_2, n_3, \ldots$, there exist functions $S_i : \mathcal{X}^{m_i} \to [2^{n_i}] \times [2^{c \cdot n_i}]$

and $R_i : \mathcal{Y}^{m_i} \times [2^{c \cdot n_i}] \to [2^{n_i}]$ for each n_i such that, when $(X^{m_i}, Y^{m_i}) \sim p_{XY}^{\otimes m_i}$, $S_i(X^{m_i}) = (K_i, M_i)$ and $R_i(Y^{m_i}, M_i) = L_i$,

$$\Pr\left[K_i \neq L_i\right] \leq \epsilon_i, \tag{19}$$

$$\sum_{k \in [2^{n_i}]} \left| \Pr\left[K_i = k\right] - \frac{1}{2^{n_i}} \right| \leq \epsilon_i, \tag{20}$$

where, $\epsilon_i \to 0$ as $i \to \infty$. Fix some n_i in the sequence; we will drop the subscript to avoid redundancy and denote m_i by m, S_i by S, K_i by K and so on.

$$\begin{aligned} H(K|Y^m) &= I(K; M|Y^m) + H(K|Y^m, M) \\ &\leq H(M) + H(K|L) \\ &\leq c \cdot n + h(\epsilon) + n\epsilon. \end{aligned} \tag{21}$$

Since M is distributed over $2^{c \cdot n}$, $H(g(X^n)) \leq c \cdot n$. The next bound follows from a fundamental result in information theory–Fano's inequality. K takes values in $[2^n]$, by Eq. (19) and Fano's inequality [20, Theorem 2.10.1], $H(K|L) \leq h(\epsilon) + n\epsilon$.

The last inequality follows from these observations.

Next, since $H(K|Y^m) = H(K) - I(Y^m; K) = I(X^m; K) - I(Y^m; K) = H(K|Y^m) - H(K|X^m)$, expanding RHS using a telescopic summation,

$$\begin{aligned} H(K|Y^m) &= H(K|Y^m) - H(K|X^m) \\ &= \sum_{i=1}^{m} H(K|X^{i-1}Y_i^m) - H(K|X^i, Y_{i+1}^m) \\ &= \sum_{i=1}^{m} H(K|X^{i-1}, Y_{i+1}^m) - H(K|X^i, Y_{i+1}^m) \\ &\qquad\qquad - H(K|X^{i-1}, Y_{i+1}^m) + H(K|X^{i-1}Y_i^m) \\ &= \sum_{i=1}^{m} I(K; X_i|X^{i-1}, Y_{i+1}^m) - I(K; Y_i|X^{i-1}, Y_{i+1}^m). \end{aligned} \tag{22}$$

Let J be uniformly distributed over the set $[m]$ independent of all the previously defined random variables.

$$\begin{aligned} &\sum_{i=1}^{m} I(K; X_i|X^{i-1}, Y_{i+1}^m) - I(K; Y_i|X^{i-1}, Y_{i+1}^m) \\ &\qquad = mI(K; X_J|X^{J-1}, Y_{J+1}^m, J) - mI(K; Y_J|X^{J-1}, Y_{J+1}^m, J) \\ &\qquad = mI(K, X^{J-1}, Y_{J+1}^m, J; X_J) - mI(K, X^{J-1}, Y_{J+1}^m, J; Y_J) \end{aligned}$$

The last equality used the independence between Y_J (and X_J) and X^{J-1}, Y_{J+1}^1, J. Hence, defining $U = KX_1, \ldots, X_{J-1}, Y_{J+1}, \ldots, Y_m J$, by Eq. (21) and Eq. (22),

$$c \cdot n + h(\epsilon) + n\epsilon \geq mI(U; X_J) - mI(U; Y_J). \tag{23}$$

Furthermore, by *Eq.* (20), $f(X^m)$ is ϵ-close to the uniform distribution over $[2^n]$. Hence, by Fano's inequality, $H(K) = H(f(X^m)) \geq n - h(\epsilon) - n\epsilon$. Thus,

$$\begin{aligned} n - h(\epsilon) - n\epsilon \leq H(K) &= I(K; X^m) \\ &= \sum_{i=1}^{m} I(K; X_i | X_1, \ldots, X_{i-1}) \\ &= mI(K; X_J | X_1, \ldots, X_{J-1}) \\ &\leq mI(U; X_J). \end{aligned}$$

Finally, since K is a function of X^m, the following Markov chain holds:

$$U \leftrightarrow X_J \leftrightarrow Y_J.$$

We may identify (X_J, Y_J) with generic random variables $(X, Y) \sim p_{XY}$ and conclude from the above observations that there exists a random variable U satisfying the Markov chain $U \leftrightarrow X \leftrightarrow Y$ such that $n - h(\epsilon) - n\epsilon \leq mI(U; X)$ and $c \cdot n + h(\epsilon) + n\epsilon \geq mI(U; X_J) - mI(U; Y_J)$. Hence, we conclude that, for each n_i,

$$\frac{c + \epsilon_i + h(\epsilon_i)/n_i}{1 - \epsilon_i - h(\epsilon_i)/n_i} \geq \frac{I(U_i; X) - I(U_i; Y)}{I(U_i; X)} \geq \inf_{p(u|x)} \frac{I(U; X) - I(U; Y)}{I(U; X)}.$$

Since $\epsilon_i \to 0$ as $i \to \infty$, by definition of $\mathrm{S}^*(X, Y)$,

$$c = \lim_{i \to \infty} \frac{R + \epsilon_i + h(\epsilon_i)/n_i}{1 - \epsilon_i - h(\epsilon_i)/n_i} \geq \inf_{p(u|x)} \frac{I(U; X) - I(U; Y)}{I(U; X)} = 1 - \mathrm{S}^*(X, Y).$$

This concludes the proof.

C.2 Proof of Lemma 6.5

Let (X^k, Y) be a 1-out-of-k correlation OT over $\mathbb{Z}_n$. That is, X_i is i.i.d. according to uniform distribution over $\mathbb{Z}_n$ for all $i \in [k]$ and $Y = (B, X_B)$ where B is uniform over [k] independent of X^k.

$\mathrm{S}^*(X^k, Y) \geq \frac{1}{k}$ - Let $U^* = X^k$. Then, U^* satisfies the Markov chain $U^* \leftrightarrow X^k \leftrightarrow Y$. We have,

$$\mathrm{S}^*(X^k, Y) = \sup_{p(u|x^k)} \frac{I(U; Y)}{I(U; X^k)} \geq \frac{I(U^*; Y)}{I(U^*; X^k)} = \frac{I(X^k; X_B, B)}{I(X^k; X^k)} = \frac{1}{k}.$$

$\mathrm{S}^*(X^k, Y) \leq \frac{1}{k}$ - Let U be any random variable satisfying the Markov chain $U \leftrightarrow X^k \leftrightarrow Y$. U can be equivalently described as $U = f(X^k, R)$ where R is independent of both X^k and Y.

$$\frac{I(U; Y)}{I(U; X^k)} = \frac{I(f(X^k, R); Y)}{I(f(X^k, R); X^k)} = \frac{\sum_r \Pr[R = r] I(f(X^k, r); Y)}{\sum_r \Pr[R = r] I(f(X^k, r); X^k)} \leq \max_r \frac{I(f(X^k, r); Y)}{I(f(X^k, r); X^k)}$$

where the inequality follows from the fact that, for any finite k, $\frac{\sum_{i=1}^k a_i}{\sum_{i=1}^k b_i} \leq \frac{a_k}{b_k}$ if $\frac{a_i}{b_i} \leq \frac{a_{i+1}}{b_{i+1}}$ for all i, which in turn can be proved using induction.

