Secure Non-Interactive Reducibility is Decidable

Kaartik Bhushan[1], Ankit Kumar Misra[1], Varun Narayanan[2⋆], and Manoj Prabhakaran[1]

[1] Indian Institute of Technology Bombay, India
{kbhushan,ankitkmisra,mp}@cse.iitb.ac.in
[2] Technion, Israel
varunnkv@gmail.com

Abstract. Secure Non-Interactive Reductions (SNIR) is a recently introduced, but fundamental cryptographic primitive. The basic question about SNIRs is how to determine if there is an SNIR from one 2-party correlation to another. While prior work provided answers for several pairs of correlations, the possibility that this is an undecidable problem in general was left open. In this work we show that the existence of an SNIR between *any pair of correlations* can be determined by an algorithm.

At a high-level, our proof follows the blueprint of a similar (but restricted) result by Khorasgani et al. But combining the spectral analysis of SNIRs by Agrawal et al. (Eurocrypt 2022) with a new variant of a "junta theorem" by Kindler and Safra, we obtain a complete resolution of the decidability question for SNIRs. The new junta theorem that we identify and prove may be of independent interest.

1 Introduction

The notion of Secure Non-Interactive Reductions (SNIR) has only recently been formally defined [2,1,18], but it is a fundamental cryptographic primitive that lies at the intersection of several major lines of research in information-theory and cryptography. On the one hand, it is a model of information-theoretically secure 2-party computation, using correlated randomness [13,14,19,16]. It is a minimal model *without any communication*, pushing the limits of minimalism in secure computation, as initiated by the influential work of Feige et al. [8]. Its non-secure counterpart, called *non-interactive simulation* commands a rich literature in both information-theory and computer science literature spanning half a century [10,27,29,3,17,11,7,26]. Another important motivation behind SNIR is also its relevance to *cryptographic complexity* [5,23,22,4,24] – namely, measuring the complexity of a function in terms of the number of samples of a correlation that need to be used in an (interactive) secure 2-party computation protocol for the function. As pointed out in [1], understanding the power of SNIR is an important part of understanding the interactive secure 2-party computation protocol for an inputless function: such a protocol consists of an interaction phase

⋆ Supported by ERC Project NTSC (742754) and ISF Grants 1709/14 & 2774/20.

(with no security requirements of its own) followed by an SNIR used to securely sample the output from the correlated views at the end of the interaction.

Finally, and significantly, studying the minimalistic model of SNIR leads us to mathematical tools that are relatively unexploited in classical cryptography, including tools from spectral graph theory and harmonic analysis [1,2,18]. Conversely, as is the case in this work, studying SNIRs can lead to contributions back to the development of these tools and their applicability.

Decidability of SNIR. SNIR is a notion of reduction from a (2-output) *target* distribution D to a *source* distribution C. It is simply a statistically secure 2-party computation protocol for sampling from D, when the parties are given access to samples from C, with the restriction that the parties cannot communicate at all. (In this model, semi-honest security and UC security are equivalent.)

The most fundamental question about SNIRs is the decidability of the following problem:

> **SNIR Problem:** *Given a pair of correlations (C, D), does there exist a statistical SNIR from D to C?*

In the works that defined SNIR, this question was tackled for specific pairs of correlations, using arguments specialized for them [2,1,18]. In [2,18], the authors insightfully observed that in certain cases, *a statistically secure SNIR implies a perfectly secure one*, which can in turn be used to design an algorithm to decide the existence of an SNIR. In this work too, we follow the same high-level approach. Further, [2,18] showed that Fourier analytic techniques can be used to prove the statistical-to-perfect security result. However, the decidability results in [2,18] were restricted to two specific target distributions, and did not cover weak notions of security (with only "vanishing" error).

In this work, starting from the spectral analysis of [1] (involving eigenvectors, or more precisely, the singular value decomposition of the "correlation operator"), we apply Fourier analytic techniques to SNIRs in an alternate fashion, to obtain a *full answer* to the fundamental decidability question.

Our Contributions. We summarize our contributions below:

- Our main technical result is a statistical-to-perfect security result for SNIRs, which shows that, for a pair of correlations D and C, a statistically secure SNIR (possibly with weak security) exists from D to C iff there is a perfectly secure SNIR from D to $C^{\otimes \ell}$ for some finite ℓ (that can be computed from D and C). The formal statement, in Theorem 1, involves certain technical restrictions on D and C, which are essential (but not a barrier to the decidability result).
  - In order to prove this, we formulate and prove a new "junta theorem" for "generalized Fourier transforms," that may be of independent interest. This is stated as Theorem 2 and proven in Section 5.
- Based on the above, we show that the SNIR problem is decidable.[3]

[3] For simplicity, we assume a computational model in which real numbers can be represented, computed upon (w.r.t. addition, multiplication and division), and com-

– We also illustrate how the statistical-to-perfect security result can be used to obtain new combinatorial necessary conditions for an SNIR to exist between a pair of correlations; these combinatorial conditions can in turn be used to rule out an SNIR from OT correlation to Rabin OT correlation, that was not covered by prior results.

We remark that our decidability result subsumes that of [18] in a couple of ways: it works for all pairs of correlations, and further works even for a very weak notion of security. That is, when our algorithm says "No" it rules out an SNIR with error that goes to 0 however slowly, and when we say "Yes" we obtain an SNIR with either perfect security (in the absence of common information) or negligible error. In contrast, the algorithm in [18] could not rule out an SNIR with error going to 0 slower than $1/n$ where n denotes the number copies of the source correlation used. More significantly, the algorithm of [18] is restricted to two special target correlations.

Related Work. As already mentioned, several lines of work in information-theoretically secure cryptography intersect with SNIR. Here we clarify the connection with some recent works.

SNIR was defined independently in two concurrent works [2,1], and was further developed in [18], which explicitly addressed the decidability of the SNIR problem. The approach of employing a statistical-to-perfect security result, and the general idea of using Fourier analysis to prove it, were both present in [18].

A similar sounding concept, called *Secure Zero Communication Reduction* (SZCR) was introduced in [24]. It is instructive to compare both SNIR and SZCR with the standard notion of (semi-honest) secure reduction (SR) to a correlation like OT (more familiarly known as 2-PC in the OT-hybrid model). Roughly put,

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

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

– Lower bounds (or impossibility results) for SNIR are an "easier" target than those for SR, and would provide a platform for nurturing new techniques; as and when we completely settle a question for SNIR (as we do here), we can approach SR by relaxing the model (e.g., allow one-directional communication).
– Lower bounds for SZCR are formally (but not necessarily conceptually) harder than those for SR. Here we seek to develop new techniques by asking simpler

pared exactly. The results would extend to all reasonable models of computing with a subset of real numbers, that is closed under these operations.

variants of the lower bound question: e.g., existential questions (a la the "invertible rank conjecture" of [24]) or lower bounds for randomized functions (as in [15]).[4] Also, the new perspective provided by SZCR may lead to fresh approaches to the original hard lower bound problems of SR.

2 Technical Overview

Recap of SNIR. We start with a brief recap of SNIR, as defined in [1], largely borrowing from the overview in that paper. As shown in [1], there are in fact multiple perspectives of SNIR, and we profit from switching among them as appropriate.

- An SNIR is simply a statistically secure 2-party computation protocol for an inputless functionality (namely, sampling from a 2-output *target* distribution D), in which the parties have access to a setup in the form of another inputless functionality (namely, sampling i.i.d. samples from a *source* distribution C), with the restriction that the parties cannot exchange messages.
- *Equivalently*, an SNIR can be specified as a pair of stochastic "protocol" matrices (A, B) representing Alice and Bob's actions (mapping a symbol from the source to a symbol in the target), and a pair of "simulation" matrices (U, V) such that – restricting here to the case of perfect security – they satisfy the following correctness and privacy conditions:

$$A^\intercal \boldsymbol{C}\, B = D \qquad A^\intercal \boldsymbol{C} = D\, V \qquad \boldsymbol{C}\, B = U^\intercal\, D. \tag{1}$$

