Secure Computation based on Leaky Correlations: High Resilience Setting

Alexander R Block* Hemanta K. Maji† Hai Nguyen‡

February 9, 2017

Abstract

Correlated private randomness, or correlation in short, is a fundamental cryptographic resource that helps parties compute securely over their private data. An offline preprocessing step, which is independent of the eventual secure computation, generates correlated secret shares for the parties and the parties use these shares during the final secure computation step. However, these secret shares are vulnerable to leakage attacks.

Inspired by the quintessential problem of privacy amplification, Ishai, Kushilevitz, Ostrovsky, and Sahai (FOCS 2009) introduced the concept of correlation extractors. Correlation extractors are interactive protocols that take leaky correlations as input and produce secure independent copies of oblivious transfer (OT), the building blocks of secure computation protocols. Although their initial feasibility result is resilient to linear leakage and produces a linear number of "fresh" OTs, the constants involved are minuscule. The output of this correlation extractor can be used to perform only small secure computation tasks, because the number of OTs needed to evaluate a functionality securely is roughly proportional to its circuit size. Recently, Gupta, Ishai, Maji, and Sahai (CRYPTO 2015) constructed an extractor that is resilient to 1/4 fractional leakage and has near-linear production rate. They also constructed an extractor from a large correlation that has 1/2 fractional resilience but produces only one OT, which does not suffice to compute even constant size functionalities securely.

In this paper, we show the existence of a correlation that produces n-bit shares for the parties and allows the extraction of $n^{1-o(1)}$ secure OTs, despite $n/2$ bits of leakage. The key technical idea is to embed several multiplications over a field into one multiplication over an extension field. The packing efficiency of this embedding directly translates into the production rate of our correlation extractor. Our work establishes a connection between this problem and a rich vein of research in additive combinatorics on constructing dense sets of integers that are free of arithmetic progressions, a.k.a. 3-free sets. We introduce a new combinatorial problem that suffices for our multiplication embedding, and produces concrete embeddings that beat the efficiency of the embeddings inspired by the reduction to 3-free sets.

Finally, the paper introduces a graph-theoretic measure to upper-bound the leakage resilience of correlations, namely the *simple partition number*. This measure is similar in spirit to graph covering problems like the biclique partition number. If the simple partition number of a correlation is 2^λ, then it is impossible to extract even one OT if parties can perform λ-bits of leakage. We compute tight estimates of the simple partition number of several correlations that are relevant to this paper, and, in particular, show that our extractor and the extractor for the large correlation by Gupta et al. have optimal leakage resilience and (qualitatively) optimal simulation error.

*Department of Computer Science, Purdue University. block9@purdue.edu.
†Department of Computer Science, Purdue University. hmaji@purdue.edu.
‡Department of Computer Science, Purdue University. nguye245@purdue.edu.

Contents

1 Introduction

Secure multi-party computation [Yao82, GMW87] helps mutually distrusting parties to compute securely over their private data. Unfortunately, it is impossible to securely compute most functionalities in the information-theoretic plain model even against parties who honestly follow the protocol but are curious to find additional information about the other parties' private input [Dol82, Kil88, IL89, Kus89, Bea89, MPR09, KMQR09]. However, we can securely compute any functionality if honest parties are in the majority [BOGW88, CCD88, RBO89, DI06], parties use some trusted setup [CLOS02, Kat07, IPS08, CGS08, MS08, DNW08, GIS+10] or correlated private randomness [Kil00, WW06, CMW05, MPR12], or there are bounds on the computational power of the parties [GMW87, IPS08].

The study of secure computation using correlated private randomness, primarily initiated due to efficiency concerns, has produced several success stories, for example FairPlay [MNPS04, BDNP08], TinyOT [NNOB12] and SPDZ [DPSZ12] (pronounced, Speedz). These secure computation protocols offload most of the computational and cryptographic complexity to an offline preprocessing phase. During this preprocessing phase, a trusted dealer samples two shares (r_A, r_B) from the joint distribution (R_A, R_B), namely the *correlated private randomness*, or *correlation* in short, and provides the secret shares r_A to Alice and r_B to Bob. During the online secure computation phase, parties use their respective secret shares in an interactive protocol to securely compute the intended functionality. Note that the preprocessing phase is independent of the functionality or the inputs fed to the functionality by the parties.

A prominent and extremely well-studied correlation is the *random oblivious transfer correlation*, represented by ROT. It samples three bits x_0, x_1, b independently and uniformly at random, and provides the secret shares (x_0, x_1) to Alice and (b, x_b) to Bob. Note that Alice does not know the choice bit b, and Bob does not know the other bit $x_{\overline{b}}$. Intuitively, ROT is an input-less functionality that implements a randomized version of *oblivious transfer* functionality, where the sender sends (x_0, x_1) as input to the functionality and the receiver picks x_b out of the two input bits. Given, m independent samples from this distribution, parties can securely compute any functionality with circuit complexity (roughly) m. For example, we can utilize the randomized self-reducibility of oblivious transfer to reimagine the GMW protocol [GMW87] in this framework naturally.

However, the storage of the secret shares by the parties brings to fore several vulnerabilities. For instance, parties can leak additional information from the secret shares of the other parties. We emphasize that the leakage need not necessarily reveal individual bits of the other party's share. The leakage can be on the entire share and encode crucial global information that can potentially jeopardize the security of the secure computation protocol.

To address these concerns, Ishai, Kushilevitz, Ostrovsky, and Sahai [IKOS09], introduced the notion of *correlation extractors*. Correlation extractors distill leaky correlations into independent samples of the ROT correlation that are secure. That is, for each of the new samples Alice does not know Bob's choice bit and Bob does not know Alice's other bit. This problem is a direct analog of the quintessential problems of privacy amplification and randomness extraction problems in the secure computation setting. With the exception that, correlation extractors ensure security against insider attacks, i.e., the parties who perform the leakage are participants in the secure protocol itself. This additional requirement makes the task of correlation extraction significantly more challenging. It is, thus, not surprising that relatively few results are known in the field of correlation extractor construction.

For example, in the setting of privacy amplification, if Alice and Bob start with a secret n-bit string then, in the presence of t-bits of arbitrary leakage to an eavesdropper, parties can re-establish a fresh m-bit secret key such that the advantage of the eavesdropper in guessing the secret key is

	Correlation Description	Number of OTs Produced (m)	Number of Leakage bits (t)	Simulation Error (ε)	Round Complexity
IKOS [IKOS09]	$\mathsf{ROT}^{n/2}$	αn	βn	$2^{-\gamma n}$	4
GIMS [GIMS15]	$\mathsf{ROT}^{n/2}$	$n/\operatorname{poly}\log n$	$(1/4 - g)n$	$2^{-gn/m}$	2
	$\mathsf{IP}\left(\mathbb{GF}\left[2\right]^n\right)$	1	$(1/2 - g)n$	2^{-gn}	2
Our Work	$\mathsf{IP}\left(\mathbb{F}^{n/\log\lvert\mathbb{F}\rvert}\right)$	$n^{1-o(1)}$	$(1/2 - g)n$	2^{-gn}	2

Figure 1: A qualitative summary of prior relevant works in correlation extractors and a comparison to our correlation extractor construction. All correlations have been normalized so that each party gets an n-bit secret share. The positive constants α, β, and γ are minuscule. And $g < 1/2$ is an arbitrary positive constant.

roughly $2^{-\Delta} \approx 2^{-(n-t-m)}$. Intuitively, the sum of "entropy deficiency" (t), "entropy of production" (m), and "$-\log$ of the adversarial advantage" (Δ) is roughly n, the initial entropy of the secret. Analogous results also exist in the setting of randomness extraction, where we can extract nearly all of the min-entropy of a source. But similar tight extraction results are not known for correlation extractors. In fact, the task of designing correlations that simultaneously support high leakage resilience and production rate with exponential security has been elusive.

The number of the output ROT samples and their high security are crucial for the secure computation protocol. For example, protocols with exponential security can increase the statistical security parameter only slightly to prohibitively increase the effort needed by adversaries to break them. Furthermore, the number of these ROT samples limit the size of the eventual functionality that can be securely computed, because the number of ROT samples needed to implement a functionality securely is directly proportional to its circuit size. As highlighted in [GIMS15], the initial feasibility result of Ishai et al. [IKOS09], though asymptotically linear in leakage resilience and production rate, has unsatisfactorily low resilience and production rate for realistic values of n, the size of the original share of the parties. The subsequent work of Gupta et al. [GIMS15], improves the resilience to (roughly) $n/4$ but trades-off the security of the protocol for high production rate and, consequently, achieves only negligible (and, not exponentially low) insecurity. They also consider a new correlation, namely the *inner-product correlation* where the secret shares of the parties are random n-bit binary vectors subject to the constraint that they are orthogonal to each other.[1] They construct a correlation extractor for the inner-product correlation with resilience $n/2$ and exponential security. However, it is inherently limited to producing one ROT sample as output, which is not adequate for the end goal of performing interesting secure computations. Our work shows that the inner-product correlation over an *appropriately large field* admits a correlation extractor that is resilient to $n/2$ bits of leakage, has high concrete production rate, and has exponentially high security. Figure 1 summarizes the entire preceding discussion tersely. Finally, similar to Gupta et al. [GIMS15], although our construction is stated in the information-theoretic setting, it is also relevant to the setting where computationally secure protocol generate the correlations or use the output OTs.

However, is the upper-bound of $n/2$ resilience inherent to the inner-product correlation? For example, $n/2$ samples of the ROT correlation cannot be resilient to more than $n/4$ bits of leakage.

[1] The actual inner-product correlation is defined slightly differently. Parties get shares $(x_0, x_1, \ldots, x_n)$ and $(y_0, y_1, \ldots, y_n)$ such that $x_0 + y_0 = \sum_{i=1}^{n} x_i y_i$. That is, x_0 and y_0 are additive secret shares of the inner product of $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$. But for intuition, it suffices to consider the correlation where the secret shares of the parties are orthogonal vectors instead.

Correlation Description	Secret Share Size (s)	Simple Partition Number (sp)	Upper Bound on the Max. Fractional Leakage ($\log \mathsf{sp}/s$)
$\mathsf{ROT}^{n/2}$	n	$2^{n/4}$	$1/4$
$\mathsf{ROLE}\,(\mathbb{F})^{n/2}$	$n \log \lvert\mathbb{F}\rvert$	$\lvert\mathbb{F}\rvert^{n/4}$	$1/4$
$\mathsf{IP}\,(\mathbb{F}^n)$	$n \log \lvert\mathbb{F}\rvert$	$\lvert\mathbb{F}\rvert^{n/2}$	$1/2$

Figure 2: A summary of the estimates of the simple partition number for the correlations relevant to our work.

A partition argument can demonstrate this upper bound of the maximum resilience of this correlation [IMSW14]. In this partition argument, Alice emulates the generation of $n/4$ (i.e., half of $n/2$) independent samples (x_0, x_1) and (c, x_c) from the ROT correlation and sends the corresponding (c, x_c) to Bob. Moreover, Bob emulates the generation of the remaining $n/4$ samples and sends the corresponding (x_0, x_1) shares to Alice. Finally, we reimagine any correlation extractor that is resilient to $n/4$ bits of leakage and produces even one secure ROT sample as a secure ROT protocol in the plain model where Alice implements $n/4$ ROT samples, and Bob implements the remaining $n/4$ ROT samples; which is impossible. Typically, the partition argument applies to "multiple independent samples of small correlations," but its extension to one huge global correlation is not apparent.

To address this question, we introduce a new graph-theoretic measure for the maximum resilience of a correlation, namely its *simple partition number*. In particular, a correlation with simple partition number $\leqslant 2^{\lambda}$ cannot be resilient to λ bits of leakage (refer to Figure 2 for a summary of these estimates). Finally, we prove the optimality of the resilience demonstrated by the correlation extractors for the inner-product correlation presented in [GIMS15] and our work. Refer to Section 5.7 for a discussion on how the relation between simple partition number and maximum resilience is similar to the connection between biclique partition number and Wyner's common information [Wyn75]. The existence of correlation extractors for a slightly lesser amount of leakage implies the tightness of our upper bounds on leakage resilience. Finally, we leverage the simple partition number bounds and use an averaging argument to show that the decay in simulation security with entropy gap as achieved by [GIMS15] and our correlation extractor are qualitatively optimal.

1.1 Model

This section presents the standard model of Ishai et al. [IKOS09] for correlation extractors, which subsequent works also use. We consider 2-party semi-honest secure computation in the preprocessing model. In the preprocessing step, a trusted dealer draws a sample (r_A, r_B) from the joint distribution (R_A, R_B). The joint distribution (R_A, R_B) is referred to as the correlated private randomness, and r_A and r_B, respectively, are the secret shares of Alice and Bob. The dealer provides the secret share r_A to Alice and r_B to Bob. An adversarial party can perform arbitrary t-bits of leakage on the secret share of the other party at the end of the preprocessing step.

We represent this leaky correlation hybrid as $(R_A, R_B)^{[t]}$. In the leaky correlation $(R_A, R_B)^{[t]}$ hybrid, during the secure computation phase, parties perform an interactive protocol to realize their target functionality securely. No leakage occurs during the execution of the secure computation protocol. In this work, we consider the functionality that implements m independent oblivious transfers between the parties, referred to as the OT^m functionality.

Definition 1 (Correlation Extractor)**.** *Let (R_A, R_B) be a correlated private randomness such that*

the secret share size of each party is n-bits. An (n, m, t, ε)-correlation extractor for (R_A, R_B) *is a two-party interactive protocol in the* $(R_A, R_B)^{[t]}$ *hybrid that securely implements the* OT^m *functionality against information-theoretic semi-honest adversaries with* ε*-simulation error.*

1.2 Our Contribution

Our work makes a two-fold contribution to enhancing the state-of-the-art for correlation extractors. First, we construct a highly resilient correlation extractor that produces a large number of secure OTs as output and has exponential security. Finally, we provide a general graph-theoretic measure that upper bounds the maximal resilience of any correlation.

1.2.1 Correlation Extraction Construction.

For any field $(\mathbb{F}, +, \cdot)$, the *inner-product correlation over* $\mathbb{F}^{n+1}$, represented by $\mathsf{IP}\left(\mathbb{F}^{n+1}\right)$, is a correlation that samples random $r_A = (x_0, x_1, \ldots, x_n) \in \mathbb{F}^{n+1}$ and $r_B = (y_0, y_1, \ldots, y_n) \in \mathbb{F}^{n+1}$ such that $x_0 + y_0 = \sum_{i=1}^{n} x_i y_i$. That is, x_0 and y_0 are the additive secret shares of the inner product of $x_{[n]} := (x_1, \ldots, x_n)$ and $y_{[n]} := (y_1, \ldots, y_n)$. Gupta et al. [GIMS15] consider a special case of the inner-product correlation, where $\mathbb{F} = \mathbb{GF}[2]$. Note that each party receives $(n+1)$ field elements as its secret share. In particular, if $\mathbb{F} = \mathbb{GF}[2^a]$, then each party gets an $a(n+1)$-bit secret share.