Claim. *For any finite k, $\frac{\sum_{i=1}^{k} a_i}{\sum_{i=1}^{k} b_i} \leq \frac{a_k}{b_k}$ if $\frac{a_i}{b_i} \leq \frac{a_{i+1}}{b_{i+1}}$ for all i*

Proof. We prove this using induction: The claim clearly holds for $k = 1$. Suppose the claim holds for $k - 1$, then

$$\frac{\sum_{i=1}^{k-1} a_i}{\sum_{i=1}^{k-1} b_i} \leq \frac{a_{k-1}}{b_{k-1}} \leq \frac{a_k}{b_k}$$

But, since $\frac{x+x'}{y+y'} \leq \frac{x'}{y'}$ whenever $\frac{x}{y} \leq \frac{x'}{y'}$, we get

$$\frac{\sum_{i=1}^{k} a_i}{\sum_{i=1}^{k} b_i} \leq \frac{a_k}{b_k}.$$

□

Note that, we can get rid of r such that $I(f(X^k, r); X^k) = 0$ (which implies $H(f(X^k, r) = 0)$ from both numerator and denominator before applying the mentioned induction.

The above discussion essentially shows that we need to consider only U such that $U = f(X^k)$ for some deterministic function f. Hence, proceeding with $U = f(X^k)$ for some function f,

$$\frac{I(U;Y)}{I(U;X^k)} = \frac{H(f(X^k)) - H(f(X^k)|X_B, B)}{H(f(X^k))}.$$

We first establish a claim to help us bound this term.

Claim. $\sum_{i=1}^{k} H(f(X^k)|X_i) \geq (k-1)H(f(X^k))$.

Proof. Firstly, for every $i \in [k]$, denoting $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k$ by X^{-i},

$$\begin{aligned} H(f(X^k)) = I(f(X^k); X^k) &= I(f(X^k); X_i) + I(f(X^k); X^{-i}|X_i) \\ &= I(f(X^k); X_i) + H(f(X^k)|X_i). \end{aligned}$$

The last equality is due to the fact that $H(f(X^k)|X^k) = 0$. Next,

$$\begin{aligned} \sum_{i=1}^{k} I(f(X^k); X_i) &\leq \sum_{i=1}^{k} I(f(X^k), X_1^{i-1}; X_i) \\ &= \sum_{i=1}^{k} I(f(X^k); X_i|X_1^{i-1}) \\ &= I(f(X^k); X^k) = H(f(X^k)). \end{aligned}$$

The first equality in the second line is due to the chain rule for mutual information and $\{X_i\}_{i=[k]}$ being independent of each other. Now we can conclude the proof by

$$\begin{aligned} kH(f(X^k)) = \sum_{i=1}^{k} H(f(X^k)) &= \sum_{i=1}^{k} H(f(X^k)|X_i)) + \sum_{i=1}^{k} I(f(X^k); X_i) \\ &\leq \sum_{i=1}^{k} H(f(X^k)|X_i) + H(f(X^k)) \end{aligned}$$

This proves the claim. □

Using the above claim,

$$\frac{H(f(X^k)) - H(f(X^k)|X_B, B)}{H(f(X^k))} = 1 - \frac{\sum_{i=1}^k H(f(X^k)|X_i)}{kH(X^k)} \leq 1 - \frac{k-1}{k} = \frac{1}{k}.$$

Thus, we conclude that S* of 1-out-of-k OT is $\frac{1}{k}$.

C.3 Proof of Lemma 6.6

Let (X, Y) be an OLE correlation over a field $\mathbb{F}$. That is $X = (A, S), Y = (B, R)$, where A, B, S are i.i.d. according to uniform distribution over $\mathbb{F}$ and $R = S - AB$. Note that this is isomorphic to the definition we gave in the preliminaries.

$\mathrm{S}^*(X, Y) \geq \frac{1}{2}$ - Let $U^* = X$. Then, U^* satisfies the Markov chain $U^* \leftrightarrow X \leftrightarrow Y$. We have,

$$\mathrm{S}^*(X, Y) = \sup_{p(u|x)} \frac{I(U;Y)}{I(U;X)} \geq \frac{I(U^*;Y)}{I(U^*;X)} = 1 - \frac{H(X|Y)}{H(X)} = 1 - \frac{\log|\mathbb{F}|}{2\log|\mathbb{F}|} = \frac{1}{2}. \tag{24}$$

$\mathrm{S}^*(X, Y) \leq \frac{1}{2}$ - Let U be any random variable satisfying the Markov chain $U \leftrightarrow X \leftrightarrow Y$. Similarly to the discussion in Appendix C.2, it is sufficient to consider U in the form of $U = f(X)$ for some deterministic function f. Thus, let us continue with the analysis:

$$\frac{I(U;Y)}{I(U;X)} = \frac{H(f(A, S)) - H(f(A, S)|B, R)}{H(f(A, S))}.$$

Claim. $2H(f(A, S)|B, R) \geq H(f(A, S))$

Proof. We have

$$\begin{aligned} 2H(f(A, S)|B, R) &= 2H(f(A, S)|S - AB, B) \\ &= 2\sum_b \Pr[B = b] H(f(A, S)|S - bA) \\ &= \frac{1}{|\mathbb{F}|}\sum_b H(f(A, S)|S - bA) + H(f(A, S)|S - (1 + b)A) \end{aligned}$$

Let $U = S - bA$ and $V = S - (1 + b)A$. We will show that, for each $b \in \mathbb{F}$, W and V are independent, and $H(f(A, S)|W, V) = 0$. Then, the claim follows from the same argument as in the proof of Appendix C.2.

Firstly, since S and A are independent and uniform over $\mathbb{F}$, for any b, $U - V = A$ is independent of $U = S - bA$. Hence, U and V are independent. Secondly, for each b, $A = U - V$ and $S = U + bA$. Hence, (A, S) is fully determined by (U, V). Hence, $H(f(A, S)|U, V) = 0$ for all b. This concludes the proof. □

C.4 Proof of Lemma 6.7

Let (X, Y) be an additive correlation over a group $\mathbb{G}$, and let D be the distribution promised from the definition. We will prove that if $H(D) < \frac{log|\mathbb{G}|}{2}$ then $\mathrm{S}^*(X, Y) > \frac{1}{2}$.

$$\mathrm{S}^*(X,Y) = \sup_{p(u|x)} \frac{I(U;Y)}{I(U;X)} \geq \frac{I(X;Y)}{I(X;X)} = 1 - \frac{H(X|Y)}{H(X)} = 1 - \frac{H(D)}{\log|\mathbb{G}|} > \frac{1}{2}$$

C.5 Proof of Lemma 6.9

Proof of Lemma 6.9. The construction of f, g from π in our proof is similar to the construction of zero communication protocol in the proof of [18, Theorem 2.6] with a notable difference: In the latter, the output of the interactive protocol is a pair of random vectors that agree with high probability, whereas in our case, the n-bit vector correlation output by π disagree on at most one coordinate with high probability. Hence, to get agreement, we flip the value at a random coordinate of one of the output vectors; this achieves agreement with protbability $\frac{1}{n}$.