  Here we have written $\boldsymbol{C}$ to denote $C^{\otimes n}$ where n is the number of i.i.d. samples from C that the protocol uses. In the general case of statistical security, there is a family of protocols indexed by the security parameter (n is allowed to increase with the security parameter), and the equalities above admit an additive (matrix) error term, whose (suitably defined) norms can be bounded by vanishing quantities.
- Finally, there is a *spectral perspective* of an SNIR. This is a set of *necessary* conditions on a pair of matrices $(\widehat{A}, \widehat{B})$ derived from an SNIR (A, B), and which satisfy a set of conditions analogous to the original security conditions as follows (restricting here to perfect security):

$$\begin{gathered} \widehat{A} = \boldsymbol{F}_C A \boldsymbol{F}_D^{-1} \qquad \widehat{B} = \boldsymbol{G}_C B \boldsymbol{G}_D^{-1} \\ \widehat{A}^\intercal \widehat{A} = I \qquad \widehat{B}^\intercal \widehat{B} = I \\ \widehat{A}^\intercal \boldsymbol{\Sigma}_C \widehat{B} = \boldsymbol{\Sigma}_D \qquad \widehat{A}^\intercal \boldsymbol{\Sigma}_C = \boldsymbol{\Sigma}_D \widehat{B}^\intercal \qquad \boldsymbol{\Sigma}_C \widehat{B} = \widehat{A} \boldsymbol{\Sigma}_D \end{gathered} \tag{2}$$

  where $(\boldsymbol{F}_C, \boldsymbol{\Sigma}_C, \boldsymbol{G}_C)$ and $(\boldsymbol{F}_D, \boldsymbol{\Sigma}_D, \boldsymbol{G}_D)$ are matrices associated with C and D, respectively, via singular value decomposition (with some careful scaling

[4] [15] is a concurrent submission to this conference and it also includes the above comparison between SNIR and SZCR.

to account for the possibly non-uniform marginal distributions of C and D). This view uses notions from spectral graph theory to study correlations using their singular values.

In this work we shall exploit yet another perspective of an SNIR: namely, a Fourier analytic perspective. While closely related to the spectral perspective above, this perspective focuses on the case when the source distribution is of the form $C^{\otimes n}$, and treats the protocols as *functions* that take n-tuples as inputs. The Fourier analytic perspective is crucial in investigating if protocols can actually use an increasing number of copies of C. This perspective was already insighfully exploited in [2] for some specific correlations which could be related to the Fourier basis; however, starting from our spectral perspective above, we discover that any source correlation can be related to an appropriate *generalized* Fourier basis.

Statistical to Perfect Security. At a high level, the plan for decidability follows that of [2,18], namely, to show that a statistical reduction from D to C implies a deterministic, perfect reduction, using only a *constant number* of copies of C. (The number of copies of C needed should be effectively determinable from the correlations D and C.) Then, to see if there is a reduction, it is enough to search among a finite number of protocols. The outline of how we carry this out is as follows:

1. Our starting point is the *spectral protocol* characterization from [1] shown in (2). We focus on $\boldsymbol{F}_C$. We observe that multiplying by $\boldsymbol{F}_C$ corresponds to a "generalized Fourier transform." Hence we can interpret the columns of $\widehat{A}$ as a generalized Fourier transform applied to the columns of the matrix $\mathring{A} := A\boldsymbol{F}_D^{-1}$. ($\mathring{A}$ could be thought of as corresponding to a "half-way spectral protocol.")
   $\mathring{A}$ is a matrix with real entries, whose rows are indexed by symbols in $\mathcal{X}^n$, where $\mathcal{X}$ is the alphabet of the distribution C (on Alice's side). So each column of $\mathring{A}$ can be interpreted as a function $\boldsymbol{a} : \mathcal{X}^n \to \mathbb{R}$. A generalized Fourier transform writes this function as a linear combination of basis functions of the form $\boldsymbol{\gamma} : \mathcal{X}^n \to \mathbb{R}$. Nominally, each basis function takes n inputs from $\mathcal{X}$, but may depend on fewer of them (e.g., the basis contains the constant function which depends on 0 inputs); the number of inputs it actually depends on is called the *degree* of a basis function.
2. Then we use the spectral protocol conditions of [1] to obtain "approximate degree bounds" on the columns of $\mathring{A}$. That is, we show that under the generalized Fourier transform mentioned above, the contribution from higher degree basis functions has low "energy" (Lemma 7).
   There is a caveat: Each column of $\mathring{A}$ is associated with a singular value of (a normalized version of) D; the degree bound holds only for columns for which the singular value associated with them is non-zero. Below, we write $\breve{A}$ to denote $\mathring{A}$ restricted to these columns with the degree bound.
3. Next, we appeal to a "junta theorem" to argue that each column of $\breve{A}$ (interpreted as a function, for which the approximate degree bound holds) can be well-approximated by a "junta" — i.e., a function which depends only on a

constant number of its inputs. Here, the approximation guarantee given by the junta is in the sense that it matches the original function exactly on most inputs.

- While there are several junta theorems available in the literature, the version we need (for generalized Fourier transforms) has not been previously stated or proved. As such, we adapt a proof of the Kindler-Safra junta theorem [21] by Filmus [9] for our purposes. Presenting this more general version of the junta theorem is a contribution of ours that may be of independent interest.

4. The next step is to translate the junta approximations of the columns of $\check{A}$ to such an approximation of the protocol matrix A itself. For this step, we invoke an important insight which is evident from the cryptographic perspective: An SNIR exists from D to C iff there is one from D' to C, where D' is a "non-redundant" version of D, which merges output symbols which are "equivalent." This means that for the decidability question, we can w.l.o.g. restrict ourselves to the case when D is non-redundant. This insight was already crucially used in [1]. In our case, we further rely on it at this step: We show that if D is non-redundant, then each row of A is fully determined by the corresponding row of $\check{A}$ (Lemma 5). This also relies on another assumption that [1] showed can be made w.l.o.g. — that A is deterministic — thanks to a determinization process that retains statistical security (with a polynomially bounded increase in error). Then, an approximation of $\check{A}$ (in which most rows are correct) yields a similar approximation of A; further, since the approximations are juntas, so is each column of the approximation of A.
   The upshot of this step is that Alice's protocol matrix A can be replaced by one which consults only a constant number of the n copies of C that it is given access to, without increasing the error too much. This still yields a statistically secure protocol family.
5. The final step is to convert the protocol to one in which both Alice and Bob consult only a constant number of copies of C (Lemma 8).
   This is easiest to see from the cryptographic perspective: If we simply remove the copies of C that Alice ignores, and require Bob to locally sample his side of C for those copies from the marginal distribution, we obtain a protocol that is at least as secure as the original one. Note that this transformation results in a protocol that has only a constant number of copies of C, but does require Bob to be randomized. We can determinize this protocol again (increasing the error in a bounded manner) to obtain a statistically secure, deterministic protocol using only a constant number of copies of C.
   Finally, we note that there are only finitely many such protocols, and hence, to form a statistically secure protocol family (with error that approaches 0), at least one of those protocols should have perfect security.

A Counterexample. Before proceeding further, we point out an apparent contradiction to the above claim: Consider C to be a uniformly random bit (both Alice and Bob get the same bit) and D to be a bit that is 0 with probability,

say, 1/3. Now, there is a statistical SNIR from D to C, by sampling more and more uniform bits from C to sample from D with increasingly better accuracy. However, 3 not being a power of 2 prevents a perfectly secure protocol from existing.

The reason the statistical-to-perfect security argument above breaks down in this case is that the argument requires C to have *no common information*. Common information in a correlation refers to a value that Alice and Bob can always agree on, when each of them is given only a sample from their side of the correlation. The restriction that C has no common information comes at the very first step of the sequence of arguments above, where we interpreted $\boldsymbol{F}_C$ as a "generalized Fourier transform."

Nevertheless, using a result from [1], we can handle correlations with common information. Suppose C has common information – say, w.l.o.g., it is in the form of sampling an index $i \in [k]$ according to some fixed distribution, and then sampling from a correlation C_i, where $C_1, \cdots, C_k$ are correlations without common information, and over disjoint alphabets (called the *components* of C). Then, it is not hard to see (from the cryptographic perspective) that for the purposes of statistically secure SNIR, C is equivalent to a correlation C' which samples $(k+1)$-tuples $(x_0, \cdots, x_k)$ for Alice and $(y_0, \cdots, y_k)$ for Bob, where $x_0 = y_0$ is a single uniform random bit, and for $i > 0$, (x_i, y_i) is a sample from C_i. That is, $C' = C_{\text{coin}} \otimes C^{\|}$, where C_{coin} is the uniform common coin, and $C^{\|}$ is a correlation without any common information. Then, using a result from [1], it follows that a correlation D has an SNIR to C' iff each of the components of D has an SNIR to $C^{\|}$. Since $C^{\|}$ has no common information this can be tested using the statistical-to-perfect security argument, as discussed above.

With this we obtain an algorithm that can decide the existence of SNIR for any pair of source and target correlations. This is detailed in Section 4.2.