Theorem 1 (High Resilience High Production Correlation Extractor)**.** *For all constants* $0 < \delta < g < 1/2$*, there exists a correlation* (R_A, R_B)*, where each party gets n-bit secret share, such that there exists a two-round* (n, m, t, ε)*-correlation extractor for* (R_A, R_B)*, where* $m = (\delta n)^{1-o(1)}$*,* $t = (1/2 - g)n$*, and* $\varepsilon = 2^{-(g-\delta)n/2}$*.*

We use $(R_A, R_B) = \mathsf{IP}\left(\mathbb{GF}\left[2^{\delta n}\right]^{1/\delta}\right)$ in this theorem. Note that we maintain the dependence on δ explicitly in the theorem statement to enable computation of concrete efficiency. As we shall see later, this theorem achieves high production rate of $(\delta n)^{\log 2/\log 3} \approx (\delta n)^{0.63}$ even for realistic values of n. The simulation error is exponentially low in the difference between the entropy gap gn and the parameter δn. Our construction achieves $(\delta n)^{1-o(1)}$ production asymptotically, which is close to the ideal target of δn production. Qualitatively, the decay in our simulation error is near optimal as demonstrated by Theorem 2 and Corollary 3.

The crux of our construction is the composition of two technical contributions. First, we observe that the correlation extractor for $\mathsf{IP}\left(\mathbb{GF}[2]^n\right)$ constructed by Gupta et al. [GIMS15] extends to the $\mathsf{IP}\left(\mathbb{F}^{1/\delta}\right)$ correlation, where $\mathbb{F}$ is a large field. However, in this case, instead of producing a secure OT, it produces a generalization of oblivious transfer, namely *oblivious linear-function evaluation over* $\mathbb{F}$ [WW06] (represented as $\mathsf{OLE}(\mathbb{F})$). An oblivious linear-function evaluation is a 2-party functionality that takes $(A, B) \in \mathbb{F}$ as input from Alice and $X \in \mathbb{F}$ as input from Bob, and provides $Z = AX + B$ as output to Bob. Note that oblivious transfer is equivalent to oblivious linear-function evaluation over $\mathbb{GF}[2]$, because $x_b = (x_1 - x_0)b + x_0$, for $x_0, x_1, b \in \mathbb{GF}[2]$.

Finally, we embed m OT evaluations simultaneously into one $\mathsf{OLE}(\mathbb{F})$ evaluation. Note that, this is *not* an asymptotic reduction. Asymptotically, there are several techniques to construct multiple copies of OT using multiple copies of OLE at a good rate. Development of more efficient embeddings will directly improve the production rate of our construction. We demonstrate that dense sets of integers that avoid any arithmetic progressions, 3-free sets, provide such embedding of multiplications. We formulate a relaxed version of this combinatorial problem (see Figure 5) that suffices for our embedding problem and obtain more efficient embeddings that those that are inspired by the 3-free set constructions.

We emphasize that although we state our correlation extractor for the bounded leakage model, i.e. an adversary can perform at most t-bits of leakage, it also extends to the noisy leakage setting. As long as the noise is high enough to maintain $(n-t)$ bits of (average) min-entropy in the secret share of the parties, our extractor construction remains secure.

1.2.2 Bound on the Maximum Resilience.

The construction of Theorem 1 and the correlation extractor of Gupta et al. [GIMS15], with fractional resilience $1/2$, lead naturally to a fascinating question. Can there exist a correlation extractor for $\mathsf{IP}(\mathbb{F}^n)$ that achieves over $1/2$ fractional resilience? In fact, more generally, can we meaningfully upper-bound the maximum leakage resilience of an arbitrary correlation?

Note that if parties obtain multiple independent samples from identical correlation, then the partition argument can be leveraged to deduce an upper bound. For example, either Alice or Bob by getting adequate information on half of the other party's secret shares can break the security of the correlation extractor protocol. As discussed earlier, this argument implies that the correlation $\mathsf{ROT}^{n/2}$ is not resilient to $\lceil n/4 \rceil$ bits of leakage, because every ROT hides only one bit of information from each party [IMSW14]. However, this approach does not apply to correlation extractors for secret shares drawn from one large correlation, for example, $\mathsf{IP}(\mathbb{F}^n)$. We prove the following main result.

Theorem 2 (Hardness of Correlation Extraction)**.** *Let $(\mathbb{F}, +, \cdot)$ be an arbitrary field. There exists a universal constant $\varepsilon^* > 0$ such that, for $(R_A, R_B) = \mathsf{IP}\left(\mathbb{F}^k\right)$, any $(n, 1, (n/k)\lceil (k+1)/2 \rceil, \varepsilon)$-correlation extractor for (R_A, R_B) has $\varepsilon \geqslant \varepsilon^*$, where $n = k \log |\mathbb{F}|$.*

This result proves the optimality of the leakage resilience achieved by our extractor in Theorem 1 and the correlation extractor for $\mathsf{IP}\left(\mathbb{GF}[2]^n\right)$ proposed by Gupta et al. [GIMS15]. In fact, a more general version of this result (using averaging arguments) shows that any $(n, 1, n/2 - gn, \varepsilon)$-correlation extractor for $\mathsf{IP}\left(\mathbb{F}^k\right)$ has $\varepsilon \geqslant \varepsilon^* 2^{-gn}$ (see Corollary 3). This result proves the qualitative optimality of simulation error achieved by these two correlation extractors.

The technical heart of this result is a new graph-theoretic measure for maximum leakage resilience in correlations, namely *simple partition number* (see Definition 4 in Section 2). Theorem 2 is a consequence of precise estimation of this quantity for the $\mathsf{IP}(\mathbb{F}^n)$ correlation. This quantity is similar in spirit to the biclique partition number of a graph [GP71, GP72], the minimum number of bicliques needed to partition the edges of a graph. Moreover, the connection of simple partition number to maximum resilience is intuitively analogous to the link between biclique partition number and Wyner's common information [Wyn75]. Section 5.7 provides details on this connection.

1.3 Prior Relevant Works

This work lies at the intersection of several fields like correlation extractors, additive combinatorics, graph covering problems, and information theory. In this section, we provide only a summary of the work on combiners and extractors. The prior relevant works related to the remaining topics are covered in appropriate sections later.

1.3.1 Combiners and Extractors.

A closely related concept is the notion of OT combiners, which are a restricted variant of OT extractors in which the leakage is limited to local information about individual OT correlations, and there is no global leakage. The study of OT combiners was initiated by Harnik et al. [HKN+05]. Since then, there has been work on several variants and extensions of OT combiners [HIKN08,

$\mathsf{ROLE}(\mathbb{F})$	Given a field $\mathbb{F}$, Alice receives $r_A = (A, B)$ and Bob receives $r_B = (X, Z)$ such that A, B, X are independently and uniformly sampled from $\mathbb{F}$ and $Z = A \cdot X + B$.
$\mathsf{IP}(\mathbb{F}^n)$	Given a field $\mathbb{F}$, Alice receives $r_A = (x_0, x_1, \ldots, x_{n-1})$ and Bob receives $r_B = (y_0, y_1, \ldots, y_{n-1})$ such that $x_0, \ldots, x_{n-1}, y_0, \ldots, y_{n-1}$ are randomly selected from $\mathbb{F}$, where $x_0 + y_0 = \sum_{i=1}^{n-1} x_i \cdot y_i$.

Figure 3: A quick summary of the definitions of a few correlations that are relevant to this paper.

IPS08, MP06, MPW07, PW08]. Recently, Ishai et al. [IMSW14] constructed OT combiners with nearly optimal leakage parameters. However, combiners consider a restricted variant of leakage where the leakage function leaks only individual bits of the secret shares.

To address general leakage, Ishai, Kushilevitz, Ostrovsky, and Sahai [IKOS09], proposed the notion of correlation extractors. Their construction has a linear leakage resilience, production rate, and exponential security. However, as indicated by Gupta et al. [GIMS15], all the constants involved are minuscule. To address this concern, they [GIMS15] construct correlation extractor for $\mathsf{ROT}^{n/2}$ that has optimal leakage resilience with only a negligible (not exponentially-low) simulation error. They also provide a correlation extractor construction from a large correlation that exhibits $1/2$ leakage resilience but outputs only one OT. Our work will achieve (roughly) the best of both these constructions, i.e., fractional resilience $1/2$, (near) linear production rate, and exponential security.

1.4 Technical Overview

In this section we present a brief overview of our correlation extractor construction and the graph-theoretic measure of the maximum resilience of an arbitrary correlation.

1.4.1 Correlation Extractor Construction

Suppose we are given $0 < \delta < g < 1/2$, and parties are in the $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)^{[t]}$-hybrid, where $t = (1/2 - g)n$ and $\mathbb{K} = \mathbb{GF}\left[2^{\delta n}\right]$. For $m = (\delta n)^{1-o(1)}$, we want to implement the $\mathsf{OLE}\left(\mathbb{GF}\left[2\right]\right)^m$ functionality. Figure 4 presents the outline of our correlation extractor construction. The extraction protocol π is similar to the correlation extractor of Gupta et al. [GIMS15]. Except that, in their case the inner-product correlation was over $\mathbb{GF}\left[2\right]$ instead of a large field $\mathbb{K}$. The security of the protocol is argued in Section 3. Our correlation extractor securely computes a sample from the $\mathsf{ROLE}\left(\mathbb{K}\right)$ correlation. The protocol ρ is the standard protocol that implements the $\mathsf{OLE}\left(\mathbb{K}\right)$ functionality in the $\mathsf{ROLE}\left(\mathbb{K}\right)$-hybrid with perfect security. So, all that remains is to simultaneously embed $\mathsf{OLE}\left(\mathbb{GF}\left[2\right]\right)^m$ into one $\mathsf{OLE}\left(\mathbb{K}\right)$. This embedding relies on finding solutions to a combinatorial problem that is summarized in Figure 5. Section 4 outlines the technique of choosing the inputs to the $\mathsf{OLE}\left(\mathbb{K}\right)$ functionality so that the parties can implement the $\mathsf{OLE}\left(\mathbb{GF}\left[2\right]\right)^m$ functionality with perfect security.

1.4.2 Hardness of Computation Result

The starting point of this result is the observation that we know the exact characterization of the correlations that *do not* suffice to construct OT asymptotically [Kil88, IL89, Kus89, Bea89, MPR09, KMQR09], namely *simple correlations*. Constructing one OT given a single sample from a simple

Ensure. Let $\mathbb{F} = \mathbb{GF}[2]$ and $\mathbb{K} = \mathbb{GF}\left[2^{\delta n}\right]$ be an extension field of $\mathbb{F}$. Let $0 < \delta < g < 1/2$.

Private Input. Let $m = (\delta n)^{1-o(1)}$. Alice has private input $(a_0, \ldots, a_{m-1}) \in \mathbb{F}^m$ and $(b_0, \ldots, b_{m-1}) \in \mathbb{F}^m$. Bob has private input $(x_0, \ldots, x_{m-1}) \in \mathbb{F}^m$.

Hybrid. Parties are in the $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)^{[t]}$-hybrid, where $t = (1/2 - g)n$.

Protocol.

1. Let $\pi(\mathbb{K}, 1/\delta - 1)$ be a protocol in the $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)^{[t]}$-hybrid that securely computes $\mathsf{ROLE}(\mathbb{K})$ with simulation error $2^{-(g-\delta)n/2-1}$. Figure 7 provides the details of the protocol in Section 3.
2. Let $\rho(\mathbb{K}, A^*, B^*, X^*)$ be a perfectly secure protocol for $\mathsf{OLE}(\mathbb{K})$ in the $\mathsf{ROLE}(\mathbb{K})$-hybrid. The private input of Alice is $(A^*, B^*) \in \mathbb{K}^2$ and the private input of Bob is $X^* \in \mathbb{K}$. Bob obtains the output $Z^* = A^*X^* + B^*$. Figure 8 provides the details of the protocol in Section 3.
3. The protocol $\sigma(\mathbb{K}, 1/\delta - 1, A^*, B^*, X^*)$ is the parallel composition of $\pi(\mathbb{K}, 1/\delta - 1)$ and $\sigma(\mathbb{K}, 1/\delta - 1, A^*, B^*, X^*)$ protocol.
4. Parties run the two-round protocol $\sigma(\mathbb{K}, 1/\delta - 1, A^*, B^*, X^*)$ with Alice's private input (A^*, B^*) and Bob's private input X^*. Lemma 3 in Section 4 explains the choice of the inputs A^*, B^*, and X^*.

Output Computation. Lemma 3 in Section 4 presents Bob's algorithm to compute $(z_0, \ldots, z_{m-1})$ from Z^*.

Figure 4: For $0 < \delta < g < 1/2$, the outline of the (n, m, t, ε)-correlation extractor in the $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)^{[t]}$-hybrid, where $m = (\delta n)^{1-o(1)}$, $t = (1/2 - g)n$, $\varepsilon = 2^{-(g-\delta)n/2-1}$.

Our Combinatorial Problem. Find S and T such that

- S and T are ordered sets of non-negative integers of equal size.
- Interpret the set $S + T$ as a matrix, where the (i, j)-th entry represents the sum of the i-th entry in S and the j-th entry in T. All entries in $S + T$ are in the range $[0, n)$, and $(S + T)_{i,i}$ is not equal to any other element in $S + T$, for $i \in \{0, \ldots, |S| - 1\}$.
- Size of $|S| = |T|$ is maximum

Figure 5: Our combinatorial problem for embedding multiple OLE over small fields into one OLE over an extension field.

correlation is even more restrictive, and, hence, the hardness of computation result carries over.[2] This result holds true even when there is no leakage on (R_A, R_B). In fact, there exists a universal constant $\varepsilon^* > 0$ such that any OT protocol using any simple correlation has simulation error at least ε^*.

Intuitively, the simple partition number of a correlation (R_A, R_B), represented by $\mathsf{sp}(R_A, R_B)$, is the minimum Λ such that (R_A, R_B) can be "decomposed into a union of" Λ simple correlations. Section 5 formalizes this notion of decomposition. Next, we prove in Lemma 4 that for any correlation (R_A, R_B), in the presence of $t = \log \mathsf{sp}(R_A, R_B)$ bits of leakage, any protocol π for OT has simulation error at least ε^*. Using this result, we translate tight upper bounds on the simple partition number of relevant correlations into corresponding meaningful upper bounds on their maximum resilience. Figure 2 summarizes our results. We construct a smoother version of this technical lemma using averaging arguments, see Corollary 3. For example, if the leakage bound $t \geqslant (\log \mathsf{sp}(G)) - gn$, then any $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B) has $\varepsilon \geqslant \varepsilon^* \cdot 2^{-gn}$.

2 Preliminaries

We represent the set $\{1, \ldots, n\}$ by $[n]$. For a vector $(x_1, \ldots, x_n)$ and $S = \{i_1, \ldots, i_{|S|}\} \subseteq [n]$, the set x_S represents $(x_{i_1}, \ldots, x_{i_{|S|}})$.

2.1 Functionalities and Correlations

We introduce some useful functionalities and correlations.

Oblivious Transfer. Oblivious transfer, represented by OT, is a two-party functionality that takes as input $(x_0, x_1) \in \{0,1\}^2$ from Alice and $b \in \{0,1\}$ from Bob and outputs x_b to Bob.

Oblivious Linear-function Evaluation. For a field $(\mathbb{F}, +, \cdot)$, oblivious linear-function evaluation over $\mathbb{F}$, represented by $\mathsf{OLE}(\mathbb{F})$, is a two-party functionality that takes as input $(a, b) \in \mathbb{F}^2$ from Alice and $x \in \mathbb{F}$ from Bob and outputs $z = ax + b$ to Bob. In particular, OLE refers to the $\mathsf{OLE}(\mathbb{GF}[2])$ functionality. Note that OT is identical (functionally equivalent) to OLE because $x_b = (x_1 - x_0)b + x_0$.