Let $\pi_S(X, Y)$ and $\pi_R(X, Y)$ be the outputs of S and R, respectively. Functions f and g are defined as follows:

$f(X)$**:** Sample $\hat{Y}$ according to $p_{Y|X}$ conditioned on X and output $\hat{U} = \pi_S(X, \hat{Y})$.

$g(Y)$**:** Sample $\hat{X}$ according to $p_{X|Y}$ conditioned on Y and compute $\hat{V} = \pi_R(X, \hat{Y})$. Flip a uniformly random coordinate i of $\hat{V}$ to obtain $\tilde{V}$ and output $\tilde{V}$.

Since $(X, \hat{Y})$ and (X, Y) are identically distributed, $\hat{U} = (X, \hat{Y})$ is uniformly distributed over $\{0,1\}^n$. Similarly, $\hat{V}$ is uniformly distributed over $\{0,1\}^n$. $\tilde{V}$ is also uniformly distributed over $\{0,1\}^n$ since it is obtained by flipping a random coordinate of $\hat{V}$. Hence, to prove the lemma, it suffices to upper bound $\Pr\left[\hat{U} = \tilde{V}\right]$.

Let $\theta = (\theta_S, \theta_R)$, where θ_S and θ_R are any realization of private randomness of S and R in π. Let $|a - b|$ denote the Hamming distance between two strings a and b. Define the following events:

- E_θ is the event "$|\pi_S(X, Y) - \pi_R(X, Y)| = 1$ conditioned on private randomness of S and R being θ_S and θ_R, respectively".
- E'_θ is the event "$|\hat{U} - \hat{V}| = 1$ conditioned on private randomness used by f and g in $\pi(X, \hat{Y})$ and $\pi(\hat{X}, Y)$ are θ_S and θ_R, respectively".

Claim. *For any* $\theta = (\theta_S, \theta_R)$, $\Pr[E'_\theta] \geq \frac{(\Pr[E_\theta])^3}{32} 4^{-\ell}$.

We first prove the lemma assuming the claim and prove the claim afterwards. Since the coordinate i that was flipped to obtain $\tilde{V}$ from $\hat{V}$ was chosen uniformly and independently, by definition of

E'_θ, $\Pr\left[\hat{U} = \tilde{V} \mid E'_\theta\right] = \frac{1}{n}$. Hence,

$$\begin{aligned}
\Pr\left[\hat{U} = \tilde{V}\right] &= \sum_\theta \Pr\left[\theta\right] \Pr\left[E'_\theta\right] \Pr\left[\hat{U} = \tilde{V} \mid E'_\theta\right] \\
&= \frac{1}{n} \sum_\theta \Pr\left[\theta\right] \Pr\left[E'_\theta\right] \\
&\geq \frac{1}{n} \sum_\theta \Pr\left[\theta\right] \frac{(\Pr\left[E_\theta\right])^3}{32} 4^{-\ell} \\
&\geq \frac{4^{-\ell}}{32n} \left(\sum_\theta \Pr\left[\theta\right] \Pr\left[E_\theta\right]\right)^3 .
\end{aligned}$$

Finally,

$$\sum_\theta \Pr\left[\theta\right] \Pr\left[E_\theta\right] = \Pr\left[|\pi_S(X,Y) - \pi_R(X,Y)| = 1\right] \geq 1 - \epsilon.$$

The last inequality used correctness of π by which, when (U,V) is an n-bit unit vector correlation,

$$\begin{aligned}
&(\pi_S(X,Y), \pi_R(X,Y)) \approx^\epsilon (U,V) \\
&\qquad\qquad\qquad \implies \Pr\left[|\pi_S(X,Y) - \pi_R(X,Y)| = 1\right] \geq \Pr\left[|U - V| = 1\right] - \epsilon,
\end{aligned}$$

where the implication follows from data processing inequality for total variation distance. We conclude the proof of the lemma by proving the above claim

of claim. When private randomness is fixed to θ, for any realization $(x,y) = (X,Y)$, the transcript of π is a deterministic function of (x,y), which we denote by $t(x,y)$. For a transcript $t \in \{0,1\}^\ell$, let $Q_x(t)$ be the probability with which the transcript of π is t conditioned on $X = x$ and the private randomness of S is θ_S; similarly, let $Q_y(t)$ be the probability with which the transcript of π is t conditioned on $Y = y$ and the private randomness of R is θ_S. Let $G = \{(x,y) : |\pi_S(x,y) - \pi_R(x,y)| = 1\}$; by definition of E_θ, $\Pr\left[(X,Y) \in G\right] = E_\theta$.

Define

$$B = \left\{(x,y) : \left(Q_x(t(x,y)) < \Pr\left[E_\theta\right] 2^{-\ell}/4\right) \vee \left(Q_y(t(x,y)) < \Pr\left[E_\theta\right] 2^{-\ell}/4\right)\right\}.$$

We will show that B occurs with probability at most $\Pr\left[E_\theta\right]/2$.

$$\begin{aligned}
&\Pr\left[(X,Y) : \left(Q_X(t(X,Y)) < E_\theta 2^{-\ell}/4\right)\right] \\
&= \sum_x \sum_{t : Q_x(t) < \frac{\Pr\left[E_\theta\right]}{4} 2^{-\ell}} \sum_{y : t(x,y) = t} p_{XY}(x,y) \\
&= \sum_x p_X(x) \sum_{t : Q_x(t) < \frac{\Pr\left[E_\theta\right]}{4} 2^{-\ell}} Q_x(t) \\
&< \sum_x p_X(x) \sum_{t : Q_x(t) < \frac{\Pr\left[E_\theta\right]}{4} 2^{-\ell}} \Pr\left[E_\theta\right] 2^{-\ell}/4 \\
&\leq \Pr\left[E_\theta\right] 2^{-\ell}/4.
\end{aligned}$$

Similarly, $\left\{(x,y) : Q_y(t(x,y)) < \Pr\left[E_\theta\right] 2^{-\ell}/4\right\}$ also occurs with probability less that $\Pr\left[E_\theta\right] 2^{-\ell}/4$. Hence, B occurs with probability at most $\Pr\left[E_\theta\right]/2$. Now, the probability of E'_θ can be upper bounded as

$$\begin{aligned}
\Pr\left[E'_\theta\right] &\geq \sum_{(x,y)\in G} p_{XY}(x,y) Q_x(t(x,y)) Q_y(t(x,y)) \\
&\geq \sum_{(x,y)\in G\setminus B} p_{XY}(x,y) Q_x(t(x,y)) Q_y(t(x,y)) \\
&\geq \sum_{(x,y)\in G\setminus B} p_{XY}(x,y) \left(\frac{\Pr\left[E_\theta\right] 2^{-\ell}}{4}\right)^2 \\
&\geq \frac{(\Pr\left[E_\theta\right])^2 4^{-\ell}}{16} \left(\Pr\left[(X,Y)\in G\right] - \Pr\left[(X,Y)\in B\right]\right) \\
&\geq \frac{(\Pr\left[E_\theta\right])^3 4^{-\ell}}{32}.
\end{aligned}$$