New Necessary Conditions and an Example of Interest. Despite its general and fundamental nature, our decidability result is practically unsatisfactory, as the underlying algorithm is hugely inefficient: it involves a brute-force search over a finite but large space of protocols. An important focus in prior work on SNIR has been to derive simpler necessary conditions for an SNIR to exist. In particular, in [1] it was shown that for there to be an SNIR from D to C, the *singular values of the correlation operator corresponding to D should all appear as singular values of that corresponding to $C^{\otimes \ell}$* for some ℓ. Such a result can be used to "manually" infer impossibility results for examples of interest.

While the results from [1,2,18] covered several cryptographically interesting source-target pairs, they also left out some. For instance, it was not known whether one can reduce the correlation D corresponding to random $\binom{2}{1}$ bit-OT to the correlation C corresponding to Rabin OT – i.e., an erasure channel with erasure probability 0.5, for uniformly random input bit. The results in [1] did not cover this example, as the non-zero singular values of the correlation operator corresponding to D, namely 1 and $\frac{1}{\sqrt{2}}$ also happen to be those associated with

C.[5] The results in [2,18] also do not cover the case of the target correlation being $\binom{2}{1}$-OT.

Our statistical-to-perfect security result for SNIR plugs this gap easily: It is easy to see that there is no perfectly secure SNIR from D to $C^{\otimes \ell}$ for any ℓ, and by our result, the impossibility extends to statistical security. Indeed, this readily generalizes to a broader class of target-source correlation pairs, as captured in Lemma 10.

Overview of the Proof of the Junta Theorem. In Section 5, we prove the version of the junta theorem mentioned above. We closely follow a recent proof of the Kindler-Safra junta theorem [21,20] by Filmus [9], making several suitable generalizations based on results (and exercises) in [25]. Compared to the statement proven in [9], the main difference is that we do not restrict it to functions over the domain $\{0,1\}^n$.

The theorem we seek to prove (roughly) states the following: Suppose $\boldsymbol{f} : \Omega^n \to \mathcal{T} \subseteq \mathbb{R}$ is an approximately degree d function as mentioned above, with the higher degree components $\boldsymbol{f}^{>d}$ having only ϵ energy (above we defined the degree of a function w.r.t. a generalized Fourier transform, but it is in fact a basis-invariant quantity; it however does depend on the distribution $\boldsymbol{\pi}$ over Ω w.r.t. which the fourier basis is defined). Then there is a degree d function $\boldsymbol{h} : \Omega^n \to \mathcal{T}$ that is in fact a function of only $O(1)$ of its n inputs (the hidden constants depending on $\mathcal{T}, d$ and $\boldsymbol{\pi}$, and not on $\boldsymbol{f}$ or n), and on all but $O(\epsilon)$ fraction of the domain Ω^n (as measured using the distribution $\boldsymbol{\pi}^{\otimes n}$) $\boldsymbol{h}$ equals $\boldsymbol{f}$.

Below we expose the high-level structure of the proof.

⋆ For each coordinate i, we will show that its *influence* on $\boldsymbol{f}^{\leq d}$ is either $O(\epsilon)$ or $\Omega(1)$ where the constants depend on $\mathcal{T}, d, \lambda$.

• But the degree bound on $\boldsymbol{f}^{\leq d}$ implies that the total influence can be at most $d\|\boldsymbol{f}\|^2 = O(1)$. So at most $O(1)$ coordinates i can have $\Omega(1)$ influence on $\boldsymbol{f}^{\leq d}$.

⋆ Outside of these $O(1)$ coordinates, the function is shown to have low *variance*.

• Then averaging over those coordinates gives a function $\boldsymbol{g}$ that does not depend on those coordinates, and is a good approximation in the sense that the function $\boldsymbol{f} - \boldsymbol{g}$ has small energy.

• This is not quite in the form of the approximation we desire, since we would like to ensure that $\Pr_{x \leftarrow \boldsymbol{\pi}^{\otimes n}}[\boldsymbol{f}(x) \neq \boldsymbol{g}(x)]$ is small. This is ensured by considering a function $\boldsymbol{h}$ which rounds off $\boldsymbol{g}$ to use values in the set $\mathcal{T}$. Since the variance is small, this can be done in a way that keeps the energy of $\boldsymbol{f} - \boldsymbol{h}$ still small. Now, since $\mathcal{T}$ is a finite set, there is an $\Omega(1)$ lower bound on $|\boldsymbol{f}(x) - \boldsymbol{h}(x)|$ whenever $\boldsymbol{f}(x) \neq \boldsymbol{h}(x)$.

Above, apart from the starred items, the others rely on mostly elementary arguments. The first starred step relies on a *hypercontractivity* result. It is applied to the so-called Laplacians of the function w.r.t. each coordinate, to prove

[5] There were additional interesting examples that the singular value condition did not cover, but were handled in [1] using another necessary condition – called the Mirroring Lemma. But the above example evaded those approaches as well.

a dichotomy for each coordinate, between having very low influence and high influence. For the second starred step, a result called the *Invariance Principle* is invoked to translate the low influence of the variables to low variance. One of our technical contributions is to flesh out an appropriately generalized version of the invariance principle (Lemma 15), to complete this step.

3 Preliminaries

Notation. We extensively employ linear algebraic notation, carefully adapted to allow precise expression of Fourier analytic definitions in terms of matrix multiplications. Some of the following is borrowed from [1].

We write $[n]$ to denote $\{1, \cdots, n\}$ and $[\![n]\!]$ to denote the set $\{0, \ldots, n-1\}$. $\mathbb{R}$ stands for the set of real numbers. Throughout the paper, all sets defined are finite. We typically denote such sets as $\mathcal{X}, \mathcal{Y}$, and so on, and a member of $\mathcal{X}$ is denoted as x.

Vectors and matrices are indexed by elements of finite sets. For a set $\mathcal{X}$, we write $\boldsymbol{v} \in \mathbb{R}^{\mathcal{X}}$ to mean that $\boldsymbol{v}$ is a column vector with real numbers indexed by the elements of $\mathcal{X}$ as its entries (i.e., $\boldsymbol{v}$ is, essentially, a function $\boldsymbol{v} : \mathcal{X} \to \mathbb{R}$); we will often refer to $\boldsymbol{v}$ as an $\mathcal{X}$ dimensional vector. For an $\mathcal{X}$ dimensional vector $\boldsymbol{v}$, the entry at the position x is denoted by $(\boldsymbol{v})_x$. Similarly, for sets $\mathcal{X}$ and $\mathcal{Y}$, we write $H \in \mathbb{R}^{\mathcal{X} \times \mathcal{Y}}$ to mean that H is an $\mathcal{X} \times \mathcal{Y}$ dimensional matrix with real numbers as entries. The row of H indexed by x and the column indexed by y are denoted as $(H)_{(x,\cdot)}$ and $(H)_{(\cdot,y)}$, respectively, and the element indexed by (x, y) is denoted as $(H)_{(x,y)}$. The transpose is denoted by $H^\intercal$. Finally, $|H|$ denotes the absolute value of H, *i.e.,* $(|H|)_{(i,j)} = |(H)_{(i,j)}|$, for all $i \in [m]$ and $j \in [n]$. The parentheses are removed whenever there is no scope for confusion and the vector/matrix itself is subscripted; *i.e.,* $(\boldsymbol{v})_x, (H)_{(\cdot,x)}$ and $(H)_{(x,y)}$ are simplified to $\boldsymbol{v}_x, H_{(\cdot,x)}$ and $H_{(x,y)}$, respectively.

A column vector over the set $\mathcal{X}$ with all elements being 1 (resp. 0) is denoted by $\mathbf{1}^{\mathcal{X}}$ (resp. $\mathbf{0}^{\mathcal{X}}$). For $x \in \mathcal{X}$, $\boldsymbol{\xi}_x^{\mathcal{X}}$ denotes the $\mathcal{X}$ dimensional unit vector along the 'direction x'; i.e., $\left(\boldsymbol{\xi}_x^{\mathcal{X}}\right)_x = 1$ and $\left(\boldsymbol{\xi}_x^{\mathcal{X}}\right)_{x'} = 0$ for all $x' \neq x$. The superscript is dropped when there is no scope for confusion regarding the dimension of these vectors.

We write $O_D(\epsilon)$ to denote an upper bound of the form $f(D) \cdot \epsilon$, for some fixed non-negative function f.