Random Oblivious Transfer Correlation. Random oblivious transfer, represented by ROT, is a correlation that samples x_0, x_1, b uniformly and independently at random. It provides Alice the secret share $r_A = (x_0, x_1)$ and provides Bob the secret share $r_B = (b, x_b)$.

Random Oblivious Linear-function Evaluation. For a field $(\mathbb{F}, +, \cdot)$, random oblivious linear-function evaluation over $\mathbb{F}$, represented by $\mathsf{ROLE}(\mathbb{F})$, is a correlation that samples $a, b, x \in \mathbb{F}$ uniformly and independently at random. It provides Alice the secret share $r_A = (a, b)$ and provides Bob the secret share $r_B = (x, z)$, where $z = ax + b$. In particular, ROLE refers to the $\mathsf{ROLE}(\mathbb{GF}[2])$ correlation. Note that ROT and ROLE are identical (functionally equivalent) correlations.

Inner-product Correlation. For a field $(\mathbb{F}, +, \cdot)$ and $n \in \mathbb{N}$, inner-product correlation over $\mathbb{F}$ of size n, represented by $\mathsf{IP}(\mathbb{F}^n)$, is a correlation that samples random $r_A = (x_0, \ldots, x_{n-1}) \in \mathbb{F}^n$ and $r_B = (y_0, \ldots, y_{n-1}) \in \mathbb{F}^n$ subject to the constraint that $x_0 + y_0 = \sum_{i=1}^{n-1} x_i y_i$. The secret shares of Alice and Bob are, respectively, r_A and r_B.

For $m \in \mathbb{N}$, the functionality $\mathcal{F}^m$ represents the functionality that implements m independent copies of any functionality/correlation $\mathcal{F}$.

[2] The problem of characterizing correlations whose single sample suffice to construct OT is a fascinating open problem that lies beyond the purview of this study.

2.2 Toeplitz Matrix Distribution

Given a field $\mathbb{F}$, the distribution $\mathbb{T}_{(k,n)}$ represents a uniform distribution over all matrices of the form $[I_{k\times k}|P_{k\times n-k}]$, where $I_{k\times k}$ is the identity matrix and $P_{k\times n-k}$ is a Toeplitz matrix with each entry in $\mathbb{F}$. The distribution $\mathbb{T}_{\perp,(k,n)}$ is the uniform distribution over all matrices of the form $[P_{n-k\times k}|I_{n-k\times n-k}]$, where $I_{n-k\times n-k}$ is the identity matrix and $P_{n-k\times k}$ is a Toeplitz matrix with each entry in $\mathbb{F}$.

2.3 Graph Representation of Correlations

We introduce a graph-theoretic representation of correlations for a more intuitive presentation.

Definition 2 (Graph of a Correlation)**.** *Let (R_A, R_B) be the joint distribution for a correlation. The* graph of the correlation (R_A, R_B) *is the weighted bipartite graph $G = (L, R, E)$ defined as follows.*

1. *The left partite set L is the set of all possible secret shares r_A for Alice,*
2. *The right partite set R is the set of all possible secret shares r_B for Bob, and*
3. *The weight connecting the vertices r_A and r_B is the probability of sampling the shares (r_A, r_B) according to the distribution (R_A, R_B).*

In this paper, the notation (R_A, R_B) also represents the bipartite graph corresponding to it. If the correlation is a uniform distribution over a subset E of all possible edges, then we normalize the entire graph such that the weights on each edge is 1. For example, consider the correlations presented in Figure 3. Henceforth, for the ease of presentation, we assume that the graph of a correlation is an unweighted bipartite graph. The left-most graph in Figure 12 is the graph of the ROLE correlation.

A bipartite graph $G = (L, R, E)$ is a *biclique* if there exists $L' \subseteq L$ and $R' \subseteq R$ such that that edge-set $E(G) = L' \times R'$.

Definition 3 (Simple Graph)**.** *A* simple graph *is a bipartite graph such that each of its connected components is a biclique.*

For example, consider the graph in Figure 6.[3] A *simple correlation* is a correlation whose graph is simple.

Definition 4 (Simple Partition Number)**.** *The* simple partition number *of a graph G, represented by $\mathsf{sp}(G)$, is the minimum number of simple graphs needed to partition its edges.*

Figure 12 and Figure 13 show that the simple partition number for both $\mathsf{ROLE}(\mathbb{GF}[2])$ and $\mathsf{ROLE}(\mathbb{GF}[2])^2$ is 2.

In this work, we use the tensor product of bipartite graphs defined as follows.

Definition 5 (Tensor Product Graph)**.** *For bipartite graphs $G = (L_G, R_G, E_G)$ and $H = (L_H, R_H, E_H)$ the* tensor product of G and H *is the bipartite graph $J = (L_J, R_J, E_J)$ defined as follows.*

1. *The left partite set $L_J := L_G \times L_H$, the right partite set $R_J := R_G \times R_H$, and*
2. *The vertices $(u, v) \in L_J$ and $(u', v') \in R_J$ are connected if $(u, u') \in E_G$ and $(v, v') \in E_H$.*

Applying this definition recursively, we define $G^m := \overbrace{G \times \cdots \times G}^{m-\text{times}}$.

[3] This definition naturally generalizes to weighted graphs. Suppose $p(r_A, r_B)$ represents the probability of jointly sampling (r_A, r_B) from the correlation (R_A, R_B). Then a simple graph has $p(r_A, r_B) = p(r_A) \cdot p(r_B)$, for every (r_A, r_B) edge with positive weight.

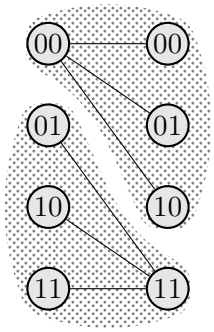

Figure 6: A representative example of a simple graph.

3 Extracting One OLE over a Large Field

In this section we will build some of the building blocks needed to construct the correlation extractor claimed in Theorem 1. In particular, we outline the extraction protocol that, given a leaky $\mathsf{IP}\left(\mathbb{K}^{\eta+1}\right)^{[t]}$ correlation, realizes a secure $\mathsf{OLE}\left(\mathbb{K}\right)$ functionality.

1. First, given the $\mathsf{IP}\left(\mathbb{K}^{\eta+1}\right)$ correlation where parties can perform t-bits of arbitrary leakage, we construct a secure sample of an $\mathsf{ROLE}\left(\mathbb{K}\right)$ correlation. This protocol $\pi(\mathbb{K}, \eta)$ is presented in Figure 7. At the end of the protocol Alice has $(\widetilde{A}_0, \widetilde{B}_0) \in \mathbb{K}^2$ and Bob has $(\widetilde{X}_0, \widetilde{Z}_0) \in \mathbb{K}^2$, such that $\widetilde{A}_0, \widetilde{B}_0, \widetilde{X}_0$ are uniformly random elements in $\mathbb{K}$ and $\widetilde{Z}_0 = \widetilde{A}_0\widetilde{X}_0 + \widetilde{B}_0$. The simulation error of this protocol is $\frac{1}{2}\sqrt{\frac{|\mathbb{K}|2^t}{|\mathbb{K}|^{\eta/2}}}$, refer to Lemma 2.

2. Next, starting with the private shares $(\widetilde{A}_0, \widetilde{B}_0)$ with Alice and $(\widetilde{X}_0, \widetilde{Z}_0)$ with Bob, we implement a protocol $\rho(\mathbb{K}, A^*, B^*, X^*)$. Alice has private inputs (A^*, B^*) that are arbitrary elements in $\mathbb{K}^2$. Bob has private input X^* that is an arbitrary element in $\mathbb{K}$. The protocol $\rho(\mathbb{K}, A^*, B^*, X^*)$, described in Figure 8 is a perfectly secure protocol where Bob outputs $Z^* = A^*X^* + B^*$.

We emphasize that both $\pi(\mathbb{K}, \eta)$ and $\rho(\mathbb{K}, A^*, B^*, X^*)$ are 2-round protocols and we can compose these two protocols in parallel. The resultant protocol $\sigma(\mathbb{K}, \eta, A^*, B^*, X^*)$ is an extraction protocol that takes as input a leaky $\mathsf{IP}\left(\mathbb{K}^{\eta+1}\right)^{[t]}$ correlation where parties can perform t-bits of arbitrary leakage and implements the $\mathsf{ROLE}\left(\mathbb{K}\right)$ functionality with simulation error $\frac{1}{2}\sqrt{\frac{|\mathbb{K}|2^t}{|\mathbb{K}|^{\eta/2}}}$. This is formalized in the following lemma and the proof is included below.

Lemma 1 (Security of Correlation Extractor)**.** *The protocol $\sigma(\mathbb{K}, \eta, A^*, B^*, X^*)$ obtained by the parallel composition of the protocols $\pi(\mathbb{K}, \eta)$ (see Figure 7) and $\rho(\mathbb{K}, A^*, B^*, X^*)$ (see Figure 8) is a secure protocol in the $\mathsf{IP}\left(\mathbb{K}^{\eta+1}\right)^{[t]}$ hybrid that implements the $\mathsf{OLE}\left(\mathbb{K}\right)$ functionality with simulation error at most $\frac{1}{2}\sqrt{\frac{|\mathbb{K}|2^t}{|\mathbb{K}|^{\eta/2}}}$.*

Section 4 elaborates the exact technique to choose appropriate $\mathbb{K}, \eta, A^*, B^*, X^*$ to imply Theorem 1.

3.1 Extraction of one secure $\mathsf{ROLE}(\mathbb{K})$ correlation

The protocol is provided in Figure 7. The security of the protocol is analogous to the proof in [GIMS15] that reduces to the unpredictability lemma over fields. We state this lemma in our context.

Lemma 2 (Unpredictability Lemma)**.** *Let $\mathcal{G} \in \left\{\mathbb{T}_{(k,\eta+1)}, \mathbb{T}_{\perp,(k,\eta+1)}\right\}$. Consider the following game between an honest challenger and an adversary:*

1. *$\mathcal{H}$ samples $m_{[\eta]} \sim U_{\mathbb{K}^\eta}$.*
2. *$\mathcal{A}$ sends a leakage function $\mathcal{L}\colon \mathbb{K}^\eta \to \{0,1\}^t$.*
3. *$\mathcal{H}$ sends $\mathcal{L}\left(m_{[\eta]}\right)$ to $\mathcal{A}$.*
4. *$\mathcal{H}$ samples $x_{[k]} \sim U_{\mathbb{K}^k}$, $G \sim \mathcal{G}$, and computes $y_{\{0\}\cup[n]} = x \cdot G + (0, m_{[\eta]})$. $\mathcal{H}$ sends $(y_{[\eta]}, G)$ to $\mathcal{A}$. $\mathcal{H}$ picks $b \xleftarrow{\$} \{0,1\}$. If $b = 0$, then she sends $\mathsf{chal} = y_0$ to $\mathcal{A}$; otherwise (if $b = 1$) then she sends $\mathsf{chal} = u \sim U_{\mathbb{K}}$ to $\mathcal{A}$.*
5. *$\mathcal{A}$ replies with an element $\tilde{b} \in \{0,1\}$.*

The adversary $\mathcal{A}$ wins the game if $b = \tilde{b}$. For any $\mathcal{A}$, the advantage of the adversary is $\leqslant \frac{1}{4}\sqrt{\frac{|\mathbb{K}|2^t}{|\mathbb{K}|^k}}$.

Similar to the security proof provided by Gupta et al. [GIMS15], the simulation error of the protocol in Figure 7 is the bound provided by the unpredictability lemma over fields (Lemma 2). Refer to Appendix A for a proof of correctness.

3.2 Securely Realizing $\mathsf{OLE}(\mathbb{K})$ using $\mathsf{ROLE}(\mathbb{K})$ Correlation

The protocol presented in Figure 8 is a perfectly semi-honest secure protocol for $\mathsf{OLE}(\mathbb{K})$ in the $\mathsf{ROLE}(\mathbb{K})$ correlation hybrid. Note that the protocols $\pi(\mathbb{K}, \eta)$ in Figure 7 and $\rho(\mathbb{K})$ in Figure 8 can be composed in parallel. Let $\sigma(\mathbb{K}, \eta, A^*, B^*, X^*)$ be the parallel composition of the protocols $\pi(\mathbb{K}, \eta)$ and $\rho(\mathbb{K}, A^*, B^*, X^*)$. This completes the proof of Lemma 1.

4 Embedding multiple OLEs into an OLE over an Extension Field

One of the primary goals in this section is to prove the following lemma.

Lemma 3 (Embedding Multiple small OLE into a Large OLE)**.** *Let $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree n. There exists a perfectly secure protocol for $\mathsf{OLE}(\mathbb{F})^m$ in the $\mathsf{OLE}(\mathbb{K})$-hybrid that makes only one call to the $\mathsf{OLE}(\mathbb{K})$ functionality and $m = n^{1-o(1)}$.*

Proof. Section 4.3 provides this lemma and proves Theorem 1.

4.1 Intuition of the Embedding

We illustrate the main underlying ideas of this embedding problem and our proposed solution using the representative field $\mathbb{F} = \mathbb{GF}[2]$ and its extension field $\mathbb{K} = \mathbb{GF}[2^n]$. Suppose we are provided with an oracle that takes as input $A^*, B^* \in \mathbb{K}$ from Alice and $X^* \in \mathbb{K}$ from Bob, and outputs $Z^* := A^* \cdot X^* + B^*$ to Bob. Our aim is to implement the following functionality. Alice has inputs

Pseudocode of the extraction protocol $\pi(\mathbb{K}, \eta)$.

Given. Alice has $(X_0, X_1, \ldots, X_\eta)$ and Bob has $(Y_0, Y_1, \ldots, Y_\eta)$ such that $X_0 + Y_0 = \sum_{i=1}^{\eta} X_i Y_i$, where $X_0, \ldots, X_\eta, Y_0, \ldots, Y_\eta \in \mathbb{K}$. For ease of presentation assume that η is odd and set $w = (\eta+1)/2$. An adversarial party can obtain arbitrary t-bit leakage on the share of the other party.

Interactive Protocol.

1. **First Round.** Bob sample a random generator matrix G from the distribution $\mathbb{T}_{w\times(\eta+1)}$ such that its elements are in $\mathbb{K}$. Let $\mathcal{C}$ be the code generated by G, and $\mathcal{C}^\perp$ be its dual code. Let H be the generator matrix for the code $\mathcal{C}^\perp$. If the first column of H is $0^{\eta+1-w}$ (i.e., all zeros), then abort the protocol. Bob picks a random codeword $(\widetilde{X}_0, \widetilde{X}_1, \ldots, \widetilde{X}_\eta) \in \mathcal{C}^\perp$ and calculates $M_{[\eta]} = Y_{[\eta]} - \widetilde{X}_{[\eta]}$.

   Bob sends $M_{[\eta]}$ and G to Alice.

2. **Second Round.** Alice samples a random codeword $(\widetilde{A}_0, \widetilde{A}_1, \ldots, \widetilde{A}_\eta) \in \mathcal{C}$ and a random field element $\widetilde{B}_0 \in \mathbb{K}$. Alice computes $\alpha_{[\eta]} = X_{[\eta]} + \widetilde{A}_{[\eta]}$ and $\beta = \langle X_{[\eta]}, M_{[\eta]} \rangle - \widetilde{B}_0 - X_0$.

   Alice sends $\alpha_{[\eta]}$ and β to Bob.

Output Computation. Alice outputs $(\widetilde{A}_0, \widetilde{B}_0)$ and Bob outputs $(\widetilde{X}_0, \widetilde{Z}_0)$, where $\widetilde{Z}_0 = -\left\langle \alpha_{[\eta]}, \widetilde{X}_{[\eta]} \right\rangle - \beta + Y_0$.