□

□

D Proof of Theorem 7.1 omitted from Section 7

Let $(\widehat{U}^n, M) = S(X,Q)$, $\widehat{V}^n = R(M,Y,Q)$, and let (U_i, V_i) be i.i.d. according to the correlation p_{UV} for all $i \in [n]$. By ϵ-correctness of $\langle S, R\rangle$,

$$\left(\widehat{U}^n, \widehat{V}^n\right) \approx^\epsilon \left(U^n, V^n\right).$$

We will denote $u^n \in \mathcal{U}^n$ and $v^n \in \mathcal{V}^n$ by $\boldsymbol{u}$ and $\boldsymbol{v}$. By ϵ-security against R, denoting the private randomness by Q'

$$\sum_{x^m, q', q} \Pr\left[X^m = x^m, Q' = q', Q = q\right] \sum_{u^n} \Pr\left[\hat{U}^n = u^n | X^m = x^m, Q' = q', Q = q\right]$$
$$\cdot \mathsf{TVD}\left(\left(\hat{V}^n \middle| X^m = x^m, Q' = q', Q = q\right), \left(\hat{V}^n \middle| \hat{U}^n = u^n\right)\right) \leq \epsilon.$$

Marginalizing over Q' and X^m, and using the definition of TVD,

$$\sum_{q\in Q} \Pr\left[Q = q\right] \sum_{\boldsymbol{u}\in\mathcal{U}^n} \Pr\left[\hat{U}^n = \boldsymbol{u} | Q = q\right]$$
$$\sum_{\boldsymbol{v}\in\mathcal{V}^n} \left| \Pr\left[\hat{V}^n = \boldsymbol{v} | \hat{U}^n = \boldsymbol{u}, Q = q\right] - \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \right| \leq 2\epsilon.$$

Similarly,

$$\sum_{q\in Q} \Pr\left[Q = q\right] \sum_{\boldsymbol{u}\in\mathcal{V}^n} \Pr\left[\hat{V}^n = \boldsymbol{v} | Q = q\right]$$
$$\sum_{\boldsymbol{u}\in\mathcal{U}^n} \left| \Pr\left[\hat{U}^n = \boldsymbol{u} | \hat{V}^n = \boldsymbol{v}, Q = q\right] - \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \right| \leq 2\epsilon.$$

By a Markov bound, there exists $q \in Q$ such that,

$$\Pr\left(\hat{U}^n = \boldsymbol{u} : \sum_{\boldsymbol{v} \in \mathcal{V}^n} \left| \Pr\left[\hat{V}^n = \boldsymbol{v} | \hat{U}^n = \boldsymbol{u}, Q = q\right] - \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \right| \right.$$
$$\left. > 2\sqrt{\epsilon} \middle| Q = q \right) \leq 2\sqrt{\epsilon}, \tag{25}$$

$$\Pr\left(\hat{V}^n = \boldsymbol{v} : \sum_{\boldsymbol{u} \in \mathcal{U}^n} \left| \Pr\left[\hat{U}^n = \boldsymbol{u} | \hat{V}^n = \boldsymbol{v}, Q = q\right] - \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \right| \right.$$
$$\left. > 2\sqrt{\epsilon} \middle| Q = q \right) \leq 2\sqrt{\epsilon}. \tag{26}$$

Let $S_q(X) = S(X, q)$ and $R_q(M, Y) = R(M, (Y, q))$. We will show that $\langle S_q, R_q \rangle$ is an ϵ'-OMSR, where $\epsilon' = O(n^2\sqrt{\epsilon})$. Since Q is independent of (X^m, Y^m), the distribution over the source correlation does not change when conditioning on $Q = q$. The above inequalities will imply that this reduction satisfies security against S and R with $O(n^2\sqrt{\epsilon})$ error if we show that

$$\left(\widehat{U}^n, \widehat{V}^n | Q = q\right) \approx^{n^2\sqrt{\epsilon}} (U^n, V^n).$$

We now make a few definitions that we would use in the analysis: In the sequel, we allow the rows and columns of matrices to be indexed by sets for ease of presentation. Rows, columns and indices in the matrix will be referred to using the members from appropriate sets. For example, if M is a $\mathcal{X} \times \mathcal{Y}$ dimensional matrix, the $(M)_{x,y}$ is the entry in the matrix at row index x and column index y. Similarly, if η is a $\mathcal{X}$ dimensional row vector, then $(\mu)_x$ is the entry in the column x of the vector. We will drop the parentheses if there is no confusion. Let $\widehat{P}$ and P be correlation matrices of dimensions $\mathcal{U}^n \times \mathcal{V}^n$ corresponding to correlations $(\widehat{U}^n, \widehat{V}^n)$ and (U^n, V^n), respectively. That is, for each $\boldsymbol{u} \in \mathcal{U}^n, \boldsymbol{v} \in \mathcal{V}^n$,

$$\widehat{P}_{\boldsymbol{u},\boldsymbol{v}} = \Pr\left[\widehat{U}^n = \boldsymbol{u}, \widehat{V}^n = \boldsymbol{v} | Q = q\right] \qquad P_{\boldsymbol{u},\boldsymbol{v}} = \Pr\left[U^n = \boldsymbol{u}, V^n = \boldsymbol{v}\right].$$

Let W, W' be matrices of dimensions $\mathcal{U}^n \times \mathcal{V}^n$ and $\mathcal{V}^n \times \mathcal{U}^n$, respectively, corresponding to conditional distributions $p_{U^n|V^n}$ and $p_{V^n|U^n}$; that is, for any $\boldsymbol{u}, \boldsymbol{v}$

$$W_{\boldsymbol{u},\boldsymbol{v}} = \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \qquad W'_{\boldsymbol{v},\boldsymbol{u}} = \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right].$$

Similarly, define matrices $\widehat{W}, \widehat{W}'$ corresponding to conditional distributions $p_{\widehat{U}^n|\widehat{V}^n}$ and $p_{\widehat{V}^n|\widehat{U}^n}$ such that, for any $\boldsymbol{u}, \boldsymbol{v}$,

$$\widehat{W}_{\boldsymbol{u},\boldsymbol{v}} = \Pr\left[\widehat{V}^n = \boldsymbol{v} | \widehat{U}^n = \boldsymbol{u}, Q = q\right] \qquad \widehat{W}'_{\boldsymbol{v},\boldsymbol{u}} = \Pr\left[\widehat{U}^n = \boldsymbol{u} | \widehat{V}^n = \boldsymbol{v}, Q = q\right].$$

Finally, let μ and $\widehat{\mu}$ be the $\mathcal{U}^n$ dimensional vectors corresponding to the marginal distribution of U^n and $\hat{U}^n$, and let D and $\hat{D}$ be the $\mathcal{U}^n \times \mathcal{U}^n$ dimensional diagonal matrices corresponding to μ and $\widehat{\mu}$. That is,

$$\mu_{\boldsymbol{u}} = D_{\boldsymbol{u},\boldsymbol{u}} = \Pr\left[U^n = \boldsymbol{u}\right] \qquad \widehat{\mu}_{\boldsymbol{u}} = \widehat{D}_{\boldsymbol{u},\boldsymbol{u}} = \Pr\left[\hat{U}^n = \boldsymbol{u} | Q = q\right].$$

The above matrices satisfy the following conditions:

(i). $DW = P$ and $\widehat{D}\widehat{W} = \widehat{P}$.