Probability. We only consider distributions over finite sets in this paper. A distribution over $\mathcal{X}$ is completely described by a *distribution vector* $\boldsymbol{\pi} \in \mathbb{R}_{\geq 0}^{\mathcal{X}}$ such that $\sum_{x \in \mathcal{X}} \boldsymbol{\pi}_x = 1$, and the probability of $x \in \mathcal{X}$ is $\boldsymbol{\pi}_x$. Sampling x according to the distribution $\boldsymbol{\pi}$ independent of all previously defined random variables is denoted by $x \sim \boldsymbol{\pi}$. The statistical distance or total variation distance between two distributions $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ over the same set $\mathcal{X}$ is denoted by $\mathrm{SD}(\boldsymbol{\pi}, \boldsymbol{\pi}')$, and is computed as

$$\mathrm{SD}(\boldsymbol{\pi}, \boldsymbol{\pi}') = \frac{1}{2} \sum_{x \in \mathcal{X}} |\boldsymbol{\pi}_x - \boldsymbol{\pi}'_x|.$$

Throughout this paper, we are interested in correlations, which are joint distributions over the product of two finite sets. A correlation over $\mathcal{X} \times \mathcal{Y}$ is completely described by a joint distribution matrix $H \in \mathbb{R}_{\geq 0}^{\mathcal{X}\times\mathcal{Y}}$ such that

$$\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} H_{(x,y)} = 1.$$

In the sequel, we will always refer to a correlation by its joint distribution matrix. The left marginal of H or the marginal distribution of the first coordinate of the joint distribution is given by the distribution vector $H\mathbf{1}$; the right marginal of H is given by the row vector $\mathbf{1}^\intercal H$ or equivalently the column vector $H^\intercal \mathbf{1}$. We write $(X, Y) \sim H$ to imply that the random variables (X, Y) are distributed according to the distribution H; i.e., $\mathrm{P}_{X,Y}(x, y) = H_{(x,y)}$ for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

When we say Alice and Bob receive a correlation (X, Y), we mean Alice and Bob receive random variables X and Y, respectively. The objective of non-interactive secure reductions is for Alice and Bob to *securely realize* a desired correlation among themselves using (potentially many copies) of the correlation at hand without communicating with each other.

Definition 1 (Norms)**.** For an $\mathcal{X} \times \mathcal{Y}$ dimensional matrix H, 1-norm of the matrix, denoted by $\|H\|_{1,1}$, is the sum of the absolute values of all elements in H, *i.e.,*

$$\|H\|_{1,1} = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} |H_{(x,y)}| = (\mathbf{1}^{\mathcal{X}})^\intercal |H| \mathbf{1}^{\mathcal{Y}}.$$

The 2-norm of an n dimensional vector $\boldsymbol{v}$ is defined as $\|\boldsymbol{v}\|_2 = \left(\sum_{i\in[n]} \boldsymbol{v}_i^2\right)^{\frac{1}{2}}$. ◁

Definition 2. A matrix $H \in \mathbb{R}_{\geq 0}^{\mathcal{X}\times\mathcal{Y}}$ with non-negative entries is said to be *stochastic* if $H\mathbf{1}^{\mathcal{Y}} = \mathbf{1}^{\mathcal{X}}$. A stochastic matrix in which every entry is either 0 or 1 is called a *deterministic* stochastic matrix or simply a deterministic matrix. ◁

Definition 3. For a (row or column) vector $\boldsymbol{v} \in \mathbb{R}^{\mathcal{X}}$, we define $\mathrm{diag}(\boldsymbol{v}) \in \mathbb{R}^{\mathcal{X}\times\mathcal{X}}$ as the diagonal matrix given by

$$(\,\mathrm{diag}(\boldsymbol{v}))_{(x,x')} = \begin{cases} \boldsymbol{v}_x & \text{if } x = x', \\ 0 & \text{otherwise.} \end{cases}$$

For $H \in \mathbb{R}^{\mathcal{X}\times\mathcal{Y}}$, we define $\boldsymbol{\Delta}_H$ as the $\mathcal{Y} \times \mathcal{Y}$ dimensional diagonal matrix

$$\boldsymbol{\Delta}_H = \mathrm{diag}(\mathbf{1}^\intercal H). \qquad \triangleleft$$

Tensor product. When $G \in \mathbb{R}^{\mathcal{X}\times\mathcal{Y}}$ and $H \in \mathbb{R}^{\mathcal{R}\times\mathcal{S}}$, *tensor (Kronecker) product* of G and H, denoted as $G \otimes H$, is an $(\mathcal{X} \times \mathcal{R}) \times (\mathcal{Y} \times \mathcal{S})$ dimensional matrix such that, for all $(x, r) \in \mathcal{X} \times \mathcal{R}$ and $(y, s) \in \mathcal{Y} \times \mathcal{S}$,

$$(G \otimes H)_{((x,r),(y,s))} = G_{(x,y)} \cdot H_{(r,s)}.$$

When G and H are joint distribution matrices, the distribution matrix of the product distribution– independent draws from distributions G and H–is $G \otimes H$. Hence, the distribution of $n \in \mathbb{N}$ i.i.d. samples drawn from a correlation with distribution matrix G is described by the joint distribution matrix $G^{\otimes n}$. We will use the following identity which follows from the definitions of matrix multiplication and tensor product.

Claim 1. *For matrices G, H, G', H', $(GH) \otimes (G'H') = (G \otimes G')(H \otimes H')$. In particular, for $t \in \mathbb{N}$, $(GH)^{\otimes t} = G^{\otimes t} H^{\otimes t}$.*

Definition 4. A correlation H over $\mathcal{X} \times \mathcal{Y}$ is said to be *redundant* if there exist distinct $x, x' \in \mathcal{X}$ and $c \in \mathbb{R}_{\geq 0}$ such that $H_{(x,\cdot)} = c \cdot H_{(x',\cdot)}$ or there exist $y, y' \in \mathcal{Y}$ and $c \in \mathbb{R}_{\geq 0}$ such that $H_{(\cdot,y)} = c \cdot H_{(\cdot,y')}$. ◁

By this definition, both the marginal distributions of a non-redundant distribution have full support since an all zero column (or row) is trivially a scalar multiple of any other column (or row). For a redundant correlation, we define its non-redundant *core* as the correlation obtained by collapsing redundant symbols (on both sides) to their equivalence classes.

A correlation is said to have non-zero common information if two parties can agree on a bit with non-trivial entropy using the correlation without communicating. We formally define this notion below:

Definition 5. Correlation H over $\mathcal{X} \times \mathcal{Y}$ has common-information if there exist functions $f : \mathcal{X} \to \{0,1\}$ and $g : \mathcal{Y} \to \{0,1\}$ such that, when $(X, Y) \sim H$,

$$\mathrm{P}[f(X) = g(Y)] = 1 \text{ and } 0 < \mathrm{P}[f(X) = 0] < 1.$$

H has non-zero common information if and only if there exist $\emptyset \subset \mathcal{X}_0 \subset \mathcal{X}$ and $\emptyset \subset \mathcal{Y}_0 \subset \mathcal{Y}$, joint distribution matrices H_0 and H_1 over $\mathcal{X}_0 \times \mathcal{Y}_0$ and $(\mathcal{X} \setminus \mathcal{X}_0) \times (\mathcal{Y} \setminus \mathcal{Y}_0)$, respectively, and $0 < \alpha < 1$ such that H can be written as

$$H = \begin{bmatrix} \alpha H_0 & \mathbf{0} \\ \mathbf{0} & (1-\alpha) H_1 \end{bmatrix}.$$

A correlation that does not admit such a decomposition is said to be common-information free. ◁

3.1 Generalized Fourier Transform

Let $\boldsymbol{\pi} \in \mathbb{R}^{\Omega}_{\geq 0}$ be a distribution over a finite set Ω. We consider the normed vector space $L^2(\Omega, \boldsymbol{\pi})$. The elements of this space are $\boldsymbol{v} \in \mathbb{R}^{\Omega}$ – i.e., real-valued vectors indexed by Ω, or equivalently, functions $\boldsymbol{v} : \Omega \to \mathbb{R}$. The inner product between two such vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{\Omega}$, denoted by $\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\boldsymbol{\pi}}$, is given by

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{\boldsymbol{\pi}} = \sum_{\omega \in \Omega} \boldsymbol{\pi}_{\omega} \cdot \boldsymbol{u}_{\omega} \cdot \boldsymbol{v}_{\omega}.$$

A set of vectors $\{\boldsymbol{\gamma}_{\alpha} \in \mathbb{R}^{\Omega} : \alpha \in [\![|\Omega|]\!]\}$ constitute a *Fourier basis* of $L^2(\Omega, \boldsymbol{\pi})$ if

1. $\boldsymbol{\gamma}_0$ is the constant function; i.e., $\boldsymbol{\gamma}_0 = \mathbf{1}^\Omega$.
2. For all $\alpha \in [\![|\Omega|]\!]$, $\boldsymbol{\gamma}_\alpha$ is unit norm; i.e., $\langle \boldsymbol{\gamma}_\alpha, \boldsymbol{\gamma}_\alpha \rangle_{\boldsymbol{\pi}} = 1$.
3. For all distinct $\alpha, \alpha' \in [\![|\Omega|]\!]$, $\boldsymbol{\gamma}_\alpha$ is orthogonal to $\boldsymbol{\gamma}_{\alpha'}$; i.e., $\langle \boldsymbol{\gamma}_\alpha, \boldsymbol{\gamma}_{\alpha'} \rangle_{\boldsymbol{\pi}} = 0$.