Figure 7: Protocol to securely extract one random sample of the $\mathsf{ROLE}(\mathbb{K})$ functionality from the leaky $\mathsf{IP}\left(\mathbb{K}^{\eta+1}\right)^{[t]}$ correlation.

Pseudocode of the OLE protocol $\rho(\mathbb{K}, A^*, B^*, X^*)$

Given. Alice has $(\widetilde{A}_0, \widetilde{B}_0)$ and Bob has $(\widetilde{X}_0, \widetilde{Z}_0)$, where $\widetilde{A}_0, \widetilde{B}_0, \widetilde{X}_0$ are random elements in $\mathbb{K}$ and $Z_0 = A_0 X_0 + B_0$.

Private Inputs. Alice has private input $(A^*, B^*) \in \mathbb{K}^2$ and Bob has $X^* \in \mathbb{K}$.

Interactive Protocol.

1. **First Round.** Bob sends $M' = \widetilde{X}_0 - X^*$ to Alice.
2. **Second Round.** Alice sends $\alpha' = \widetilde{A}_0 + A^*$ and $\beta' = \widetilde{A}_0 M + B^* + \widetilde{B}_0$.

Output Computation. Bob outputs $Z^* = \alpha' X^* + \beta' - \widetilde{Z}_0$.

Figure 8: Perfectly secure protocol to realize $\mathsf{OLE}(\mathbb{K})$ in the $\mathsf{ROLE}(\mathbb{K})$ correlation hybrid.

$(a_0, \ldots, a_{m-1}) \in \mathbb{F}^m$ and $(b_0, \ldots, b_{m-1}) \in \mathbb{F}^m$, and Bob has inputs $(x_0, \ldots, x_{m-1}) \in \mathbb{F}^m$. We want Bob to obtain $(z_0, \ldots, z_{m-1}) \in \mathbb{F}^m$, where each $z_i = a_i \cdot x_i + b_i$, for $i \in \{0, \ldots, m-1\}$. Intuitively, we want maximize m and embed $\mathsf{OLE}(\mathbb{F})^m$ into one $\mathsf{OLE}(\mathbb{K})$.

Preliminary Idea. Consider the following simple preliminary embedding. Let $m = \sqrt{n}$. Alice defines $A^* = a_0 + a_1\zeta + \cdots + a_{m-1}\zeta^{m-1}$, where $a_0, \ldots, a_{m-1} \in \mathbb{F}$. And, Alice defines $B^* = \sum_{i=0}^{n-1} r_i \zeta^i$, where each r_i is a random element in $\mathbb{F}$; except when $(m+1)$ divides i, then we set $r_{t(m+1)} = b_t$, for $t \in \{0, \ldots, m-1\}$. Bob defines $X^* = x_0 + x_1\zeta^m + \cdots + x_{m-1}\zeta^{(m-1)m}$, where $x_0, \ldots, x_{m-1} \in \mathbb{F}$.

Now, the parties compute $Z^* = A^*X^* + B^*$ using one oracle call to $\mathsf{OLE}(\mathbb{K})$ and Bob obtains the output Z^*. Note that the intended $z_i = a_i \cdot x_i + b_i$ is the coefficient of $\zeta^{i(m+1)}$ in Z^*, for each $i \in \{0, \ldots, m-1\}$. Coefficients of all other powers of ζ contain no information about $a_0, \ldots, a_{m-1}, b_0, \ldots, b_{m-1}$, because they are masked with random elements in $\mathbb{F}$. So, for $m = \sqrt{n}$, we have embedded $\mathsf{OLE}(\mathbb{F})^m$ into one $\mathsf{OLE}(\mathbb{K})$.

Better Embedding. Observe that $(a_0 + a_1\zeta) \cdot (x_0 + x_1\zeta) = a_0x_0 + (a_0x_1 + a_1x_0)\zeta + a_1x_1\zeta^2$. So, we can embed $\mathsf{OLE}(\mathbb{F})^2$ into one $\mathsf{OLE}(\mathbb{K})$, where $\mathbb{K}$ is an extension field of $\mathbb{F}$ of degree 3, as follows. Alice chooses $A^* = a_0 + a_1\zeta \in \mathbb{GF}[2^2]$ and $B^* = b_0 + r\zeta + b_1\zeta^2$ (where r is a random element from $\mathbb{F}$), and Bob chooses $X^* = x_0 + x_1\zeta$. Note that the coefficients of ζ^0 and ζ^2 in Z^*, respectively, correspond to $a_0x_0 + b_0$ and $a_1x_1 + b_1$. Recursively applying this idea, we can construct an embedding of $\mathsf{OLE}(\mathbb{GF}[2])^{2^k}$ into one $\mathsf{OLE}\left(\mathbb{GF}\left[2^{3^k}\right]\right)$. Asymptotically, this scheme embeds $m = n^{\log 2/\log 3} \approx n^{0.63}$ copies of $\mathsf{OLE}(\mathbb{GF}[2])$ into one $\mathsf{OLE}(\mathbb{GF}[2^n])$.

Generalization to 3-free sets. Consider the previous solution when $n = 3^k$. Let $S = \{s_0 < s_1 < \cdots < s_{m-1}\}$ be the set of indices. The set S corresponding to the previous solution contains all integers less than 3^k whose ternary representation does not contain the digit 2. This is the famous greedy sequence of integers that does not include an arithmetic progression of length 3; namely, 3-free sets. In fact, there is nothing sacrosanct about the S chosen in the previous embedding, and *any* 3-free set suffices.

For example, let $S = \{s_0 < s_1 < \cdots < s_{m-1}\}$ be any 3-free set such that each entry is in the range $[0, n/2)$, $\mathbb{F} = \mathbb{GF}[2]$, and $\mathbb{K} = \mathbb{GF}[2^n]$. Alice prepares $A^* = \sum_{i=0}^{m-1} a_i\zeta^{s_i}$ and $B^* = \sum_{k=0}^{n-1} r_k\zeta^k$, where $r_{2s_i} = b_i$; otherwise it is a random element in $\mathbb{F}$. Bob prepares $X^* = \sum_{i=0}^{m-1} x_i\zeta^{s_i}$. Using one call to $\mathsf{OLE}(\mathbb{K})$ Bob obtains Z^*. The coefficient of ζ^{2s_i} is $a_ix_i + b_i$, because no other $s_j + s_k = 2s_i$. Now, we can embed $m = n^{1-o(1)}$ copies of $\mathsf{OLE}(\mathbb{F})$ into $\mathsf{OLE}(\mathbb{K})$ using the state-of-the-art constructions of 3-free sets [Beh46, Elk10]. However, this approach cannot give us $m = \Theta(n)$ due to sub-linear upper bounds on m [Rot53, HB87, Sze90, Bou99, Bou08, San11].

New Problem. Note that although solutions to the 3-free set problem imply embeddings in our setting, our embedding problem is potentially less restrictive. For example, the solution for $m = \sqrt{n}$ presented above is not obtained by the reduction to 3-free sets. Are we missing something?

Suppose $S = (s_0, \ldots, s_{m-1})$ and $T = (t_0, \ldots, t_{m-1})$ be tuples of indices in the range $[0, n/2)$. Consider the combinatorial problem proposed in Figure 5.

Given S and T that are solutions to the problem in Figure 5, Alice and Bob use the strategy explained in Figure 9. Note that the initial solution for $m = \sqrt{n}$ indeed corresponds to the solution $S = \{0, \ldots, m-1\}$ and $T = \{0, m, \ldots, (m-1)m\}$. Restricted to $S = T$, our combinatorial problem is identical to the 3-free set problem. We numerically solve this problem for small values of n and, indeed, it produces more efficient embeddings than the embedding based on the optimal 3-free set

Given. Two sets S and T of size m that is a solution to the combinatorial problem presented in Figure 5. Let $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree n.

Private input. Alice has private input $(a_0, \ldots, a_{m-1}) \in \mathbb{F}^m$ and $(b_0, \ldots, b_{m-1}) \in \mathbb{F}^m$. Bob has private input $(x_0, \ldots, x_{m-1}) \in \mathbb{F}^m$.

Hybrid. Parties are in the $\mathsf{OLE}(\mathbb{K})$-hybrid.

Private Input Construction.

1. Alice creates private input $A^* = \sum_{i=0}^{m-1} a_i \zeta^{s_i} \in \mathbb{K}$.
2. Alice chooses r_i, for $i \in \{0, \ldots, n-1\}$, as follows.

$$r_i = \begin{cases} b_k & \text{, if } i = s_k + t_k \text{ for some } k \in \{0, \ldots, m-1\} \\ U_{\mathbb{F}} & \text{, otherwise.} \end{cases}$$

   Alice creates private input $B^* = \sum_{i=0}^{n-1} r_i \zeta^i \in \mathbb{K}$.
3. Bob creates private input $X^* = \sum_{i=0}^{m-1} x_i \zeta^{t_i} \in \mathbb{K}$.
4. Both parties invoke the $\mathsf{OLE}(\mathbb{K})$ functionality with respective Alice input (A^*, B^*) and Bob input X^*. Bob receives $Z^* = A^* X^* + B^*$.

Output Decoding. Bob outputs $(z_0, \ldots, z_{m-1})$, where z_i is the coefficient of $\zeta^{s_i + t_i}$ and $i \in \{0, \ldots, m-1\}$.

Figure 9: Embedding $\mathsf{OLE}(\mathbb{F})^m$ into one $\mathsf{OLE}(\mathbb{K})$, where $\mathbb{K}$ is an extension field of $\mathbb{F}$ of degree n.

constructions. We emphasize that we compare our solutions against the largest 3-free set computed by *exhaustive search*. We summarize our observations in Figure 10.

4.2 Relevant Prior Work on 3-free Sets

Our asymptotic construction for Theorem 1 relies on constructing a dense subset S of $\{0, 1, \cdots, n-1\}$ that does not contain any arithmetic progression, namely 3-free sets. Erdős and Turán introduced this problem in 1936 and presented a greedy construction with $|S| = \Omega\left(n^{\log 2/\log 3}\right) \approx n^{0,63}$. Salem and Spencer [SS42] showed that the surface of high-dimensional convex bodies can be embedded in the integers to construct 3-free sets of size $n^{1-o(1)}$. Later, Behrend [Beh46] noticed that points lying on the surface of a sphere of suitable radius are a particularly good choice, and gave a construction with $|S| = \Omega\left(\frac{n}{2^{2\sqrt{2\log n}} \cdot \log^{1/4} n}\right)$. Recently, after a gap of over sixty years, Elkin [Elk10] improved this further by a factor of $\Theta(\sqrt{\log n})$ by thickening the spheres to produce the largest known 3-free set. The proofs of Behrend [Beh46] and Elkin [Elk10] are constructive in nature and the sets can be constructed in poly(n) time. Although the greedy construction is asymptotically worse than these two constructions, it performs well for realistic values of n. See Figure 11 for details.

Roth [Rot53] provided the first nontrivial upper bound of $O\left(\frac{n}{\log \log n}\right)$ on the size of 3-free sets. More than thirty years later, Heath-Brown [HB87] showed that $|S| = O\left(\frac{n}{\log^c n}\right)$, for some constant

m	$n(m)$	Solution Sets	$n'(m)$	3-free Set
1	1	$S = \{0\}$ $T = \{0\}$	1	$S = \{0\}$
2	3	$S = \{0, 1\}$ $T = \{0, 1\}$	3	$S = \{0, 1\}$
3	7	$S = \{0, 1, 3\}$ $T = \{0, 1, 2\}$	7	$S = \{0, 1, 3\}$
4	9	$S = \{0, 1, 3, 4\}$ $T = \{0, 1, 2, 4\}$	9	$S = \{0, 1, 3, 4\}$
5	14	$S = \{0, 1, 3, 5, 8\}$ $T = \{0, 1, 4, 5, 3\}$	17	$S = \{0, 1, 3, 7, 8\}$
6	19	$S = \{0, 1, 3, 4, 7, 9\}$ $T = \{0, 1, 3, 9, 7, 8\}$	21	$S = \{0, 1, 3, 4, 9, 10\}$
7	24	$S = \{0, 1, 3, 4, 11, 6, 10\}$ $T = \{0, 1, 5, 10, 6, 12, 9\}$	25	$S = \{0, 1, 3, 4, 9, 10, 12\}$
8	27	$S = \{0, 1, 3, 4, 9, 10, 12, 13\}$ $T = \{0, 1, 3, 4, 9, 10, 12, 13\}$	27	$S = \{0, 1, 3, 4, 9, 10, 12, 13\}$
9	34	$S = \{0, 1, 3, 4, 9, 12, 14, 16, 17\}$ $T = \{0, 1, 3, 4, 13, 11, 12, 15, 16\}$	39	$S = \{0, 1, 5, 6, 8, 13, 14, 17, 19\}$

Figure 10: Let $\mathbb{K}$ be an extension field of $\mathbb{F}$ of degree n. Our goal is to embed m copies of $\mathsf{OLE}(\mathbb{F})$ into one $\mathsf{OLE}(\mathbb{K})$ using minimum n. The number $n(m)$ represents the minimum n obtained by using solutions to our combinatorial problem in Figure 5. The number $n'(m)$ represents the minimum n obtained by using the optimum solutions to the 3-free set problem.

$c > 0$, and then Szemeredi [Sze90] produced an explicit value $c = 1/20$. Bourgain [Bou99, Bou08] improved the upper bound by polylog factors. Sanders [San11] proved the current best known upper bound of $O\left(\frac{n(\log\log n)^5}{\log n}\right)$. Nathan [McN] provides a comprehensive summary for both 3-free set size constructions and upper bounds.

4.3 Generating Explicit Embedding and Proof of Theorem 1

First, we prove Lemma 3. Let $S(n)$ be a 3-free set with elements in the range $[0, n/2)$. Behrend [Beh46] and Elkin [Elk10] provide constructions for $S(n)$ such that $|S(n)| \geqslant n^{1-o(1)}$. Note that $S = T = S(n)$ is a solution to the combinatorial problem proposed in Figure 5. Now, we use the protocol described in Figure 9.

It is clear that the protocol is correct. The coefficients of all other ζ^i in Z^* are random elements in $\mathbb{F}$, if $i \neq s_k + t_k$, for all $k \in \{0, \ldots, m-1\}$. It is, therefore, easy to see that this is a perfectly secure protocol for $\mathsf{OLE}(\mathbb{F})^m$ in the $\mathsf{OLE}(\mathbb{K})$-hybrid. □

Remark. We provide a short discussion on how to pick the 3-free set S for *concrete* values of n. The greedy construction is the fastest and runs in $O(n \log n)$ time. The proofs of Behrend [Beh46] and Elkin [Elk10] are also constructive in nature and the set can be constructed in poly(n) time. However, their performance for realistic values of n are worse than the greedy algorithm.