(ii). $\mu W W' = \mu$ and $\widehat{\mu}\widehat{W}\widehat{W}' = \widehat{\mu}$.

For each $\boldsymbol{u}, \boldsymbol{v}$,

$$P_{\boldsymbol{u},\boldsymbol{v}} = \Pr\left[U^n = \boldsymbol{u}, V^n = \boldsymbol{v}\right] = \Pr\left[U^n = \boldsymbol{u}\right]\Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] = (DW)_{\boldsymbol{u},\boldsymbol{v}};$$

the other equality can be similarly shown to prove (i). For each $\boldsymbol{u}$, w.r.t. correlation (U^n, V^n),

$$\begin{aligned}(\mu W W')_{\boldsymbol{u}} &= \sum_{\boldsymbol{u}'} \mu_{\boldsymbol{u}'} \left(\sum_{\boldsymbol{v}} W_{\boldsymbol{u}',\boldsymbol{v}} W'_{\boldsymbol{v},\boldsymbol{u}}\right) = \sum_{\boldsymbol{u}'} \Pr\left[\boldsymbol{u'}\right] \left(\sum_{\boldsymbol{v}} \Pr\left[\boldsymbol{v}|\boldsymbol{u'}\right]\Pr\left[\boldsymbol{u}|\boldsymbol{v}\right]\right)\\ &= \sum_{\boldsymbol{u}',\boldsymbol{v}} \Pr\left[\boldsymbol{u'}, \boldsymbol{v}\right]\Pr\left[\boldsymbol{u}|\boldsymbol{v}\right] = \sum_{\boldsymbol{v}} \Pr\left[\boldsymbol{v}\right]\Pr\left[\boldsymbol{u}|\boldsymbol{v}\right] = \Pr\left[\boldsymbol{u}\right] = \mu_{\boldsymbol{u}};\end{aligned}$$

the other equality can be similarly shown to prove (ii).
Using (i),

$$\begin{aligned}&\mathsf{2TVD}\left(\left(U^n, V^n\right), \left(\widehat{U}^n, \widehat{V}^n | Q = q\right)\right)\\ &= \sum_{\boldsymbol{u},\boldsymbol{v}} |P_{\boldsymbol{u},\boldsymbol{v}} - \widehat{P}_{\boldsymbol{u},\boldsymbol{v}}|\\ &= \sum_{\boldsymbol{u},\boldsymbol{v}} |(DW)_{\boldsymbol{u},\boldsymbol{v}} - \left(\widehat{D}\widehat{W}\right)_{\boldsymbol{u},\boldsymbol{v}}|\\ &= \sum_{\boldsymbol{u},\boldsymbol{v}} |\mu_{\boldsymbol{u}}(W)_{\boldsymbol{u},\boldsymbol{v}} - \widehat{\mu}_{\boldsymbol{u}}\left(\widehat{W}\right)_{\boldsymbol{u},\boldsymbol{v}}|\\ &= \sum_{\boldsymbol{u},\boldsymbol{v}} |\widehat{\mu}_{\boldsymbol{u}}\left(W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}\right) + \hat{W}_{\boldsymbol{u},\boldsymbol{v}}(\mu_{\boldsymbol{u}} - \widehat{\mu}_{\boldsymbol{u}}) + \left(W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}\right)(\mu_{\boldsymbol{u}} - \widehat{\mu}_{\boldsymbol{u}})|.\end{aligned}$$

We will denote the matrix obtained by taking absolute of each entry in a matrix M by $|M|$; similarly, absolute value of a vector μ is denoted by $|\mu|$. We get the following upper bound:

$$\begin{aligned}&\mathsf{2TVD}\left(\left(U^n, V^n\right), \left(\widehat{U}^n, \widehat{V}^n | Q = q\right)\right)\\ &\leq \sum_{\boldsymbol{u},\boldsymbol{v}} \widehat{\mu}_{\boldsymbol{u}}|W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}| + \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}|\mu_{\boldsymbol{u}} - \widehat{\mu}_{\boldsymbol{u}}| + |W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}||\mu_{\boldsymbol{u}} - \widehat{\mu}_{\boldsymbol{u}}|\\ &\leq \left(\sum_{\boldsymbol{v}}\sum_{\boldsymbol{u}} \widehat{\mu}_{\boldsymbol{u}}|W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}|\right) + \left(\sum_{\boldsymbol{v}}\sum_{\boldsymbol{u}} |\mu_{\boldsymbol{u}} - \widehat{\mu}_{\boldsymbol{u}}|2\widehat{W}_{\boldsymbol{u},\boldsymbol{v}} + W_{\boldsymbol{u},\boldsymbol{v}}\right)\\ &\leq \sum_{\boldsymbol{v}} \left(\widehat{\mu}|W - \widehat{W}|\right)_{\boldsymbol{v}} + \sum_{\boldsymbol{v}} \left(|\mu - \widehat{\mu}|\left(2\widehat{W} + W\right)\right)_{\boldsymbol{v}}.\end{aligned}$$

Using the definitions of $\widehat{\mu}, W, \widehat{W}$, the first term in the RHS can be bounded as

$$\begin{aligned}
&\sum_{\boldsymbol{v}} \left(\widehat{\mu}|W-\widehat{W}|\right)_{\boldsymbol{v}} \\
&= \sum_{\boldsymbol{v}} \sum_{\boldsymbol{u}} \widehat{\mu}_{\boldsymbol{u}} |W_{\boldsymbol{u},\boldsymbol{v}} - \widehat{W}_{\boldsymbol{u},\boldsymbol{v}}| \\
&= \sum_{\boldsymbol{v}} \sum_{\boldsymbol{u}} \Pr\left[\hat{U}^n = \boldsymbol{u} | Q = q\right] \\
&\qquad\qquad \left| \Pr\left[\hat{V}^n = \boldsymbol{v} | \hat{U}^n = \boldsymbol{u}, Q = q\right] - \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \right| \\
&= \sum_{\boldsymbol{u}} \Pr\left[\hat{U}^n = \boldsymbol{u} | Q = q\right] \\
&\qquad \sum_{\boldsymbol{v}} \left| \Pr\left[\hat{V}^n = \boldsymbol{v} | \hat{U}^n = \boldsymbol{u}, Q = q\right] - \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \right| \leq 4\sqrt{\epsilon},
\end{aligned}$$

where the final inequality follows from Eq. (25). Next, we bound the second term. Let $B = WW'$ and $E = \widehat{W}\widehat{W}' - WW'$. For any N, repeatedly using (ii),

$$\widehat{\mu}(B+E)^N = \widehat{\mu}(\widehat{W}\widehat{W}')^N = \widehat{\mu}(\widehat{W}\widehat{W}')^{N-1} = \ldots = \widehat{\mu}(\widehat{W}\widehat{W}') = \widehat{\mu} \tag{27}$$