We shall identify the set Γ with a matrix $\Gamma \in \mathbb{R}^{\Omega \times [\![|\Omega|]\!]}$ with $\boldsymbol{\gamma}_\alpha$ as its rows: i.e., $\Gamma_{(\alpha,\cdot)} = \boldsymbol{\gamma}_\alpha^\intercal$.

Definition 6. The *generalized Fourier transform* w.r.t. a Fourier basis $\Gamma = \{\boldsymbol{\gamma}_\alpha \in \mathbb{R}^\Omega : \alpha \in [\![|\Omega|]\!]\}$ of $L^2(\Omega, \boldsymbol{\pi})$ is a linear operation that maps a vector $\boldsymbol{v} \in \mathbb{R}^\Omega$ to $\widehat{\boldsymbol{v}} \in \mathbb{R}^{[\![|\Omega|]\!]}$ such that

$$\widehat{\boldsymbol{v}}_\alpha = \langle \boldsymbol{\gamma}_\alpha, \boldsymbol{v} \rangle_{\boldsymbol{\pi}} \text{ for all } \alpha \in [0, |\Omega|).$$

The linear operator $\boldsymbol{F}$ that effects this transformation – i.e., $\boldsymbol{F} \in \mathbb{R}^{[\![|\Omega|]\!] \times \Omega}$ such that for all $\boldsymbol{v} \in \mathbb{R}^\Omega$, $\boldsymbol{F}\boldsymbol{v} = \widehat{\boldsymbol{v}}$ – is called the *Fourier transform operator* for Γ in $L^2(\Omega, \boldsymbol{\pi})$. ◁

Proposition 1. *Suppose* $\Gamma \in \mathbb{R}^{\Omega \times [\![|\Omega|]\!]}$ *and its rows* $\boldsymbol{\gamma}_\alpha := \Gamma_{(\alpha,\cdot)}$ *form a Fourier basis of* $L^2(\Omega, \boldsymbol{\pi})$. *Then, the matrix* $\boldsymbol{F} \in \mathbb{R}^{[\![|\Omega|]\!] \times \Omega}$ *defined as* $\boldsymbol{F} = \Gamma \operatorname{diag}(\boldsymbol{\pi})$ *is the Fourier transform operator for* Γ.

Proof: For all $\boldsymbol{v} \in L^2(\Omega, \boldsymbol{\pi})$ and $\alpha \in [\![|\Omega|]\!]$,

$$(\boldsymbol{F}\boldsymbol{v})_\alpha = \boldsymbol{F}_{(\alpha,\cdot)}\boldsymbol{v} = \boldsymbol{\gamma}_\alpha^\intercal \operatorname{diag}(\boldsymbol{\pi})\boldsymbol{v} = \sum_{\omega \in \Omega} \boldsymbol{\pi}_\omega \, (\boldsymbol{\gamma}_\alpha)_\omega \, \boldsymbol{v}_\omega = \langle \boldsymbol{\gamma}_\alpha, \boldsymbol{v} \rangle_{\boldsymbol{\pi}} = \widehat{\boldsymbol{v}}_\alpha.$$

Thus, $\boldsymbol{F}\boldsymbol{v} = \widehat{\boldsymbol{v}}$. □

Energy and Degree. The energy of a vector $\boldsymbol{v} \in L^2(\Omega, \boldsymbol{\pi})$ is defined as

$$\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle_{\boldsymbol{\pi}} = \sum_{\omega \in \Omega} \boldsymbol{\pi}_\omega \cdot \boldsymbol{v}_\omega \cdot \boldsymbol{v}_\omega.$$

Parseval's theorem refers to the following alternative for computing the energy of $\boldsymbol{v} \in L^2(\Omega, \boldsymbol{\pi})$:

$$\|\boldsymbol{v}\|^2 = \sum_{\alpha \in [\![|\Omega|]\!]} \widehat{\boldsymbol{v}}_\alpha \cdot \widehat{\boldsymbol{v}}_\alpha.$$

For any Fourier basis $\Gamma = \{\boldsymbol{\gamma}_\alpha \mid \alpha \in [\![|\Omega|]\!]\}$ over $L^2(\Omega, \boldsymbol{\pi})$ and for any $n \in \mathbb{N}$, the following is a generalized Fourier basis over $L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$:

$$\{\boldsymbol{\gamma}_{\alpha_1} \otimes \boldsymbol{\gamma}_{\alpha_2} \otimes \ldots \otimes \boldsymbol{\gamma}_{\alpha_n} : \alpha_i \in [\![|\Omega|]\!] \text{ for all } i \in [n]\}.$$

For any $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in [\![|\Omega|]\!]^n$, degree of $\boldsymbol{\alpha}$ denoted by $\deg(\boldsymbol{\alpha})$ is given by

$$\deg(\boldsymbol{\alpha}) = |\{i \in [n] : \alpha_i \neq 0\}|.$$

We can project a vector to its low-degree and high-degree components. For a vector $\boldsymbol{v} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ and $d \in [0, n]$,

$$\boldsymbol{v}^{\leq d} = \sum_{\boldsymbol{\alpha} \in [\![|\Omega|]\!]^n : \deg(\boldsymbol{\alpha}) \leq d} \widehat{\boldsymbol{v}}_\alpha \cdot \boldsymbol{\gamma}_{\boldsymbol{\alpha}}$$

$$\text{and } \boldsymbol{v}^{>d} = \sum_{\boldsymbol{\alpha} \in [\![|\Omega|]\!]^n : \deg(\boldsymbol{\alpha}) > d} \widehat{\boldsymbol{v}}_\alpha \cdot \boldsymbol{\gamma}_{\boldsymbol{\alpha}}$$

Even though we have written the low-degree component $\boldsymbol{v}^{\leq d}$ in terms of the basis vectors, it should be noted that this is actually the same for all bases [25]. Furthermore, we say that a vector $\boldsymbol{v}$ *has degree* d when $\boldsymbol{v} = \boldsymbol{v}^{\leq d}$.

3.2 Secure Non-Interactive Reduction

In this section, we formally define SNIR and import a set of statements established in [1] that we will need to prove our main result. The definitions and statement of theorem have been adapted to the current notations, but are otherwise imported verbatim from the older work.

Definition 7. Let C and D be correlations over $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{R} \times \mathcal{S}$, respectively. For any $\epsilon \geq 0$, an *ϵ-secure non-interactive reduction* (ϵ-SNIR) from D to C is a pair of probabilistic algorithms $\mathfrak{A} : \mathcal{X} \to \mathcal{R}$ and $\mathfrak{B} : \mathcal{Y} \to \mathcal{S}$ such that, when $(X, Y) \sim C$ and $(R, S) \sim D$,

ϵ-Correctness:

$$\mathrm{SD}\left((\mathfrak{A}(X), \mathfrak{B}(Y)), (R, S)\right) \leq \epsilon. \tag{3}$$

ϵ-Security: There exist a pair of probabilistic algorithms, $\mathsf{Sim}_A : \mathcal{R} \to \mathcal{X}$ and $\mathsf{Sim}_B : \mathcal{S} \to \mathcal{Y}$ such that,

$$\mathrm{SD}\left((X, \mathfrak{B}(Y)), (\mathsf{Sim}_A(R), S)\right) \leq \epsilon, \tag{4}$$

$$\mathrm{SD}\left((\mathfrak{A}(X), Y), (R, \mathsf{Sim}_B(S))\right) \leq \epsilon. \tag{5}$$

0-SNIR is alternatively called a perfect SNIR. ◁

Definition 8. Let C and D be correlations over $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{R} \times \mathcal{S}$, respectively. D is said to have a statistical SNIR to C if, for all $\epsilon > 0$, there exists a sufficiently large n for which, D has an ϵ-SNIR to $C^{\otimes n}$. ◁

Suppose $(\mathfrak{A}, \mathfrak{B})$ is an SNIR from correlation D distributed over $\mathcal{U} \times \mathcal{V}$ to C distributed over $\mathcal{X} \times \mathcal{Y}$. The probabilistic algorithm $\mathfrak{A}$ employed by Alice can be equivalently thought of as a $\mathcal{X} \times \mathcal{Y}$ dimensional stochastic matrix A with $A_{(x,u)} = \mathrm{P}_{\mathfrak{A}}(u|x)$ for each x, u. Similarly, probabilistic algorithm $\mathfrak{B}$ can be thought of as a $\mathcal{Y} \times \mathcal{U}$ dimensional stochastic matrix B. The simulators Sim_A and Sim_B can also be equivalently thought of as stochastic matrices U and V of dimensions $\mathcal{U} \times \mathcal{X}$ and $\mathcal{V} \times \mathcal{Y}$, respectively. The following proposition shows how the correctness and security conditions of SNIR translates to linear algebraic constraints in terms of these stochastic matrices.