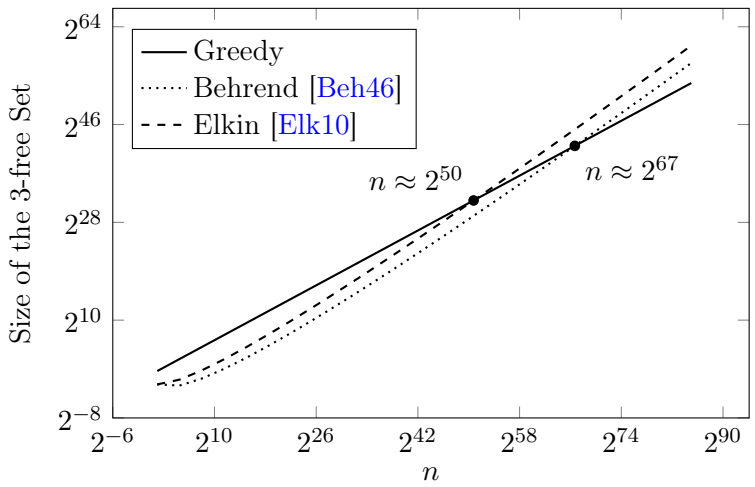

Figure 11: A logarithmic scaled graph of the size of the 3-free sets produced by the greedy, Behrend [Beh46], and Elkin [Elk10] constructions.

4.3.1 Proof of Theorem 1

Suppose we given n, $0 < \delta < g < 1/2$, and $t = (1/2 - g)n$. Let $\mathbb{K} = \mathbb{GF}\left[2^{\delta n}\right]$ and $\mathbb{F} = \mathbb{GF}[2]$. We construct $A^*, B^*, X^* \in \mathbb{K}$ using Lemma 3 and $m \geqslant (\delta n)^{1-o(1)}$. Perform the protocol $\sigma(\mathbb{K}, 1/\delta - 1, A^*, B^*, X^*)$ in the $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)^{[t]}$-hybrid.[4] The simulation error is

$$\varepsilon \leqslant \frac{1}{2}\sqrt{\frac{2^{\delta n}2^{t}}{2^{\delta n(1/\delta-1)/2}}} = 2^{-(g-\delta)n/2-1}$$

This is an (n, m, t, ε)-correlation extractor for the correlation $\mathsf{IP}\left(\mathbb{K}^{1/\delta}\right)$.

5 Simple Partition Number

This section defines the simple partition number of a graph, provides estimates of this quantity for correlations relevant to our work, and proves Theorem 2.

5.1 Intuition of the Hardness of Computation Result

We know that if parties have multiple independent samples of secret shares sampled according to a simple correlation, then the parties cannot securely compute OT [Kil88, IL89, Kus89, Bea89, MPR09, KMQR09]. Constructing one OT given a single sample from such a correlation is even more restrictive, and, hence, the hardness of computation result carries over.[5] This result holds true even when there is no leakage on (R_A, R_B). More precisely, we import the following result that we restate in our context.

Imported Theorem 1 ([MPR09])**.** *Let (R_A, R_B) be a simple correlation with n-bit secret shares for each party. There exists a universal constant $\varepsilon^* > 0$, such that any $(n, 1, 0, \varepsilon)$-correlation extractor for (R_A, R_B) has $\varepsilon \geqslant \varepsilon^*$.*

[4] Recall that in the protocol $\pi(\mathbb{K}, \eta)$, all parties have share size $(\eta + 1)\log|\mathbb{K}|$.

[5] The problem of characterizing correlations whose single sample suffice to construct OT is a fascinating open problem that lies beyond the purview of this study.

Suppose (R_A, R_B) is a correlation that has simple partition number $\mathsf{sp}(G) = 2^\lambda$ and $G = G^{(1)} + \cdots + G^{(2^\lambda)}$, where each $G^{(i)}$ is a simple graph. Then we consider the leakage function $\mathcal{L}(r_A, r_B) = \ell$, where $\ell \in \{1, \ldots, 2^\lambda\}$ is the unique index such that $(r_A, r_B) \in E(G^{(\ell)})$. Note that $\mathcal{L}$ is a λ-bit leakage function and conditioned on the leakage being ℓ, for any $\ell \in \{1, \ldots, 2^\lambda\}$, the correlation $(R_A, R_B|\ell)$ is a simple correlation. So, one of the parties can break the security of any purported OT protocol where parties get secret shares sampled from the $(R_A, R_B|\ell)$ correlation. Overall, with probability half, one of the parties can break the security of any purported OT protocol where parties get secret shares sampled from the (R_A, R_B) by performing the leakage $\mathcal{L}$ described above. This technique upper-bounds the leakage resilience of (R_A, R_B) and we summarize it as follows.

Lemma 4 (Connection between Maximum Leakage Resilience and Simple Partition Number)**.** *Let (R_A, R_B) is a correlated private randomness that provides n-bit private shares to Alice and Bob. Let G be the bipartite graph corresponding to the correlation (R_A, R_B). There exists a universal constant $\varepsilon^* > 0$ such that any $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B) with $t \geqslant \lceil \lg \mathsf{sp}(G) \rceil$ has $\varepsilon \geqslant \varepsilon^*$.*

We construct a smoother version of this technical lemma using averaging arguments. For example, if the leakage bound t is roughly $(\log \mathsf{sp}(G)) - gn$, then we consider a subset of simple graphs of size $\mathsf{sp}(G) \cdot 2^{-gn}$ from the set $\left\{G^{(1)}, \ldots, G^{(\mathsf{sp}(G))}\right\}$ that covers at least 2^{-gn} fraction of the edges of G. Applying the previous lemma, we can conclude that $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B) with $t \geqslant \lceil \log \mathsf{sp}(G) - gn \rceil$ has $\varepsilon \geqslant \varepsilon^* \cdot 2^{-gn}$.

Corollary 3 ((Smooth Version of the) Connection between Maximum Leakage Resilience and Simple Partition Number)**.** *Let (R_A, R_B) is a correlated private randomness that provides n-bit private shares to Alice and Bob. Let G be the bipartite graph corresponding to the correlation (R_A, R_B). There exists a universal constant $\varepsilon^* > 0$ such that any $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B) with $t \geqslant \lceil \lg \mathsf{sp}(G) - gn \rceil$ has $\varepsilon \geqslant \varepsilon^* \cdot 2^{-gn}$.*

5.2 Relevant Prior Work on Graph Covering Problems

The graph-theoretic measure proposed in our work to measure the maximum resilience of correlations in best presented in the framework of graph covering problems. Several problems in graph theory, for example, clique partition number, biparticity, arboricity, edge-chromatic number, vertex cover number and biclique partition number, can be expressed as covering a graph with subgraphs from a family of graphs. Of these representative examples, the concept of *biclique partition number* is most relevant to our paper. For a graph G, its biclique partition number, represented by $\mathsf{bp}(G)$, is the minimum number of bicliques that suffice to partition it.

Refer to [KRW88] for a comprehensive survey on graph covering problems. Motivated by network addressing problem and graph storage problem, Graham and Pollak [GP71, GP72] introduced the biclique partition problem (see also [BF92, VL85, YY06, vLW01]). The celebrated Graham-Pollak Theorem states that $\mathsf{bp}(K_n) = (n-1)$ [GP72, Tve82, Pec84, Vis08, Vis13], but all proofs are algebraic, and no purely combinatorial proof is known. In general, $\mathsf{bp}(G) \geqslant \max\{n_+(G), n_-(G)\}$ [GP72, Hof72, Tve82, Pec84], where $n_+(\cdot)$ and $n_-(\cdot)$, respectively, represents the number of positive and negative eigenvalues of the adjacency matrix of the graph. Determining the $\mathsf{bp}(G)$ of a general graph is a hard problem [KRW88], but it admits a trivial upper bound $\mathsf{bp}(G) \leqslant$ the size of the smallest vertex cover of G. Variants of this quantity have been considered recently by [CT13].

This quantity is closely related to the recently disproved [HS12, CT11] Alon-Saks-Seymour Conjecture [Kah91] that $\mathsf{bp}(G)+1$ colors suffice to color a graph. This conjecture can be interpreted as

a generalization of the Graham-Pollak Theorem and has close relations to computational complexity [HS12, NW95, Raz92]. In the context of this paper, intuitively, the biclique partition number is a combinatorial version of the *Wyner's Common Information* [Wyn75] that corresponds to the minimum description complexity of the information that kills the mutual information of correlations. We interpret a correlation as a weighted bipartite graph with the left-partite set being all possible values of r_A, and the right partite set being all possible values of r_B. The weight on an edge joining r_A and r_B represents the probability of jointly sampling (r_A, r_B). This graph-theoretic interpretation of correlations helps establish connections between combinatorial and information-theoretic concepts.

5.3 Relation to Leakage resilience: Proof of Lemma 4

In this section we prove Lemma 4, i.e. the maximum leakage resilience of a correlation (R_A, R_B) is at most $\lg \mathsf{sp}(R_A, R_B)$.

Let G be the bipartite graph corresponding to the correlation (R_A, R_B). Let π be a $(n, 1, t, \varepsilon)$-correlation extractor for G, where $t = \lceil \log \mathsf{sp}(G) \rceil$. Let $G = G^{(1)} + \cdots + G^{(\mathsf{sp}(G))}$ be the simple partition of G. Define the leakage function $\mathcal{L}\colon E(G) \to \{1, \ldots, \mathsf{sp}(G)\}$ as follows. For $e \in E(G)$, we have $\mathcal{L}(e) = \ell$, where ℓ is the unique index in $\{1, \ldots, \mathsf{sp}(G)\}$ such that $e \in E(G^{(\ell)})$.

Consider an interactive protocol that runs π between Alice and Bob with secret samples drawn from the correlation G, and *both parties* receive the leakage $\mathcal{L}(r_A, r_B)$.

Note that this is identical to the interactive protocol, where the correlation G^+ that samples $\ell \in \{1, \ldots, \mathsf{sp}(G)\}$ with probability proportional to $\left|E(G^{(\ell)})\right|$, samples $(u, v) \equiv e \xleftarrow{\$} E(G^{(\ell)})$, and provides (u, ℓ) to Alice and (v, ℓ) to Bob.

The functionality G^+ itself is simple, because each $G^{(\ell)}$ is simple. So, we can use Imported Theorem 1. Therefore, one of the parties' view cannot be simulated with less than $\varepsilon^* > 0$ simulation error when the parties follow the protocol π. Suppose, that party is Alice, without loss of generality. That is, the view of the party Alice* (to represent the semi-honest adversarial strategy) in the interactive protocol between Alice* and B incurs at least ε^* simulation error.

Now consider the case where only Alice* receives the leakage from the correlation and not Bob. The view of Alice* remains identical to the previous hybrid. Therefore, this protocol also incurs a simulation error at least ε^*

This implies that for any $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B), if $t \geqslant \log \mathsf{sp}(R_A, R_B)$, then $\varepsilon \geqslant \varepsilon^*$.

Intuitively, Lemma 4 can be summarized as follows. A small simple partition number of the correlated private randomness (R_A, R_B) implies a low maximum leakage-resilience of (R_A, R_B).

5.3.1 Proof of Corollary 3.

Suppose $2^t = \mathsf{sp}(G)/2^{gn}$ and π is an $(n, 1, t, \varepsilon)$-correlation extractor for (R_A, R_B). Now, we choose the $\mathsf{sp}(G)/(2^{gn} - 1)$ simple graphs among $\{G^{(1)}, \ldots, G^{(\mathsf{sp}(G))}\}$ that cover a subset $E' \subset E(G)$ such that $|E'|/|E(G)| \geqslant (2^{gn} - 1)^{-1}$. The leakage function $\mathcal{L}(r_A, r_B)$ outputs the index of the simple graph from which the edge $e = (r_A, r_B)$ comes, if $e \in E'$; otherwise, it returns $\perp$. Using the same proof as Lemma 4 we can conclude that the simulation error is $\varepsilon \geqslant \varepsilon^* (2^{gn} - 1)^{-1} \approx \varepsilon^* 2^{-gn}$.

5.4 Estimates of Simple Partition Number and Proof of Theorem 2

In this section we present the lemma that provides the estimates of the simple partition number of relevant correlations.

Lemma 5 (Simple Partition Number Estimates)**.** *The following holds true for arbitrary field* $\mathbb{F}$*.*

1. $\mathsf{sp}\left(\mathsf{IP}\left(\mathbb{F}^n\right)\right) \leqslant |\mathbb{F}|^{\lceil (n+1)/2 \rceil}$*, and*
2. *For even* n*,* $\mathsf{sp}\left(\mathsf{ROLE}\left(\mathbb{F}\right)^{n/2}\right) \leqslant |\mathbb{F}|^{\lceil n/4 \rceil}$*.*

The proof is provided in Appendix C

5.4.1 Proof of Theorem 2.

The proof is a direct application of Lemma 4 and Lemma 5.

5.5 Subsuming the Partition Argument

In this section, using a particular example, we want to illustrate that the simple partition number is sophisticated enough to subsume partition argument based impossibility results. To begin, let us consider an example. Let (R_A, R_B) be the random oblivious linear function evaluation over $\mathbb{GF}[2]$. So, the correlation samples $a, b, x \in \mathbb{GF}[2]$ independently and uniformly at random. The secret share of Alice is $r_A = (a, b)$ and the secret share of Bob is $r_B = (x, z)$, where $z = ax + b$. The secrecy of $\mathsf{ROLE}(\mathbb{GF}[2])$ ensures that Alice has no advantage in guessing x and Bob has no advantage in guessing a. The graph of the correlation is provided in Figure 12. The figure presents the simple decomposition corresponding to the leakage $\ell = x - a$.

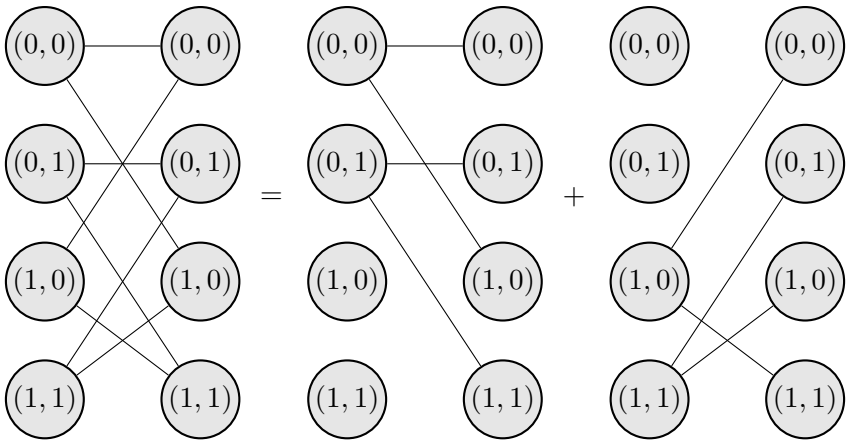

Figure 12: The graph of the correlated private randomness $\mathsf{ROLE}(\mathbb{GF}[2])$ and its decomposition into two simple graphs.

Now, let us consider $\mathsf{ROLE}(\mathbb{GF}[2])^2$, i.e. two independent samples from the $\mathsf{ROLE}(\mathbb{GF}[2])$ correlation. Alice gets secret share (a_1, b_1, a_2, b_2) and Bob gets secret share (x_1, z_1, x_2, z_2), where $z_1 = a_1 x_1 + b_1$ and $z_2 = a_2 x_2 + b_2$. Suppose in the partition argument Alice implements the first correlation and Bob implements the second correlation. This implies that Alice knows x_1 and Bob knows a_2. We want to achieve this effect using only one-bit leakage that is provided to both the parties.

Given the decomposition in Figure 12, note that we can define a two-bit leakage to achieve this. For example the first leakage bit represents $\ell_1 = x_1 - a_1$, and the second leakage bit represents $\ell_2 = x_2 - a_2$. We show in Figure 13 that even a one-bit leakage suffices. In particular, we use $\mathcal{L}(r_A, r_B) = x_1 - a_2$. In general, Lemma 10 shows that $\mathsf{sp}\left(\mathsf{ROLE}(\mathbb{GF}[2])^2\right) \leqslant |\mathbb{F}|$.