Hence,

$$\mu - \widehat{\mu} = \mu - \widehat{\mu}B^N - \widehat{\mu}\left((B+E)^N - B^N\right)$$

The plan going forward is as follows: The stochastic matrix B can be interpreted as the state transition matrix of a Markov process with stationary distribution μ by (ii). Using a convergence theorem for Markov chains, we can choose N to make $\widehat{\mu}B^N$ approach arbitrarily close to μ in total variation distance. Indeed, N can be chosen to be $O(n^2)$. Further, we show that $\widehat{\mu}(B+E)^N$ and $\widehat{\mu}B^N$ are close. Thus, we conclude that the total variation distance between μ and $\widehat{\mu}$ is at most ϵ' as required. The two conditions we require are stated in the following claims:

Claim. *When the correlation (U, V) has no common information, there exists a constant k that depends only on entries of p_{UV} such that, when $N = kn^2$, $\sum_{\boldsymbol{u}} |(\widehat{\mu}B^N)_{\boldsymbol{u}} - \mu_{\boldsymbol{u}}| \leq \epsilon$.*

Claim. $\sum_{\boldsymbol{u}} |\left(\widehat{\mu}(B+E)^N - \widehat{\mu}B^N\right)_{\boldsymbol{u}}| \leq 8N\sqrt{\epsilon}$.

Assuming the above claims,

$$\begin{aligned}
\sum_{\boldsymbol{u}} (|\mu - \widehat{\mu}|)_{\boldsymbol{u}} &\leq \sum_{\boldsymbol{u}} |(\widehat{\mu}B^N)_{\boldsymbol{u}} - \mu_{\boldsymbol{u}}| + |\left(\widehat{\mu}(B+E)^N - \widehat{\mu}B^N\right)_{\boldsymbol{u}}| \qquad (28) \\
&\leq \epsilon + 8N\sqrt{\epsilon}. \qquad (29)
\end{aligned}$$

We conclude by proving the claims.

Proof of the first claim. We will interpret the stochastic matrix B as the state transition matrix of a Markov process. By (ii), μ is the stationary distribution of this stochastic process. We prove the claim by showing an upper bound on the mixing time in terms of the eigenvalue gap of the Markov chain. Specifically, we use the following theorem:

Theorem D.1. *Consider an ergodic, reversible Markov chain with state space Ω, state transition matrix P and stationary distribution π. For any $\epsilon > 0$, $i \in \Omega$,*

$$\sum_{j\in\Omega} |P^{\tau}_{i,j} - \pi_j| \leq \epsilon \text{ when } \tau = \frac{1}{1-\lambda_2} \ln\left(\frac{1}{\pi^*\epsilon}\right),$$

where $\pi^* = \min_{i\in\Omega} \pi_i$ (the probability of i under distribution π) and λ_2 be the second largest Eigen value of P.

We first argue that B is reversible; i.e., $\mu_{\boldsymbol{u}} B_{\boldsymbol{u},\boldsymbol{u}'} = \mu_{\boldsymbol{u}'} B_{\boldsymbol{u}',\boldsymbol{u}}$ for all $\boldsymbol{u}, \boldsymbol{u}'$. Noting that $B = WW'$ and using the definitions of W and W',

$$\begin{aligned}
\mu_{\boldsymbol{u}} B_{\boldsymbol{u},\boldsymbol{u}'} &= \Pr\left[U^n = \boldsymbol{u}\right] \left(\sum_{\boldsymbol{v}} \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \cdot \Pr\left[U^n = \boldsymbol{u}' | V^n = \boldsymbol{v}\right]\right) \\
&= \sum_{\boldsymbol{v}} \Pr\left[U^n = \boldsymbol{u}\right] \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}\right] \cdot \Pr\left[U^n = \boldsymbol{u}' | V^n = \boldsymbol{v}\right] \\
&= \sum_{\boldsymbol{v}} \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \cdot \Pr\left[V^n = \boldsymbol{v}\right] \cdot \Pr\left[U^n = \boldsymbol{u}' | V^n = \boldsymbol{v}\right] \\
&= \sum_{\boldsymbol{v}} \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \cdot \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}'\right] \cdot \Pr\left[U^n = \boldsymbol{u}'\right] \\
&= \Pr\left[U^n = \boldsymbol{u}'\right] \left(\sum_{\boldsymbol{v}} \Pr\left[V^n = \boldsymbol{v} | U^n = \boldsymbol{u}'\right] \cdot \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right]\right) \\
&= \mu_{\boldsymbol{u}'} B_{\boldsymbol{u}',\boldsymbol{u}}.
\end{aligned}$$

We will show that B^k is ergodic for a constant k independent of n; this amounts to showing that all entries of B^k are non-zero. Correlation (U, V) has no common information, hence, the correlation graph of (U, V) is connected. Correlation graph is the bipartite graph on $\mathcal{U} \cup \mathcal{V}$ with an edge between u and v if $\Pr[u, v] > 0$. Let P_1, W_1 and W_1' be the correlation matrix of p_{XY}, and conditional distribution matrices for $p_{U|V}$ and $p_{V|U}$, respectively (defined analogous to P, W and W'). Since the correlation graph of (U, V) is connected, (interpreting W_1, W_1' as state transition matrices), W and W' correspond to irreducible Markov processes. Hence, state transition matrix $B_1 = W_1 W_1'$ also correspond to an irreducible Markov process. Hence, there exists a constant k determined entirely by the description of p_{UV} for which $(B_1^k)_{\boldsymbol{u},\boldsymbol{u}'} > 0$ for all $\boldsymbol{u}, \boldsymbol{u}'$. Finally, observe that $P = P_1^{\otimes n}$, $W = W_1^{\otimes n}$ and $W' = W_1^{\otimes n}$, where $P_1^{\otimes n}$ is the n fold tensor product of P_1 with itself and so on. Consequently, $B = B_1^{\otimes n}$. Thus, irrespective of the value of n, all entries of B^k are non-zero, and the corresponding Markov process is ergodic. Since B is irreversible, so is B^k.