Proposition 2 ([1, Theorem 2]). *A correlation D over $\mathcal{U} \times \mathcal{V}$ has an ϵ-SNIR to a correlation C over $\mathcal{X} \times \mathcal{Y}$ if and only if there exist stochastic matrices A, B, U, and V of dimensions $\mathcal{X} \times \mathcal{U}$, $\mathcal{Y} \times \mathcal{V}$, $\mathcal{U} \times \mathcal{X}$, and $\mathcal{V} \times \mathcal{Y}$, respectively, such that*

$$\|A^\intercal C B - D\|_{1,1} \leq \epsilon \quad (6) \qquad \|A^\intercal C - DV\|_{1,1} \leq \epsilon \quad (7) \qquad \|CB - U^\intercal D\|_{1,1} \leq \epsilon. \quad (8)$$

Identifying the stochastic matrices with the probabilistic algorithms as discussed above, conditions (6), (7), and (8) can be seen to correspond to the correctness condition (3) and security conditions (4), and (5), respectively.

A (redundant) correlation has a perfect SNIR to its core and vice-versa (Lemma 5 of [1]). This leads to the following observation in [1]:

Proposition 3. *A redundant correlation D has a statistical SNIR to a correlation C iff the core of D has a statistical SNIR to C.*

Keeping this in mind, we focus on SNIR of non-redundant target correlations throughout this work.

Given a purported perfect SNIR (A, B) from D to C, one can verify it easily, thanks to the following result [1, Lemma 8]:

Proposition 4. *Let C and D be non-redundant correlations over $\mathcal{X} \times \mathcal{Y}$ and $\mathcal{U} \times \mathcal{V}$, respectively. If deterministic matrices A, B and stochastic matrices U and V satisfy (1); i.e., (A, B) is a perfect SNIR with U and V being the simulators for Alice and Bob, respectively, then*

$$V = \boldsymbol{\Delta}_D^{-1} B^\intercal \boldsymbol{\Delta}_C \qquad U = \boldsymbol{\Delta}_{D^\intercal}^{-1} A^\intercal \boldsymbol{\Delta}_{C^\intercal}. \quad (9)$$

In [1], the authors observed that the algorithms employed by Alice and Bob in a perfect SNIR (to a non-redundant target) is determinsitic. Furthermore, given a probabilistic statistical SNIR, one can construct a deterministic SNIR with a slightly worse correctness and security error. This observation is crucially used in proving our main result.

Lemma 1 ([1, Lemma 7]). *Let D be non-redundant correlation over $\mathcal{U} \times \mathcal{V}$ and C be a correlation over $\mathcal{X} \times \mathcal{Y}$. For any $\epsilon \geq 0$, if there exist stochastic matrices A, B, U and V such that*

$$\|A^\intercal C B - D\|_{1,1} \leq \epsilon \qquad \|A^\intercal C - DV\|_{1,1} \leq \epsilon \qquad \|CB - U^\intercal D\|_{1,1} \leq \epsilon$$

then there exist deterministic stochastic matrices $\bar{A}, \bar{B}$ such that,

$$\|\bar{A}^\intercal C \bar{B} - D\|_{1,1} \leq O_D\left(\sqrt{\epsilon}\right),$$
$$\|\bar{A}^\intercal C - DV\|_{1,1} \leq O_D\left(\sqrt{\epsilon}\right),$$
$$\|C\bar{B} - U^\intercal D\|_{1,1} \leq O_D\left(\sqrt{\epsilon}\right).$$

We recall the definitions coined in [1] relating to spectral protocols:

Definition 9 (Spectral decomposition of a correlation)**.** For a correlation H distributed over $\mathcal{X} \times \mathcal{Y}$, its spectral decomposition $\widetilde{H}$ is a $\mathcal{X} \times \mathcal{Y}$ dimensional matrix

$$\widetilde{H} = \boldsymbol{\Delta}_{H^\intercal}^{-\frac{1}{2}} H \boldsymbol{\Delta}_{H}^{-\frac{1}{2}}.$$

Define $\boldsymbol{\Sigma}_H$, $\boldsymbol{\Psi}_H$ and $\boldsymbol{\Phi}_H$ to be given by a canonical singular value decomposition of $\widetilde{H}$, so that $\boldsymbol{\Sigma}_H$ is an $[\![|\mathcal{X}|]\!] \times [\![|\mathcal{Y}|]\!]$ dimensional non-negative diagonal matrix with the diagonal sorted in descending order, $\boldsymbol{\Psi}_H$ and $\boldsymbol{\Phi}_H$ are unitary matrices of dimensions $[\![|\mathcal{X}|]\!] \times \mathcal{X}$ and $[\![|\mathcal{Y}|]\!] \times \mathcal{Y}$, respectively, and

$$\widetilde{H} = \boldsymbol{\Psi}_H^\intercal \boldsymbol{\Sigma}_H \boldsymbol{\Phi}_H.$$

Finally, define $\boldsymbol{F}_H = \boldsymbol{\Psi}_H \, \boldsymbol{\Delta}_{H^\intercal}^{1/2}$. ◁

The following properties of the spectral decomposition of a correlation were observed in [1].

Lemma 2 ([1, Lemma 9])**.** *Let $|\mathcal{X}| \leq |\mathcal{Y}|$ and H be a correlation over $\mathcal{X} \times \mathcal{Y}$. Then,*

(i) $1 = (\boldsymbol{\Sigma}_H)_{(0,0)} \geq (\boldsymbol{\Sigma}_H)_{(1,1)} \geq \ldots \geq (\boldsymbol{\Sigma}_H)_{(|\mathcal{X}|-1,|\mathcal{X}|-1)} \geq 0$. *Furthermore, if H is common information free, then $(\boldsymbol{\Sigma}_H)_{(1,1)} < 1$.*

(ii) For all $\lambda \in (0,1)$, there exists $\delta > 0$ such that for all $n \in \mathbb{N}$ and for all $\alpha \in [\![|\mathcal{X}|]\!]$, either $\lambda = (\boldsymbol{\Sigma}_H)_{(\alpha,\alpha)}$ or $|\lambda - (\boldsymbol{\Sigma}_H)_{(\alpha,\alpha)}| > \delta$.

Similar to the correlations, the SNIR protocol also allows a spectral decomposition, which we now define.

Definition 10 (Spectral Image of SNIR)**.** Let D be a non-redundant correlation over $\mathcal{U} \times \mathcal{V}$, and C be a correlation over $\mathcal{X} \times \mathcal{Y}$. The *spectral image* of an SNIR (A, B) from D to C is $(\widehat{A}, \widehat{B})$, where $\widehat{A}$ and $\widehat{B}$ are matrices of dimensions $[\![|\mathcal{X}|]\!] \times [\![|\mathcal{U}|]\!]$ and $[\![|\mathcal{Y}|]\!] \times [\![|\mathcal{V}|]\!]$, respectively, defined as

$$\widehat{A} = \boldsymbol{F}_C A \boldsymbol{F}_D^{-1} \qquad \widehat{B} = \boldsymbol{G}_C B \boldsymbol{G}_D^{-1}. \qquad \triangleleft$$

A crucial observation we make in this paper is that the columns of $\widehat{A}$ can be interpreted as a *generalized Fourier Transform* applied to the columns of a matrix derived from A. Loosely speaking, the following lemma in [1] shows that this Fourier spectrum is mostly concentrated on specific coefficients.