Using this observation and the fact that $\mathsf{sp}(G \times H) \leqslant \mathsf{sp}(G) \cdot \mathsf{sp}(H)$ (see Lemma 8), Lemma 5 shows that $\mathsf{sp}(\mathsf{ROLE}(\mathbb{GF}[2])^n) \leqslant |\mathbb{F}|^{\lceil n/2 \rceil}$. This demonstrates that the simple partition number subsumes the partition argument.

5.6 Relevant Prior Work on Common Information and Tension.

We briefly introduce a few relevant information-theoretic measures for maximum resilience and maximum production rate. For a joint distribution, the mutual information $I(R_A; R_B)$ measures the distance (KL-divergence) between the joint probability distribution $p(r_A, r_B)$ and the distribution $p(r_A) \cdot p(r_B)$. The mutual information between (R_A, R_B) represents the number of bits of the secret key that the two parties can agree. The Gács-Körner [GK73] common information, represented by $K(R_A; R_B)$, represents the largest entropy of the common random variable that each party can generate based on their respective secret share. Intuitively, this corresponds to the number of connected components in a bipartite graph representing the correlation. The Wyner common information [Wyn75], represented by $J(R_A; R_B)$, is the minimum information that, when leaked to the eavesdropper, ensures that the parties cannot establish a secret key. This quantity roughly corresponds to the biclique partition number of a bipartite graph for the correlation, where the correlation is a uniform distribution over the edges of the bipartite graph. Prabhakaran and Prabhakaran [PP10, PP11], generalizing [WW05], introduced the concept of *assisted common information* that, among its various applications, helps characterize an upper bound on the number of OTs that a correlation can produce.

5.6.1 Relation to Mutual Information.

In the setting of key-agreement, the mutual information $I(R_A; R_B)$ of a correlation (R_A, R_B) measures the length of the secret key that the two parties can agree on. We emphasize that this is a measure of production, and not a measure of resilience. For example, $I(\mathsf{IP}(\mathbb{GF}[2]^n)) = 1$. Since, secure OT implies one-bit key-agreement, mutual information is also an upper bound on the OT production that a correlation can support. However, production capacity and resilience to leakage are extremely disparate quantities. For example, in the secure computation setting, the correlation $\mathsf{IP}(\mathbb{GF}[2]^n)$ is resilient to $n/2$ bits of leakage but can only produce one OT. Additionally, mutual information significantly overestimates the maximum OT production capacity. For example, n-bit shared private key cannot produce one OT even without any leakage. However, it has n-bits of mutual information.

We emphasize that the simple partition number is only a measure for the maximum leakage resilience of correlations in the setting of secure computation. Our measure *does not* provide any estimates on the OT production. The most relevant measure for OT production is the notion of assisted common information proposed by Prabhakaran and Prabhakaran [PP10, PP11].

5.7 Analogy of Biclique Partition Number and Wyner's Common Information

A correlation that is a biclique has no mutual-information and, hence, is useless for parties to agree on a secret key even asymptotically. In particular, one sample from a correlation that is a biclique is also useless for key-agreement. Suppose (R_A, R_B) is an arbitrary correlation and has biclique partition complexity $\mathsf{bp}(R_A, R_B)$. Similar to Lemma 4, in the presence of $t = \log \mathsf{bp}(R_A, R_B)$ bits of leakage there is not even a one-bit secure key-agreement protocols using (R_A, R_B). The random variable J for the leakage function $\mathcal{L}(R_A, R_B)$ outputs the index of the biclique that contains the edge $e = (r_A, r_B)$.

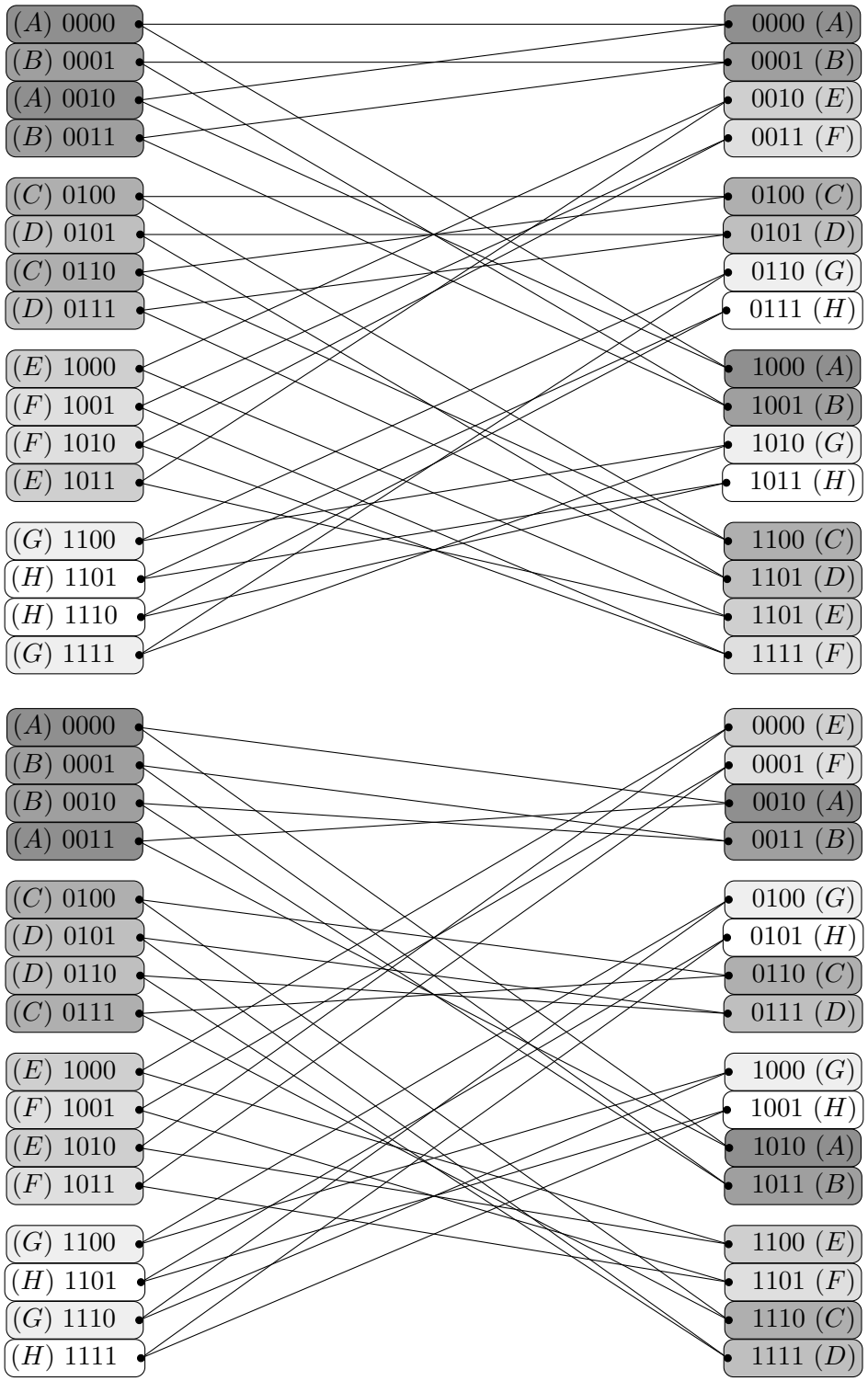

Figure 13: A simple decomposition of $\mathsf{ROLE}\left(\mathbb{GF}\left[2\right]\right)^2$, into two simple graphs. Each collection of nodes with identical shade of gray and letter represents a connected component.

Wyner's common information [Wyn75] is defined to be the minimum entropy random variable J that suffices to ensure $I(R_A; R_B|J) = 0$. If the bicliques that partition G have roughly equal number of edges then these two concepts are identical. Analogously, $\mathsf{sp}\,(R_A, R_B)$ can be interpreted as the analog for Wyner's common information in the secure computation setting.

6 Conclusions and Open Problems

This paper helps progress towards the ideal goal of designing correlations such that there are (n, m, t, ε)-correlation extractors, for $n \approx m + t - \log \varepsilon$. We need to explore new correlations and extractor design principles for this task. A fascinating open question is whether we can break the barrier of 1/2 fractional leakage.

Regarding the technical tools and techniques introduced in this paper, it will be interesting to obtain asymptotic and concrete constructions for our combinatorial problem that are better than the solutions produced by 3-free set constructions. It will also be interesting to upper-bound the maximum achievable solution size for our combinatorial problem. For concrete constructions, linear programming methods and perturbation techniques need to be explored.

We also need algorithmic techniques to explore the hardness, estimate and approximate the simple partition number of general correlations. Analogous to the biclique partition number, it will be helpful to obtain algebraic or combinatorial techniques to lower bounds the simple partition number of a graph. An ideal result in this field, though it is perceived extremely difficult by the authors, will be to demonstrate that the simple partition number also forms a lower-bound to leakage resilience.

References

[BDNP08] Assaf Ben-David, Noam Nisan, and Benny Pinkas. FairplayMP: a system for secure multi-party computation. In Peng Ning, Paul F. Syverson, and Somesh Jha, editors, *ACM CCS 08: 15th Conference on Computer and Communications Security*, pages 257–266, Alexandria, Virginia, USA, October 27–31, 2008. ACM Press. 1

[Bea89] Donald Beaver. Perfect privacy for two-party protocols. In Joan Feigenbaum and Michael Merritt, editors, *Proceedings of DIMACS Workshop on Distributed Computing and Cryptography*, volume 2, pages 65–77. American Mathematical Society, 1989. 1, 6, 16

[Beh46] Felix A Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proceedings of the National Academy of Sciences*, 32(12):331–332, 1946. 13, 14, 15, 16

[BF92] László Babai and Péter Frankl. *Linear Algebra Methods in Combinatorics: With Applications to Geometry and Computer Science*. Department of Computer Science, univ. of Chicag, 1992. 17

[BOGW88] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 1–10, Chicago, IL, USA, May 2–4, 1988. ACM Press. 1

[Bou99] Jean Bourgain. On triples in arithmetic progression. *Geometric and Functional Analysis*, 9(5):968–984, 1999. 13, 15

[Bou08] Jean Bourgain. Roth's theorem on progressions revisited. *Journal d'Analyse Mathématique*, 104(1):155–192, 2008. 13, 15

[CCD88] David Chaum, Claude Crépeau, and Ivan Damgård. Multiparty unconditionally secure protocols (extended abstract). In *20th Annual ACM Symposium on Theory of Computing*, pages 11–19, Chicago, IL, USA, May 2–4, 1988. ACM Press. 1

[CGS08] Nishanth Chandran, Vipul Goyal, and Amit Sahai. New constructions for UC secure computation using tamper-proof hardware. In Nigel P. Smart, editor, *Advances in Cryptology – EUROCRYPT 2008*, volume 4965 of *Lecture Notes in Computer Science*, pages 545–562, Istanbul, Turkey, April 13–17, 2008. Springer, Heidelberg, Germany. 1

[CLOS02] Ran Canetti, Yehuda Lindell, Rafail Ostrovsky, and Amit Sahai. Universally composable two-party and multi-party secure computation. In *34th Annual ACM Symposium on Theory of Computing*, pages 494–503, Montréal, Québec, Canada, May 19–21, 2002. ACM Press. 1

[CMW05] Claude Crépeau, Kirill Morozov, and Stefan Wolf. Efficient unconditional oblivious transfer from almost any noisy channel. In Carlo Blundo and Stelvio Cimato, editors, *SCN 04: 4th International Conference on Security in Communication Networks*, volume 3352 of *Lecture Notes in Computer Science*, pages 47–59, Amalfi, Italy, September 8–10, 2005. Springer, Heidelberg, Germany. 1

[CT11] Sebastian M Cioaba and Michael Tait. More counterexamples to the alon-saks-seymour and rank-coloring conjectures. *the electronic journal of combinatorics*, 18(P26):1, 2011. 17

[CT13] Sebastian M Cioabă and Michael Tait. Variations on a theme of graham and pollak. *Discrete Mathematics*, 313(5):665–676, 2013. 17

[DI06] Ivan Damgård and Yuval Ishai. Scalable secure multiparty computation. In Cynthia Dwork, editor, *Advances in Cryptology – CRYPTO 2006*, volume 4117 of *Lecture Notes in Computer Science*, pages 501–520, Santa Barbara, CA, USA, August 20–24, 2006. Springer, Heidelberg, Germany. 1

[DNW08] Ivan Damgård, Jesper Buus Nielsen, and Daniel Wichs. Isolated proofs of knowledge and isolated zero knowledge. In Nigel P. Smart, editor, *Advances in Cryptology – EUROCRYPT 2008*, volume 4965 of *Lecture Notes in Computer Science*, pages 509–526, Istanbul, Turkey, April 13–17, 2008. Springer, Heidelberg, Germany. 1

[Dol82] Danny Dolev. The byzantine generals strike again. *J. Algorithms*, 3(1):14–30, 1982. 1

[DPSZ12] Ivan Damgård, Valerio Pastro, Nigel P. Smart, and Sarah Zakarias. Multiparty computation from somewhat homomorphic encryption. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology – CRYPTO 2012*, volume 7417 of *Lecture Notes in Computer Science*, pages 643–662, Santa Barbara, CA, USA, August 19–23, 2012. Springer, Heidelberg, Germany. 1

[Elk10] Michael Elkin. An improved construction of progression-free sets. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 886–905. Society for Industrial and Applied Mathematics, 2010. 13, 14, 15, 16

[GIMS15] Divya Gupta, Yuval Ishai, Hemanta K. Maji, and Amit Sahai. Secure computation from leaky correlated randomness. In Rosario Gennaro and Matthew J. B. Robshaw, editors, *Advances in Cryptology – CRYPTO 2015, Part II*, volume 9216 of *Lecture Notes in Computer Science*, pages 701–720, Santa Barbara, CA, USA, August 16–20, 2015. Springer, Heidelberg, Germany. 2, 3, 4, 5, 6, 11, 29, 30

[GIS+10] Vipul Goyal, Yuval Ishai, Amit Sahai, Ramarathnam Venkatesan, and Akshay Wadia. Founding cryptography on tamper-proof hardware tokens. In Daniele Micciancio, editor, *TCC 2010: 7th Theory of Cryptography Conference*, volume 5978 of *Lecture Notes in Computer Science*, pages 308–326, Zurich, Switzerland, February 9–11, 2010. Springer, Heidelberg, Germany. 1

[GK73] Peter Gács and János Körner. Common information is far less than mutual information. *Problems of Control and Information Theory*, 2(2):149–162, 1973. 20

[GMW87] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game or A completeness theorem for protocols with honest majority. In Alfred Aho, editor, *19th Annual ACM Symposium on Theory of Computing*, pages 218–229, New York City, NY, USA, May 25–27, 1987. ACM Press. 1

[GP71] Ronald L Graham and Henry O Pollak. On the addressing problem for loop switching. *Bell System Technical Journal*, 50(8):2495–2519, 1971. 5, 17

[GP72] Ronald L Graham and Henry O Pollak. On embedding graphs in squashed cubes. In *Graph theory and applications*, pages 99–110. Springer, 1972. 5, 17

[HB87] David Rodney Heath-Brown. Integer sets containing no arithmetic progressions. *J. London Math. Soc.(2)*, 35(3):385–394, 1987. 13, 14

[HIKN08] Danny Harnik, Yuval Ishai, Eyal Kushilevitz, and Jesper Buus Nielsen. OT-combiners via secure computation. In Ran Canetti, editor, *TCC 2008: 5th Theory of Cryptography Conference*, volume 4948 of *Lecture Notes in Computer Science*, pages 393–411, San Francisco, CA, USA, March 19–21, 2008. Springer, Heidelberg, Germany. 6