At this point we invoke Theorem D.1 with state transition matrix B^k and stationary distribution μ (note that the stationary distribution of B and B^k are the same). Since the largest eigenvalue of any state transition matrix is 1, the second largest eigenvalue of B^k coincides with the second largest eigenvalue of B_1^k say λ_2. This roots from the fact that the Eigen value of $M \otimes M$ is the set of all pairwise products of Eigen values of M. Finally,

$$\mu^* = \min_{\boldsymbol{u}\in\mathcal{U}^n} \Pr\left[U^n = \boldsymbol{u}\right] = \left(\min_{u\in\mathcal{U}} \Pr\left[U = u\right]\right)^n.$$

Thus, for all $\boldsymbol{u}$,

$$\sum_{\boldsymbol{u}'} |(B^\tau)_{\boldsymbol{u},\boldsymbol{u}'} - \mu_{\boldsymbol{u}'}| \leq \epsilon \text{ when } \tau = \frac{n}{1-\lambda_2^k} \ln\left(\frac{1}{\epsilon \cdot \min_u(\Pr[U=u])}\right). \tag{30}$$

Let B' be a $\mathcal{U}^n \times \mathcal{U}^n$ dimensional stochastic matrix such that $(B')_{\boldsymbol{u},\boldsymbol{u}'} = \mu_{\boldsymbol{u}'}$ for all $\boldsymbol{u}, \boldsymbol{u}'$. Letting $N = \tau$,

$$\sum_{\boldsymbol{u}} |(\widehat{\mu} B^N)_{\boldsymbol{u}} - \mu_{\boldsymbol{u}}| = \sum_{\boldsymbol{u}} |(\widehat{\mu} B^N)_{\boldsymbol{u}} - (\widehat{\mu} B')_{\boldsymbol{u}}| \leq \sum_{\boldsymbol{u}} \left(\widehat{\mu}|B^N - B'|\right)_{\boldsymbol{u}} \leq \epsilon.$$

The last inequality follows the fact that each row of $|B^n - B'|$ sums up to at most ϵ (by Eq. (30)), and that the sum of all entries of (non-negative) vector μ is 1. □

Proof of the second claim. We have

$$\begin{aligned} B^N &= (B+E)B^{N-1} - EB^{N-1} \\ &= (B+E)^2 B^{N-2} - (B+E)EB^{N-2} - EB^{N-1} \\ &= (B+E)^3 B^{N-3} - (B+E)^2 EB^{N-3} - (B+E)EB^{N-2} - EB^{N-1} \\ &\vdots \\ &= (B+E)^N - \sum_{i=1}^{N} (B+E)^{i-1} EB^{N-i}. \end{aligned}$$

Since $\widehat{\mu}(B+E)^i = \widehat{\mu}(\widehat{W}\widehat{W'})^i = \widehat{\mu}$ for all i,

$$\widehat{\mu}(B+E)^N - \widehat{\mu}B^N = \sum_{i=1}^{N} \widehat{\mu}(B+E)^{i-1} EB^{N-i} = \sum_{i=1}^{N} \widehat{\mu} EB^{N-i}.$$

We have $\widehat{\mu}E = \widehat{\mu}\widehat{W}\widehat{W}' - \widehat{\mu}WW' = \widehat{\mu} - \widehat{\mu}WW'$. Hence,

$$\begin{aligned} \widehat{\mu}E = \widehat{\mu} - \widehat{\mu}WW' &= \widehat{\mu} - \widehat{\mu}(\widehat{W} + W - \widehat{W})(\widehat{W}' + W' - \widehat{W}') \\ &= \widehat{\mu} - \widehat{\mu}\widehat{W}\widehat{W}' - \widehat{\mu}\widehat{W}(W' - \widehat{W}') - \widehat{\mu}(W - \widehat{W})W' \\ &= \widehat{\mu}\widehat{W}(W' - \widehat{W}') - \widehat{\mu}(W - \widehat{W})W'. \end{aligned} \tag{31}$$

Using the definitions of $\widehat{\mu}, W', \widehat{W}'$ and $\widehat{W}$,

$$\begin{aligned}
&\sum_{\boldsymbol{u}} \left| \left(\widehat{\mu}\widehat{W}(W' - \widehat{W}')\right)_{\boldsymbol{u}} \right| \\
&\leq \sum_{\boldsymbol{u}} \left(\widehat{\mu}\widehat{W}|W' - \widehat{W}'|\right)_{\boldsymbol{u}} \\
&= \sum_{\boldsymbol{u}} \sum_{\boldsymbol{v}} \left(\widehat{\mu}\widehat{W}\right)_{\boldsymbol{v}} |W'_{\boldsymbol{v},\boldsymbol{u}} - \widehat{W}'_{\boldsymbol{v},\boldsymbol{u}}| \\
&= \sum_{\boldsymbol{u}} \sum_{\boldsymbol{v}} \left(\sum_{\boldsymbol{u}'} \Pr\left[\hat{U}^n = \boldsymbol{u}' | Q = q\right] \Pr\left[\hat{V}^n = \boldsymbol{v} | \hat{U}^n = \boldsymbol{u}', Q = q\right] \right) \\
&\qquad\qquad \left| \Pr\left[\hat{U}^n = \boldsymbol{u} | \hat{V}^n = \boldsymbol{v}, Q = q\right] - \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \right| \\
&= \sum_{\boldsymbol{u}} \sum_{\boldsymbol{v}} \Pr\left[\hat{V}^n = \boldsymbol{v} | Q = q\right] \\
&\qquad\qquad \left| \Pr\left[\hat{U}^n = \boldsymbol{u} | \hat{V}^n = \boldsymbol{v}, Q = q\right] - \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \right| \\
&= \sum_{\boldsymbol{v}} \Pr\left[\hat{V}^n = \boldsymbol{v} | Q = q\right] \\
&\qquad\qquad \sum_{\boldsymbol{u}} \left| \Pr\left[\hat{U}^n = \boldsymbol{u} | \hat{V}^n = \boldsymbol{v}, Q = q\right] - \Pr\left[U^n = \boldsymbol{u} | V^n = \boldsymbol{v}\right] \right| \leq 4\sqrt{\epsilon}, \qquad (32)
\end{aligned}$$

where the final inequality follows from Eq. (26). We have already shown that

$$\sum_{\boldsymbol{v}} \left(\widehat{\mu}|W - \widehat{W}|\right)_{\boldsymbol{v}} \leq 4\sqrt{\epsilon}. \qquad (33)$$

If M is a stochastic matrix, for any vector μ' (with non-negative entries),

$$\sum_{\boldsymbol{u}} (\mu' M)_{\boldsymbol{u}} \leq \gamma \sum_{\boldsymbol{u}} (\mu)_{\boldsymbol{u}}. \qquad (34)$$

Hence, using Eqs. (32) to (34) in Eq. (31),

$$\begin{aligned}
\sum_{\boldsymbol{u}} |\widehat{\mu}E|_{\boldsymbol{u}} &\leq \sum_{\boldsymbol{u}} \left| \left(\widehat{\mu}\widehat{W}(W' - \widehat{W}')\right)_{\boldsymbol{u}} \right| + \sum_{\boldsymbol{u}} \left| \left(\widehat{\mu}(W - \widehat{W})W'\right)_{\boldsymbol{u}} \right| \\
&\leq 4\sqrt{\epsilon} + \sum_{\boldsymbol{v}} \left(\widehat{\mu}|W - \widehat{W}|\right)_{\boldsymbol{v}} \leq 8\sqrt{\epsilon}.
\end{aligned}$$

Noting that B^{N-i} is stochastic, we conclude that

$$\sum_{\boldsymbol{u}} \left| \left(\widehat{\mu}(B + E)^N - \widehat{\mu}B^N\right)_{\boldsymbol{u}} \right| \leq \sum_{i=1}^{N} \sum_{\boldsymbol{u}} |\widehat{\mu}E|_{\boldsymbol{u}} \leq 8N\sqrt{\epsilon}.$$

□

This concludes the proof of the claim and the theorem.