Lemma 3 ([1, Lemma 11])**.** *Suppose a non-redundant correlation D over $\mathcal{U} \times \mathcal{V}$ has a deterministic ϵ-SNIR (A, B) to C over $\mathcal{X} \times \mathcal{Y}$. Then, for all $\beta \in [\![|\mathcal{U}|]\!]$,*

$$\sum_{\substack{\alpha \in [\![|\mathcal{X}|]\!] \\ (\boldsymbol{\Sigma}_C)_{(\alpha,\alpha)} \neq (\boldsymbol{\Sigma}_D)_{(\beta,\beta)}}} \left((\boldsymbol{\Sigma}_C \boldsymbol{\Sigma}_C^\intercal)_{(\alpha,\alpha)} - (\boldsymbol{\Sigma}_D \boldsymbol{\Sigma}_D^\intercal)_{(\beta,\beta)} \right)^2 \left(\widehat{A}_{(\alpha,\beta)} \right)^2 = O_D(\epsilon)\,.$$

Finally, we import a lemma that shows that presence of common randomness does not help in SNIR.

Lemma 4 ([1, Theorem 6]). *Let* $C_w = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$ *be the 1-bit common randomness correlation. If a non-redundant common-information free correlation* D *has a statistical SNIR to* $C_w \otimes C$ *for a correlation* C*, then* D *also has a statistical SNIR to* C*.*

4 Decidability of SNIR

4.1 Statistical to Perfect Security

The crucial observation we make to show the decidability of SNIR is that a statistical reduction from D to C implies a deterministic perfect reduction from D to a constant number of copies of C. This is the main result of this section.

Theorem 1. *A non-redundant correlation* D *over* $\mathcal{U} \times \mathcal{V}$ *has a statistical SNIR to a common information free correlation* C *over* $\mathcal{X} \times \mathcal{Y}$ *if and only if, for a constant* $\ell \in \mathbb{N}$ *that depends only on* D *and* C*,* D *has a perfect SNIR to* $C^{\otimes \ell}$*.*

The proof of the theorem follows the outline presented in the technical overview. Without loss of generality, we assume that $|\mathcal{U}| \leq |\mathcal{V}|$ and focus on an ϵ-SNIR (for an arbitrary $\epsilon \geq 0$) implied by the assumption that D has a statistical SNIR to C. We make several observations about the spectral image of such an ϵ-SNIR (A, B) from D to (say) $\boldsymbol{C} = C^{\otimes n}$ as defined in Definition 10. Since $|\mathcal{U}| \leq |\mathcal{V}|$, it is sufficient to focus on Alice's spectral protocol $\widehat{A} = \boldsymbol{F}_{\boldsymbol{C}} A \boldsymbol{F}_D^{-1}$. In Lemma 6, we establish that $\boldsymbol{F}_{\boldsymbol{C}}$ is a Fourier transform operator for the normed vector space $L^2(\mathcal{X}^n, \boldsymbol{\pi})$, where $\boldsymbol{\pi} = \boldsymbol{C}\mathbf{1}$ is the marginal of $\boldsymbol{C}$ at Alice. This makes $\widehat{A}$ the Fourier transform of the "half-way spectral protocol" $A\boldsymbol{F}_D^{-1}$ with respect to this operator. In Lemma 7, we use the properties of spectral protocols established in [1]–restated here as Lemma 3–to show that the columns of $A\boldsymbol{F}_D^{-1}$ associated with non-zero singular values of D concentrate most of their energy in the lower degree coefficients. We focus on this sub-matrix of $A\boldsymbol{F}_D^{-1}$ given by $A\check{\boldsymbol{F}}_D$. Lemma 5 shows that each column of A is completely determined by the corresponding row of this sub-matrix. At this point, we appeal to a "generalized junta theorem" stated as Theorem 2 to argue that each column of $A\check{\boldsymbol{F}}_D$ can be approximated by a junta–a vector in $\mathcal{X}^n$ that depends only on a constant number of coordinates of $\mathcal{X}^n$. Since A is determined by $A\check{\boldsymbol{F}}_D$ which itself is close to a junta, A itself is close to a junta. Finally, in Lemma 8, we show that if Alice's protocol A 'almost entirely' depends only on a small subset of copies of the correlations in a sequence of SNIR protocols with progressively better security error, then there is a perfectly secure SNIR, concluding the proof.

We state the lemmas mentioned above which imply the theorem.

Lemma 5. *Let D be a non-degenerate correlation over $\mathcal{U} \times \mathcal{V}$, and let $(\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0$ if and only if $\beta < k \leq |\mathcal{U}|$. Define $\check{\boldsymbol{F}}_D \in \mathbb{R}^{\mathcal{U} \times [\![k]\!]}$ such that*

$$(\check{\boldsymbol{F}}_D)_{(\cdot,\beta)} = \left(\boldsymbol{F}_D^{-1}\right)_{(\cdot,\beta)} \qquad \forall \beta \in [\![k]\!]. \tag{10}$$

There exists a function $\phi : \mathbb{R}^{[\![k]\!]} \to \{0,1\}^{\mathcal{U}}$ such that, for any $\mathcal{R}$ and $\mathcal{R} \times \mathcal{U}$ dimensional deterministic matrix A,

$$A_{(r,\cdot)} = \phi\left(A_{(r,\cdot)} \check{\boldsymbol{F}}_D\right) \qquad \forall r \in \mathcal{R}. \tag{11}$$

Proof: For any $r \in \mathcal{R}$, the row $A_{(r,\cdot)}$ is a basis vector since A is a deterministic stochastic matrix. Fix $r \in \mathcal{R}$ and let $A_{(r,\cdot)} = \boldsymbol{\xi}_u^\intercal$ for some fixed u. Then, $A_{(r,\cdot)} \check{\boldsymbol{F}}_D = \boldsymbol{\xi}_u^\intercal \check{\boldsymbol{F}}_D = (\check{\boldsymbol{F}}_D)_{(u,\cdot)}$. Suppose all the rows of $\check{\boldsymbol{F}}_D$ are distinct, i.e., $(\check{\boldsymbol{F}}_D)_{(u_1,\cdot)} \neq (\check{\boldsymbol{F}}_D)_{(u_2,\cdot)}$ whenever $u_1 \neq u_2$. Consider the map $\phi : (\check{\boldsymbol{F}}_D)_{(u,\cdot)} \mapsto \boldsymbol{\xi}_u$ for all u (and is otherwise defined arbitrarily). Then, $\phi\left(A_{(r,\cdot)} \check{\boldsymbol{F}}_D\right) = A_{(r,\cdot)}$ for all r. Thus, there exists ϕ as required in the lemma whenever all the rows of $\check{\boldsymbol{F}}_D$ are distinct.

The proof is completed by showing that D is degenerate if there exist $u_1, u_2 \in \mathcal{U}$ such that $(\check{\boldsymbol{F}}_D)_{(u_1,\cdot)} = (\check{\boldsymbol{F}}_D)_{(u_2,\cdot)}$.

$$\begin{aligned}
\boldsymbol{\Delta}_{D^\intercal}^{-1} D &= \boldsymbol{\Delta}_{D^\intercal}^{-1} \left(\boldsymbol{\Delta}_{D^\intercal}^{1/2} \widetilde{D} \boldsymbol{\Delta}_{D^\intercal}^{1/2}\right) \\
&= \boldsymbol{\Delta}_{D^\intercal}^{-1/2} \boldsymbol{\Psi}_D^\intercal \boldsymbol{\Sigma}_D \boldsymbol{\Phi}_D \boldsymbol{\Delta}_{D^\intercal}^{1/2} \\
&= \sum_{\beta : (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0} (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} \left(\boldsymbol{F}_D^{-1}\right)_{(\cdot,\beta)} \cdot \left(\boldsymbol{\Phi}_D \boldsymbol{\Delta}_{D^\intercal}^{1/2}\right)_{(\beta,\cdot)}.
\end{aligned}$$

The last equality used the outer product expansion of $\boldsymbol{\Delta}_{D^\intercal}^{-1/2} \boldsymbol{\Psi}_D^\intercal \boldsymbol{\Sigma}_D \boldsymbol{\Phi}_D \boldsymbol{\Delta}_{D^\intercal}^{1/2}$ with respect to the diagonal matrix $\boldsymbol{\Sigma}_D$. Since $(\check{\boldsymbol{F}}_D)_{(u_1,\cdot)} = (\check{\boldsymbol{F}}_D)_{(u_2,\cdot)}$, substituting $\check{\boldsymbol{F}}_D = \boldsymbol{F}_D^{-1}$ in the above equation,

$$\begin{aligned}
\left(\boldsymbol{\Delta}_{D^\intercal}^{-1} D\right)_{(u_1,\cdot)} &= \sum_{\beta : (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0} (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} \left(\check{\boldsymbol{F}}_D\right)_{(u_1,\beta)} \cdot \left(\boldsymbol{\Phi}_D \boldsymbol{\Delta}_{D^\intercal}^{1/2}\right)_{(\beta,\cdot)} \\
&= \sum_{\beta : (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0} (\boldsymbol{\Sigma}_D)_{(\beta,\beta)} \left(\check{\boldsymbol{F}}_D\right)_{(u_2,\beta)} \cdot \left(\boldsymbol{\Phi}_D \boldsymbol{\Delta}_{D^\intercal}^{1/2}\right)_{(\beta,\cdot)} \\
&= \left(\boldsymbol{\Delta}_{D^\intercal}^{-1} D\right)_{(u_2,\cdot)}.
\end{aligned}$$