[HKN+05] Danny Harnik, Joe Kilian, Moni Naor, Omer Reingold, and Alon Rosen. On robust combiners for oblivious transfer and other primitives. In Ronald Cramer, editor, *Advances in Cryptology – EUROCRYPT 2005*, volume 3494 of *Lecture Notes in Computer Science*, pages 96–113, Aarhus, Denmark, May 22–26, 2005. Springer, Heidelberg, Germany. 5

[Hof72] AJ Hoffman. Eigenvalues and partitionings of the edges of a graph. *Linear Algebra and Its Applications*, 5(2):137–146, 1972. 17

[HS12] Hao Huang and Benny Sudakov. A counterexample to the alon-saks-seymour conjecture and related problems. *Combinatorica*, 32(2):205–219, 2012. 17, 18

[IKOS09] Yuval Ishai, Eyal Kushilevitz, Rafail Ostrovsky, and Amit Sahai. Extracting correlations. In *50th Annual Symposium on Foundations of Computer Science*, pages 261–270, Atlanta, GA, USA, October 25–27, 2009. IEEE Computer Society Press. 1, 2, 3, 6

[IL89] Russell Impagliazzo and Michael Luby. One-way functions are essential for complexity based cryptography (extended abstract). In *30th Annual Symposium on Foundations of Computer Science*, pages 230–235, Research Triangle Park, North Carolina, October 30 – November 1, 1989. IEEE Computer Society Press. 1, 6, 16

[IMSW14] Yuval Ishai, Hemanta K. Maji, Amit Sahai, and Jürg Wullschleger. Single-use ot combiners with near-optimal resilience. In *2014 IEEE International Symposium on Information Theory, Honolulu, HI, USA, June 29 - July 4, 2014*, pages 1544–1548. IEEE, 2014. 3, 5, 6

[IPS08] Yuval Ishai, Manoj Prabhakaran, and Amit Sahai. Founding cryptography on oblivious transfer - efficiently. In David Wagner, editor, *Advances in Cryptology – CRYPTO 2008*, volume 5157 of *Lecture Notes in Computer Science*, pages 572–591, Santa Barbara, CA, USA, August 17–21, 2008. Springer, Heidelberg, Germany. 1, 6

[Kah91] Jeff Kahn. *Recent results on some not-so-recent hypergraph matching and covering problems.* DIMACS, Center for Discrete Mathematics and Theoretical Computer Science, 1991. 17

[Kat07] Jonathan Katz. Universally composable multi-party computation using tamper-proof hardware. In Moni Naor, editor, *Advances in Cryptology – EUROCRYPT 2007*, volume 4515 of *Lecture Notes in Computer Science*, pages 115–128, Barcelona, Spain, May 20–24, 2007. Springer, Heidelberg, Germany. 1

[Kil88] Joe Kilian. Founding cryptography on oblivious transfer. In *20th Annual ACM Symposium on Theory of Computing*, pages 20–31, Chicago, IL, USA, May 2–4, 1988. ACM Press. 1, 6, 16

[Kil00] Joe Kilian. More general completeness theorems for secure two-party computation. In *32nd Annual ACM Symposium on Theory of Computing*, pages 316–324, Portland, OR, USA, May 21–23, 2000. ACM Press. 1

[KMQR09] Robin Künzler, Jörn Müller-Quade, and Dominik Raub. Secure computability of functions in the IT setting with dishonest majority and applications to long-term security. In Omer Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 238–255. Springer, Heidelberg, Germany, March 15–17, 2009. 1, 6, 16

[KRW88] Thomas Kratzke, Bruce Reznick, and Douglas West. Eigensharp graphs: Decomposition into complete bipartite subgraphs. *Transactions of the American Mathematical Society*, 308(2):637–653, 1988. 17

[Kus89] Eyal Kushilevitz. Privacy and communication complexity. In *30th Annual Symposium on Foundations of Computer Science*, pages 416–421, Research Triangle Park, North Carolina, October 30 – November 1, 1989. IEEE Computer Society Press. 1, 6, 16

[McN] Nathan McNew. Avoiding geometric progressions in the integers. Available at https://math.dartmouth.edu/graduate-students/works/2013-14/McNew-GradPosterSession.pdf (2017/02/05). 15

[MNPS04] Dahlia Malkhi, Noam Nisan, Benny Pinkas, and Yaron Sella. Fairplay - secure two-party computation system. In Matt Blaze, editor, *Proceedings of the 13th USENIX Security Symposium, August 9-13, 2004, San Diego, CA, USA*, pages 287–302. USENIX, 2004. 1

[MP06] Remo Meier and Bartosz Przydatek. On robust combiners for private information retrieval and other primitives. In Cynthia Dwork, editor, *Advances in Cryptology – CRYPTO 2006*, volume 4117 of *Lecture Notes in Computer Science*, pages 555–569, Santa Barbara, CA, USA, August 20–24, 2006. Springer, Heidelberg, Germany. 6

[MPR09] Hemanta K. Maji, Manoj Prabhakaran, and Mike Rosulek. Complexity of multi-party computation problems: The case of 2-party symmetric secure function evaluation. In Omer Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 256–273. Springer, Heidelberg, Germany, March 15–17, 2009. 1, 6, 16

[MPR12] Hemanta K. Maji, Manoj Prabhakaran, and Mike Rosulek. A unified characterization of completeness and triviality for secure function evaluation. In Steven D. Galbraith and Mridul Nandi, editors, *Progress in Cryptology - INDOCRYPT 2012: 13th International Conference in Cryptology in India*, volume 7668 of *Lecture Notes in Computer Science*, pages 40–59, Kolkata, India, December 9–12, 2012. Springer, Heidelberg, Germany. 1

[MPW07] Remo Meier, Bartosz Przydatek, and Jürg Wullschleger. Robuster combiners for oblivious transfer. In Salil P. Vadhan, editor, *TCC 2007: 4th Theory of Cryptography Conference*, volume 4392 of *Lecture Notes in Computer Science*, pages 404–418, Amsterdam, The Netherlands, February 21–24, 2007. Springer, Heidelberg, Germany. 6

[MS08] Tal Moran and Gil Segev. David and goliath commitments: UC computation for asymmetric parties using tamper-proof hardware. In Nigel P. Smart, editor, *Advances in Cryptology - EUROCRYPT 2008, 27th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Istanbul, Turkey, April 13-17, 2008. Proceedings*, volume 4965 of *Lecture Notes in Computer Science*, pages 527–544. Springer, 2008. 1

[NNOB12] Jesper Buus Nielsen, Peter Sebastian Nordholt, Claudio Orlandi, and Sai Sheshank Burra. A new approach to practical active-secure two-party computation. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology – CRYPTO 2012*, volume 7417 of *Lecture Notes in Computer Science*, pages 681–700, Santa Barbara, CA, USA, August 19–23, 2012. Springer, Heidelberg, Germany. 1

[NW95] Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995. 18

[Pec84] GW Peck. A new proof of a theorem of graham and pollak. *Discrete mathematics*, 49(3):327–328, 1984. 17

[PP10] Vinod M. Prabhakaran and Manoj Prabhakaran. Assisted common information. In *IEEE International Symposium on Information Theory, ISIT 2010, June 13-18, 2010, Austin, Texas, USA, Proceedings*, pages 2602–2606. IEEE, 2010. 20

[PP11] Vinod M. Prabhakaran and Manoj Prabhakaran. Assisted common information: Further results. In Alexander Kuleshov, Vladimir Blinovsky, and Anthony Ephremides, editors, *2011 IEEE International Symposium on Information Theory Proceedings, ISIT 2011, St. Petersburg, Russia, July 31 - August 5, 2011*, pages 2861–2865. IEEE, 2011. 20

[PW08] Bartosz Przydatek and Jürg Wullschleger. Error-tolerant combiners for oblivious primitives. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP 2008: 35th International Colloquium on Automata, Languages and Programming, Part II*, volume 5126 of *Lecture Notes in Computer Science*, pages 461–472, Reykjavik, Iceland, July 7–11, 2008. Springer, Heidelberg, Germany. 6

[Raz92] Alexander A Razborov. The gap between the chromatic number of a graph and the rank of its adjacency matrix is superlinear. *Discrete mathematics*, 108(1):393–396, 1992. 18

[RBO89] Tal Rabin and Michael Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority (extended abstract). In *21st Annual ACM Symposium on Theory of Computing*, pages 73–85, Seattle, WA, USA, May 15–17, 1989. ACM Press. 1

[Rot53] Klaus F Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953. 13, 14

[San11] Tom Sanders. On roth's theorem on progressions. *Annals of Mathematics*, pages 619–636, 2011. 13, 15

[SS42] Raphaël Salem and Donald C Spencer. On sets of integers which contain no three terms in arithmetical progression. *Proceedings of the National Academy of Sciences*, 28(12):561–563, 1942. 14

[Sze90] Endre Szemerédi. Integer sets containing no arithmetic progressions. *Acta Mathematica Hungarica*, 56(1-2):155–158, 1990. 13, 15

[Tve82] Helge Tverberg. On the decomposition of kn into complete bipartite graphs. *Journal of Graph Theory*, 6(4):493–494, 1982. 17

[Vis08] Sundar Vishwanathan. A polynomial space proof of the graham–pollak theorem. *Journal of Combinatorial Theory, Series A*, 115(4):674–676, 2008. 17

[Vis13] Sundar Vishwanathan. A counting proof of the graham–pollak theorem. *Discrete Mathematics*, 313(6):765–766, 2013. 17

[VL85] JH Van Lint. {0, 1,*} distance problems in combinatorics. 1985. 17

[vLW01] Jacobus Hendricus van Lint and Richard Michael Wilson. *A course in combinatorics*. Cambridge university press, 2001. 17

[WW05] Stefan Wolf and Jürg Wullschleger. New monotones and lower bounds in unconditional two-party computation. In Victor Shoup, editor, *Advances in Cryptology – CRYPTO 2005*, volume 3621 of *Lecture Notes in Computer Science*, pages 467–477, Santa Barbara, CA, USA, August 14–18, 2005. Springer, Heidelberg, Germany. 20

[WW06] Stefan Wolf and Jürg Wullschleger. Oblivious transfer is symmetric. In Serge Vaudenay, editor, *Advances in Cryptology – EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 222–232, St. Petersburg, Russia, May 28 – June 1, 2006. Springer, Heidelberg, Germany. 1, 4

[Wyn75] Aaron D. Wyner. The common information of two dependent random variables. *IEEE Transactions on Information Theory*, 21(2):163–179, 1975. 3, 5, 18, 20, 22

[Yao82] Andrew Chi-Chih Yao. Protocols for secure computations (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science*, pages 160–164, Chicago, Illinois, November 3–5, 1982. IEEE Computer Society Press. 1

[YY06] Weigen Yan and Yeong-Nan Yeh. A simple proof of graham and pollak's theorem. *Journal of Combinatorial Theory, Series A*, 113(5):892–893, 2006. 17

A Outline of the Security Proof of Protocol in Figure 7

The protocol presented in [GIMS15] is specific to $\mathbb{K} = \mathbb{GF}[2]$. So, the '+' and '-' over the field $\mathbb{K}$ are identical operations. We have to be careful because our protocol is defined over arbitrary fields. So, we argue the correctness of the correlation extractor protocol in Figure 7. The correctness follows from the expression provided below.

$$
\begin{aligned}
X_0 + Y_0 &= \sum_{i=1}^{\eta} X_i Y_i \\
(\widetilde{X}_0, \widetilde{X}_1, \ldots, \widetilde{X}_\eta) &\in \mathcal{C}^{\perp} \\
(M_1, \ldots, M_\eta) &= (Y_1, \ldots, Y_\eta) - (\widetilde{X}_1, \ldots, \widetilde{X}_\eta) \\
(\widetilde{A}_0, \widetilde{A}_1, \ldots, \widetilde{A}_\eta) &\in \mathcal{C} \\
\widetilde{A}_0 \widetilde{X}_0 &= -\sum_{i=1}^{\eta} \widetilde{A}_i \widetilde{X}_i \\
(\alpha_1, \ldots, \alpha_\eta) &= (X_1, \ldots, X_\eta) + (\widetilde{A}_1, \ldots, \widetilde{A}_\eta) \\
\beta &= \sum_{i=1}^{\eta} X_i M_i - \widetilde{B}_0 - X_0 \\
&= \sum_{i=1}^{\eta} (X_i Y_i - X_i \widetilde{X}_i) - \widetilde{B}_0 - X_0 \\
&= (X_0 + Y_0) - \sum_{i=1}^{\eta} X_i \widetilde{X}_i - \widetilde{B}_o - X_0 \\
&= Y_0 - \sum_{i=1}^{\eta} X_i \widetilde{X}_i - \widetilde{B}_0 \\
-\sum_{i=1}^{\eta} \alpha_i \widetilde{X}_i &= -\sum_{i=1}^{\eta} (X_i + \widetilde{A}_i) \widetilde{X}_i \\
&= -\sum_{i=1}^{\eta} X_i \widetilde{X}_i - \sum_{i=1}^{\eta} \widetilde{A}_i \widetilde{X}_i \\
&= -\sum_{i=1}^{\eta} X_i \widetilde{X}_i + A_0 X_0 \\
Z_0 &= -\sum_{i=1}^{\eta} \alpha_i \widetilde{X}_i - \beta + Y_0 \\
&= \left(-\sum_{i=1}^{\eta} X_i \widetilde{X}_i + A_0 X_0 \right) - \left(Y_0 - \sum_{i=1}^{\eta} X_i \widetilde{X}_i - \widetilde{B}_0 \right) + Y_0 \\
&= A_0 X_0 + B_0
\end{aligned}
$$

Similar to Gupta et al. [GIMS15], using the unpredictability lemma over fields (Lemma 2), we can argue the security of the protocol in Figure 7.

B Proof of Lemma 2

Let M be the random variable for $m_{[\eta]}$. Let G_0 be the first column of G, and G' be the matrix after truncating the first column of G. Let $L = \mathcal{L}(M)$ represent the random variable for the leakage. It suffices to show that

$$2\mathsf{SD}((XG_0, XG' + M, G, L), (U_{\mathbb{K}^{\eta+1}}, G, L)) \leqslant \sqrt{\frac{|\mathbb{K}|\, 2^t}{|\mathbb{K}|^k}}$$

Utilizing the property of the random choice of the generator matrix, analogous to Gupta et al. [GIMS15], we make the following claim.