E Remarks on Correlations that Lack Common Randomness

In the informal Theorems 2.3 and 2.4 we talked about correlations that lack common randomness. However, in the formal versions of them, we meant two different notions. In Theorem 6.8 we meant that $S^*(X,Y) < 1$, and in Theorem 7.1 we meant that there don't exist functions f, g such that $f(X) = g(Y)$ with probability 1 where $H(f(X)) > 0$. We will show that these two notions are in fact equivalent.

We will show that for a correlation (X,Y) it holds that $S^*(X,Y) = 1$ if and only if there exist deterministic functions f and g such that $f(X) = g(Y)$ with probability 1 and $H(f(X)) > 0$.
$\Rightarrow$: If $\mathrm{S}^*(X,Y) = 1$, from the discussion in Appendix C.2, we know that there exists U such that $U = f(X)$ (for some function f) and it holds that

$$\frac{I(U;Y)}{I(U;X)} = \frac{H(f(X)) - H(f(X)|Y)}{H(f(X))} = 1.$$

This means that $H(f(X)|Y) = 0$. In other words, there exists a function g such that $g(Y) = f(X)$ with probability 1. Note also, that from the definition of S^* it must be that $H(U) > 0$.
$\Leftarrow$: Let us consider $U = f(X)$. It holds that $\mathrm{S}^*(X,Y) \geq \frac{I(U;Y)}{I(U;X)} = \frac{H(g(Y))}{H(f(X)} = 1$. Since $\mathrm{S}^*(X,Y) \leq 1$, we conclude that $\mathrm{S}^*(X,Y) = 1$.

F Remarks on Hyper-Contractivity

If $\mathrm{S}^*(X,Y) < 1$, there exist $1 < q < p$ such that, defining $p' = \frac{p}{p-1}$, for any pair of functions $f : \mathcal{X}^m \to \mathbb{R}$ and $g : \mathcal{Y}^m \to \mathbb{R}$,

$$\mathbb{E}_{X^m,Y^m}[f(X^m) \cdot g(Y^m)] \leq \left(\mathbb{E}_{X^m}[f^{p'}(X^m)]\right)^{\frac{1}{p'}} \left(\mathbb{E}_{Y^m}[g^q(Y^m)]\right)^{\frac{1}{q}}.$$

To show this, we will use the following simplified definitions and facts from [35] and [5]:

- We will say $(p,q) \in \mathcal{R}_{X,Y}$, for $1 \leq q \leq p$ if and only if for any $f : \mathcal{X} \to \mathbb{R}$ and $g : \mathcal{Y} \to \mathbb{R}$ it holds that

$$\mathbb{E}_{X,Y}[f(X) \cdot g(Y)] \leq \left(\mathbb{E}_{X}[f^{p'}(X)]\right)^{\frac{1}{p'}} \left(\mathbb{E}_{Y}[g^q(Y)]\right)^{\frac{1}{q}}.$$

- We will define $q^*_{X,Y}(p) = inf\{1 \leq q : (p,q) \in \mathcal{R}_{X,Y}\}$.
- $\lim_{p\to\infty} \frac{q^*_{X,Y}(p)-1}{p-1} = \mathrm{S}^*(X,Y)$.
- $\mathrm{S}^*(X^m, Y^m) = \mathrm{S}^*(X,Y)$, for any positive integer m.

Fix any positive integer m. We know that $\lim_{p\to\infty} \frac{q^*_{X^m,Y^m}(p)-1}{p-1} = \mathrm{S}^*(X^m,Y^m) = \mathrm{S}^*(X,Y) < 1$. Thus, there exists $p > 1$ such that $\frac{q^*_{X^m,Y^m}(p)-1}{p-1} < 1$, hence, $q^*_{X^m,Y^m}(p) < p$. This means that $inf\{1 \leq q : (p,q) \in \mathcal{R}_{X^m,Y^m}\} < p$, thus, there exists $q' < p$ such that $(p,q') \in \mathcal{R}_{X^m,Y^m}$. This finishes our proof.

G Secure Reduction converting OT to n-UV

In this section, we present a protocol that achieves the asymptotic lower bound proven in Theorem 6.8. Before showing the protocol we will first describe a few sub-protocols to help us out.

Basic protocol for generating n-UV. Given source correlation 1-out-of-n OT over n-bit strings, we can generate a n-bit unit vector correlation with n^2 bits of communication. This can be done by the sender to sample a random n-bit string u and a random permutation $\pi \in S_n$, and to run an OT protocol with the secrets - $(s_1, \ldots, s_n) = (u \oplus e_{\pi(1)}, \ldots, u \oplus e_{\pi(n)})$. To do it with an OT correlation of $(r_1, \ldots, r_n)$, the sender can send the receiver the vector $(k_1, \ldots, k_n) = (s_1 - r_1, \ldots, s_n - r_n)$ and now the receiver can compute $s_b = k_b + r_b$. This requires communication of n^2 bits.

Converting 1-out-of-2 OT over $\mathbb{F}_2$ to 1-out-of-k OT over k-bit string. This protocol has two stages.

- Create 1-out-of-2 OT over k-bit string: This can be done in the following way. By using k copies of 1-out-of-2 OT over 1-bit string $\{((r_0^i, r_1^i), (b^i, r_b^i))\}_{i \in [k]}$, the receiver will sample $b \leftarrow \{0, 1\}$ in random. He will send to the sender $\mathbf{b} = (b \oplus b^1, \ldots, b \oplus b^k)$. The sender will "flip" the secrets of the i^{th} OT copy if in the i^{th} position of $\mathbf{b}$ there is a 1. In this way, after the flip, the receiver will have the b^{th} secret of every copy. Afterwards, each party will concatenate the bits of r^i they have. This will result in a 1-out-of-2 OT over k.
- Create 1-out-of-k OT over k-bit string: This can be done in the following simple way. By using k copies of 1-out-of-2 OT over k-bit string the receiver will choose j uniformly over $[k]$. Then the two parties will run k times a 1-out-of-2 OT over k-bit string protocol where the receiver's choice bits will be 1 on the j^{th} run, and 0 elsewhere. We can run this protocol using the OT correlation with $O(k)$ bit communication per run. The receiver will output j and the secret he got from the j^{th} run. The sender's secrets would be the second secret in each run.

Overall the communication cost of this construction is $O(k^2)$ bits.

Full protocol. We will now describe a (two-way communication) protocol which converts from source correlation 1-out-of-2 OT to n-UV.

Protocol:

- Create two 1-out-of-$\sqrt{n}$ OT over $\sqrt{n}$-bit string.
- Create two $\sqrt{n}$-bit unit vector correlation with the basic protocol using the generated string OT and get $(u_1, v_1), (u_2, v_2)$ for the unit vectors e_{i_1} and e_{i_2}.
- Compute shares of $e_{i_1} \cdot e_{i_2}^T$ using the 1-out-of-2 OT over bit.

The overall communication of this protocol is $O(n)$. This result complements the $\Omega(n)$ bound that we got.