Claim 1.

$$\mathbb{E}_G(\widehat{XG_0, XG|G})(S)^2 \leqslant \frac{1}{|\mathbb{K}|^k |\mathbb{K}|^{2(\eta+1)}}$$

Now, we upper-bound the statistical distance as follows.

$$\begin{aligned}
&2\mathsf{SD}((XG_0, XG' + M, G, L), (U_{\mathbb{K}^{\eta+1}} G, L)) \\
&= \mathbb{E}_{G,L}\left[2\mathsf{SD}((XG_0, XG + M|G, L), (U_{\mathbb{K}^{\eta+1}}|G, L))\right] \\
&= \mathbb{E}_{G,L}\left[\sum_y |(XG_0, XG + M|G, L)(y) - (U_{\mathbb{K}^{\eta+1}}|G, L)(y)|\right] \\
&\leqslant \mathbb{E}_{G,L}\left[|\mathbb{K}|^{\eta+1}\sqrt{\mathbb{E}_y\left((XG_0, XG + M|G, L)(y) - (U_{\mathbb{K}^{\eta+1}}|G, L)(y)\right)^2}\right] && \text{[By Cauchy Schwartz]} \\
&= |\mathbb{K}|^{\eta+1} \cdot \mathbb{E}_{G,L}\sqrt{\sum_S \widehat{(XG_0, XG + M|G, L) - (U_{\mathbb{K}^{\eta+1}}|G, L)}(S)^2} && \text{[By Parseval's Identity]} \\
&= |\mathbb{K}|^{\eta+1} \cdot \mathbb{E}_{G,L}\sqrt{\sum_{S_0 \neq 0} \widehat{(XG_0, XG + M|G, L)}(S)^2} \\
&= |\mathbb{K}|^{\eta+1} \cdot \mathbb{E}_{G,L}\sqrt{\sum_{S_0 \neq 0} \left(|\mathbb{K}|^{\eta+1}(\widehat{XG_0, XG|G, L})(S) \cdot \widehat{(0, M|G, L)}(S)\right)^2} && \text{[By Convolution]} \\
&= |\mathbb{K}|^{2(\eta+1)} \cdot \mathbb{E}_{G,L}\sqrt{\sum_{S_0 \neq 0} (\widehat{XG_0, XG|G})(S)^2 \cdot \widehat{(0, M|L)}(S)^2} \\
&\leqslant |\mathbb{K}|^{2(\eta+1)}\sqrt{\sum_{S_0 \neq 0} \mathbb{E}_{G,L}\left[(\widehat{XG_0, XG|G})(S)^2 \cdot \widehat{(0, M|L)}(S)^2\right]} && \text{[By Jensen's Inequality]} \\
&= |\mathbb{K}|^{2(\eta+1)}\sqrt{\sum_{S_0 \neq 0} \mathbb{E}_L\widehat{(0, M|L)}(S)^2 \cdot \mathbb{E}_G(\widehat{XG_0, XG|G})(S)^2} \\
&\leqslant |\mathbb{K}|^{2(\eta+1)}\sqrt{\sum_{S_0 \neq 0} \mathbb{E}_L\widehat{(0, M|L)}(S)^2 \cdot \frac{1}{|\mathbb{K}|^{k+2(\eta+1)}}} \\
&\leqslant |\mathbb{K}|^{2(\eta+1)}\sqrt{\frac{2^t}{|\mathbb{K}|^{\eta+1} \cdot |\mathbb{K}|^{\eta}} \cdot \frac{1}{|\mathbb{K}|^{k+2(\eta+1)}}} \\
&= \sqrt{\frac{|\mathbb{K}|\, 2^t}{|\mathbb{K}|^k}}
\end{aligned}$$

This completes the proof of the unpredictability lemma.

C Proof of Lemma 5

In this section, we will prove the estimates of the simple partition number of a few correlations relevant to our work.

C.1 Bound on the Inner-Product Correlation.

Define $\mathsf{IP}_\lambda(\mathbb{F}^n)$ to be the bipartite graph over the partite sets $L = R = \mathbb{F}^n$ such that there is an edge between $x = (x_1, x_2, \cdots, x_n) \in L$ and $y = (y_1, y_2, \cdots, y_n) \in R$ if $\langle x, y\rangle := \sum_{i=1}^n x_i y_i = \lambda$.

Lemma 6.

$$\mathsf{sp}\left(\mathsf{IP}_\lambda(\mathbb{F}^n)\right) \leqslant |\mathbb{F}|^{\lceil n/2 \rceil}$$

We shall prove a more generalized result as following. Then Lemma 6 follows immediately as a corollary.

Lemma 7. *For any natural numbers n and m such that $n/2 \leqslant m \leqslant n$, there exists a simple decomposition of $\mathsf{IP}_\lambda(\mathbb{F}^n)$ of size $|\mathbb{F}|^m$.*

Proof. Consider any subspace U of $\mathbb{F}^n$ of dimension $(n-m) \leqslant n/2$. Let V be the dual of U and, hence, has dimension $m \geqslant n/2$. Note that $\mathbb{F}^n$ is a group under addition, so the number of cosets of U is $[\mathbb{F}^n : U] = |\mathbb{F}|^m$ and the number of cosets of V is $[\mathbb{F}^n : V] = |\mathbb{F}|^{(n-m)}$, where $n - m \leqslant m$. Let $\pi : \mathbb{Z}_{|\mathbb{F}|^m} \times \mathbb{Z}_{|\mathbb{F}|^{n-m}} \to \mathbb{Z}_{|\mathbb{F}|^m}$ be a function defined by: $\pi(k, y) = (k + y) \mod |\mathbb{F}|^m$.

For $0 \leqslant k < |\mathbb{F}|^m$, define $G^{(k)}$ as the bipartite graph over the partite sets L and R such that every edge in $\mathsf{IP}_\lambda(\mathbb{F})^n$ between the cosets $x + U$ and $y + V$, such that $\pi(k, y) = x$, is included.

To complete the proof of Lemma 7, the following claims suffice.

Claim 2. *For every k such that $0 \leqslant k < |\mathbb{F}|^m$, the graph $G^{(k)}$ is simple.*

Proof. Note that if there is an edge between $(x+u) \in (x+U)$ and $(y+v) \in (y+V)$, then $\pi(k, y) = x$. For the sake of contradiction, assume $G^{(k)}$ is not simple for some k. Then $G^{(k)}$ must contain a OR minor (or "Z" shape). More concretely, there exist four distinct nodes $x + u, x' + u' \in L$ and $y+v, y+v' \in R$, where $u, u' \in U$ and $v, v' \in V$ such that there are exactly three edges between the two left nodes and the two right nodes. Without loss of generality, suppose that the edge between $x + u$ and $y'+v'$ is missing. Since there is an edge between $x+u$ and $y+v$, by above observation $\pi(k, y) = x$; also since there is an edge between $x' + u'$ and $y + v$, by above observation $\pi(k, y) = x'$. Thus, $x = x'$. Similarly, we have $y = y'$. Next, since $\langle x + u, y + v\rangle = \langle x + u', y + v\rangle = \langle x + u', y + v'\rangle = \lambda$ and $\langle u, v\rangle = \langle u', v\rangle = \langle u', v'\rangle = 0$,

$$\begin{aligned}
\langle x, y\rangle + \langle x, v\rangle + \langle u, y\rangle &= \lambda \\
\langle x, y\rangle + \langle x, v\rangle + \langle u', y\rangle &= \lambda \\
\langle x, y\rangle + \langle x, v'\rangle + \langle u', y\rangle &= \lambda
\end{aligned}$$

The first and second equations imply that $\langle u', y\rangle = \langle u, y\rangle$. Substituting this into the third equation, we get $\langle x, y\rangle + \langle x, v'\rangle + \langle u, y\rangle = \lambda$. Notice that $\langle u, v'\rangle = 0$, therefore $\langle x + u, y + v'\rangle = \lambda$, contradiction with the assumption that $\langle x + u, y + v'\rangle \neq \lambda$. □

Claim 3. *For every edge $e \in E(\mathsf{IP}_\lambda(\mathbb{F}^n))$, there exists a unique k such that $G^{(k)}$ contains the edge e.*

Proof. Let $x+U$ be the coset that contains the left endpoint of e, where $0 \leqslant x < |\mathbb{F}|^m$, and let $y+V$ be the coset that contains the right endpoint of e, where $0 \leqslant y < |\mathbb{F}|^{n-m}$. The equation $(k+y) \mod |\mathbb{F}|^m = x$ has a unique solution, which is $(x-y) \mod |\mathbb{F}|^m$ in $\mathbb{Z}_{|\mathbb{F}|^m}$. It means that there exists a unique integer k such that $0 \leqslant k < |\mathbb{F}|^m$ and $\pi(k,y) = x$. Clearly $e \in E(G^{(k)})$. Hence, $G^{(k)}$ is the unique graph that contains the edge e. □

This completes the proof of Lemma 7. □

Now, we prove that $\mathsf{sp}\left(\mathsf{IP}\left(\mathbb{F}\right)^n\right) \leqslant |\mathbb{F}|^{\lceil (n+1)/2 \rceil}$. The graph $\mathsf{IP}\left(\mathbb{F}^n\right)$ can be partitioned into $|\mathbb{F}|$ disjoint components such that every component G_λ, for $\lambda \in \mathbb{F}$, satisfies that there is an edge between $(x_0, x_1, \cdots, x_{n-1})$ and $(y_0, y_1, \cdots, y_{n-1})$ if $x_0+y_0 = x_1y_1+\cdots x_{n-1}y_{n-1} = \lambda$. The equation $x_0 + y_0 = \lambda$ has exactly $|\mathbb{F}|$ solutions. Thus, each G_λ can be viewed as the disjoint union of $|\mathbb{F}|$ graphs $G_\lambda = \mathsf{IP}_\lambda(\mathbb{F}^{n-1})$. By Lemma 6, $\mathsf{sp}(G_\lambda) \leqslant |\mathbb{F}|^{\lceil (n-1)/2 \rceil}$. Therefore, we can conclude that $\mathsf{sp}(\mathsf{IP}\left(\mathbb{F}^n\right)) \leqslant |\mathbb{F}| \cdot |\mathbb{F}|^{\lceil (n-1)/2 \rceil} = |\mathbb{F}|^{\lceil (n+1)/2 \rceil}$.

C.2 Bound on the Random Oblivious Linear-function Evaluation Correlation.

Let $\widetilde{n} = n/2$. Then, we shall show that $\mathsf{sp}\left(\mathsf{ROLE}\left(\mathbb{F}\right)^{\widetilde{n}}\right) \leqslant |\mathbb{F}|^{\widetilde{n}/2}$. To aid our proof, we need the following lemmas.

Lemma 8. *For any bipartite graphs G and H, $\mathsf{sp}(G \times H) \leqslant \mathsf{sp}(G) \cdot \mathsf{sp}(H)$.*

Proof. Suppose $\mathsf{sp}(G) = p$ and $\mathsf{sp}(H) = q$. Then let $G^{(1)}, \ldots, G^{(p)}$ be the smallest simple partition of G, and let $H^{(1)}, \ldots, H^{(q)}$ be a smallest simple partition of H. For each (i,j) in $[p] \times [q]$, consider the graph $J^{(i,j)}$ defined by $G^{(i)} \times H^{(j)}$. Since $G^{(i)}$ and $H^{(j)}$ are simple, $G^{(i)} \times H^{(j)}$ is also simple. It is clear that the set of graphs $J^{(i,j)}$ is a partition of the graph $G \times H$. Therefore, $J^{(1,1)}, \ldots, J^{(p,q)}$ is a simple partition of $G \times H$, which implies that $\mathsf{sp}(G \times H) \leqslant p \cdot q = \mathsf{sp}(G) \cdot \mathsf{sp}(H)$. □

Applying Lemma 8 inductively for $G = H$, we have the following result as a corollary.

Corollary 4. *For any bipartite graph G, we have $\mathsf{sp}\left(G^n\right) \leqslant \mathsf{sp}\left(G\right)^n$.*

Lemma 9 (Simple partition number of $\mathsf{ROLE}\left(\mathbb{F}\right)$)**.**

$$\mathsf{sp}\left(\mathsf{ROLE}\left(\mathbb{F}\right)\right) \leqslant |\mathbb{F}|$$

Proof. Note that $\mathsf{ROLE}\left(\mathbb{F}\right)$ is a $|\mathbb{F}|$-regular bipartite graph. Thus, $|\mathbb{F}|$ perfect matchings give us a simple partition of size $|\mathbb{F}|$. □

Lemma 10 (Simple partition number of $\mathsf{ROLE}\left(\mathbb{F}\right)^2$)**.**

$$\mathsf{sp}\left(\mathsf{ROLE}\left(\mathbb{F}\right)^2\right) \leqslant |\mathbb{F}|$$

Proof. Let $G = (L, R, E)$ be the graph representing $\mathsf{ROLE}\left(\mathbb{F}\right)^2$. Consider the following partition of edges: for every $c \in \mathbb{F}$,

$$E^{(c)} := \{((a_1, b_1, a_2, b_2), (x_1, z_1, x_2, z_2)) : a_1 - x_2 = c\}$$

Then, we create the following graphs: for every $c \in \mathbb{F}$,

$$G^{(c)} := \left(L, R, E^{(c)}\right).$$

We claim that for every $c \in \mathbb{F}$, $G^{(c)}$ is simple. Suppose for any $c \in \mathbb{F}$, we have

$$((a_1, b_1, a_2, b_2), (x_1, z_1, x_2, z_2)) \in E^{(c)} \tag{1}$$
$$\left((a_1', b_1', a_2', b_2'), (x_1, z_1, x_2, z_2)\right) \in E^{(c)} \tag{2}$$
$$\left((a_1', b_1', a_2', b_2'), (x_1', z_1', x_2', z_2')\right) \in E^{(c)}. \tag{3}$$

We shall show that

$$\left((a_1, b_1, a_2, b_2), (x_1', z_1', x_2', z_2')\right) \in E^{(c)}$$

This suffices to show that the graph $G^{(c)}$ is a simple graph. Fix (a_1, b_1, a_2, b_2). Since (a_1, b_1, a_2, b_2) is connected to (x_1, z_1, x_2, z_2), we have:

$$(x_1, z_1, x_2, z_2) = (x_1, a_1 x_1 + b_1, a_1 - c, a_2(a_1 - c) + b_2)$$

Since (x_1, z_1, x_2, z_2) is connected to (a_1', b_1', a_2', b_2'), we have:

$$(a_1', b_1', a_2', b_2') = (a_1, b_1, a_2', (a_2 - a_2')(a_1 - c) + b_2)$$

Since (a_1', b_1', a_2', b_2') is connected to (x_1', z_1', x_2', z_2'), we have:

$$(x_1', z_1', x_2', z_2') = (x_1', b_1 x_1' + z_1', a_1 - c, a_2(a_1 - c) + b_2)$$

This implies that (a_1, b_1, a_2, b_2) is connected to (x_1', z_1', x_2', z_2'). This completes the proof. □

Now, we get back to proving the estimate of the simple partition number of $\mathsf{ROLE}(\mathbb{F})^{\widetilde{n}}$. Consider two cases.

1. If $\widetilde{n}$ is even, by Lemma 10, we know that $\mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})^2\right) \leqslant |\mathbb{F}|$. Then directly applying Corollary 4 with $G = \mathsf{ROLE}(\mathbb{F})^2$, we have that

$$\mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})^{\widetilde{n}}\right) \leqslant \mathsf{sp}\left(\left(\mathsf{ROLE}(\mathbb{F})^2\right)^{\widetilde{n}/2}\right) \leqslant |\mathbb{F}|^{\widetilde{n}/2}$$

2. If $\widetilde{n}$ is odd, then $\widetilde{n} - 1$ is even. We have

$$\begin{aligned}
\mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})^{\widetilde{n}}\right) &\leqslant \mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})\right) \cdot \mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})^{\widetilde{n}-1}\right) && [\text{ By } \textit{Lemma 8}] \\
&\leqslant |\mathbb{F}| \cdot \mathsf{sp}\left(\mathsf{ROLE}(\mathbb{F})^{\widetilde{n}-1}\right) && [\text{ By } \textit{Lemma 9}] \\
&\leqslant |\mathbb{F}| \cdot |\mathbb{F}|^{(\widetilde{n}-1)/2} && [\text{ By applying Case 1}] \\
&= |\mathbb{F}|^{\lceil \widetilde{n}/2 \rceil}
\end{aligned}$$