But then,

$$D_{(u_1,\cdot)} = \frac{(\boldsymbol{\Delta}_{D^\intercal})_{(u_1,u_1)}}{(\boldsymbol{\Delta}_{D^\intercal})_{(u_2,u_2)}} D_{(u_2,\cdot)};$$

hence, D is degenerate. □

Lemma 6. *Let $\boldsymbol{C}$ be the n-wise product of a correlation C over $\mathcal{X} \times \mathcal{Y}$ for some $n \in \mathbb{N}$; i.e., $\boldsymbol{C} = C^{\otimes n}$. Rows of $\Gamma = \left(\boldsymbol{\Psi}_{\boldsymbol{C}} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2}\right)$ form a generalized Fourier basis of the normed vector space $L^2(\mathcal{X}^n, \boldsymbol{\pi})$, where $\boldsymbol{\pi} = \boldsymbol{C}\mathbf{1}$.*

Proof: For any $\boldsymbol{\alpha}, \boldsymbol{\alpha}' \in [\![|\mathcal{X}|]\!]^n$,

$$\begin{aligned}\left\langle \Gamma_{(\boldsymbol{\alpha},\cdot)}, \Gamma_{(\boldsymbol{\alpha}',\cdot)} \right\rangle_{\boldsymbol{\pi}} &= \sum_{\boldsymbol{x}\in\mathcal{X}^n} \boldsymbol{\pi}_{\boldsymbol{x}}\, \Gamma_{(\boldsymbol{\alpha},\boldsymbol{x})}\, \Gamma_{(\boldsymbol{\alpha}',\boldsymbol{x})} = \Gamma_{(\boldsymbol{\alpha},\cdot)}\, \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}\, (\Gamma^\intercal)_{(\cdot,\boldsymbol{\alpha}')} \\ &= \left(\boldsymbol{\Psi}_{\boldsymbol{C}} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2}\right)_{(\boldsymbol{\alpha},\cdot)} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal} \left(\boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2} \boldsymbol{\Psi}_{\boldsymbol{C}}^\intercal\right)_{(\cdot,\boldsymbol{\alpha}')} \\ &= \left(\left(\boldsymbol{\Psi}_{\boldsymbol{C}} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2}\right) \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal} \left(\boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2} \boldsymbol{\Psi}_{\boldsymbol{C}}^\intercal\right)\right)_{(\alpha,\alpha')} \\ &= \begin{cases} 1 \text{ if } \boldsymbol{\alpha} = \boldsymbol{\alpha}', \\ 0 \text{ if } \boldsymbol{\alpha} \neq \boldsymbol{\alpha}'. \end{cases}\end{aligned}$$

We now show that $\Gamma_{(\mathbf{0},\cdot)} = \mathbf{1}^\intercal$. In the context of Lemma 1 in [1] and its proof, we observe that $\mathcal{L}(G_{\boldsymbol{C}})$ has an eigenvalue 0 corresponding to eigenvector $[(\mathbf{1}^\intercal \cdot \boldsymbol{C}^\intercal)^{1/2}, (\mathbf{1}^\intercal \cdot \boldsymbol{C})^{1/2}]^\intercal$, and thus $\left((\mathbf{1}^\intercal \cdot \boldsymbol{C}^\intercal)^{1/2}\right)^\intercal = (\boldsymbol{C} \cdot \mathbf{1})^{1/2}$ is a left singular vector of $\widetilde{C}$ corresponding to singular value 1. Assuming $\boldsymbol{C}$ has a single connected component, we get that the multiplicity of 1 in $\boldsymbol{\Lambda}_{\boldsymbol{C}}$ is only one, and this is the maximum singular value as well, implying $(\boldsymbol{\Psi}_{\boldsymbol{C}}^\intercal)_{(\cdot,\mathbf{0})} = (\boldsymbol{C} \cdot \mathbf{1})^{1/2}$, i.e., $(\boldsymbol{\Psi}_{\boldsymbol{C}})_{(\mathbf{0},\cdot)} = (\mathbf{1}^\intercal \cdot \boldsymbol{C}^\intercal)^{1/2}$. We then have

$$\Gamma_{(\mathbf{0},\cdot)} = \left(\boldsymbol{\Psi}_{\boldsymbol{C}} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2}\right)_{(\mathbf{0},\cdot)} = (\boldsymbol{\Psi}_{\boldsymbol{C}})_{(\mathbf{0},\cdot)}\, \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2} = (\mathbf{1}^\intercal \cdot \boldsymbol{C}^\intercal)^{1/2} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2} = \mathbf{1}^\intercal.$$

Thus, the rows of Γ form a generalized Fourier basis of $L^2(\mathcal{X}^n, \boldsymbol{\pi})$. Finally, by Proposition 1, $\Gamma \operatorname{diag}(\boldsymbol{\pi}) = \left(\boldsymbol{\Psi}_{\boldsymbol{C}} \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{-1/2}\right) \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal} = \boldsymbol{F}_{\boldsymbol{C}}$ is a Fourier transform operator for Γ in $L^2(\mathcal{X}^n, \boldsymbol{\pi})$.

□

Lemma 7. *Let D be a non-redundant correlation over $\mathcal{U} \times \mathcal{V}$ and $\boldsymbol{C} = C^{\otimes n}$ be the n-wise product of a common information free correlation C over $\mathcal{X} \times \mathcal{Y}$. If (A, B) is a deterministic ϵ-SNIR from D to $\boldsymbol{C}$, then there exists a number $d \in \mathbb{N}$ that depends only on C and D (and not on n) such that, for each β such that $(\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0$, the vector $\boldsymbol{a}_\beta = \left(A \boldsymbol{F}_D^{-1}\right)_{(\cdot,\beta)} \in L^2(\mathcal{X}^n, \boldsymbol{\pi})$, where $\boldsymbol{\pi} = \boldsymbol{C}\mathbf{1}$, satisfies $\|\boldsymbol{a}_\beta^{>d}\|^2 \leq O_D(\sqrt{\epsilon})$.*

Proof: Consider the Fourier basis Γ of $L^2(\mathcal{X}^n, \boldsymbol{\pi})$ described in Lemma 6 and its Fourier transform operator $\boldsymbol{F}_{\boldsymbol{C}} = \boldsymbol{\Psi}_{\boldsymbol{C}}\, \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}^{1/2}$. By Definition 10, Fourier transform of $\boldsymbol{a}_\beta$ for any β w.r.t. Γ is given by

$$\widehat{\boldsymbol{a}_\beta} = \boldsymbol{F}_{\boldsymbol{C}} \boldsymbol{a}_\beta = \boldsymbol{F}_{\boldsymbol{C}} \left(A \boldsymbol{F}_D^{-1}\right)_{(\cdot,\beta)} = \widehat{A}_{(\cdot,\beta)}. \tag{12}$$

By Lemma 3, for $\boldsymbol{\alpha} \in [\![|\mathcal{X}|]\!]^n$,

$$\sum_{\substack{\boldsymbol{\alpha}\in[\![|\mathcal{X}|]\!]^n \\ (\boldsymbol{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\boldsymbol{\Sigma}_D)_{(\beta,\beta)}}} \left((\boldsymbol{\Sigma}_{\boldsymbol{C}} \boldsymbol{\Sigma}_{\boldsymbol{C}}^\intercal)_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} - (\boldsymbol{\Sigma}_D \boldsymbol{\Sigma}_D^\intercal)_{(\beta,\beta)}\right)^2 \left(\widehat{A}_{(\boldsymbol{\alpha},\beta)}\right)^2 = O_D(\epsilon)\,. \tag{13}$$

By Lemma 2 (ii), $\exists \delta > 0$ such that, for any $\boldsymbol{\alpha}$, β, s.t. $(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\mathbf{\Sigma}_D)_{(\beta,\beta)}$ and $(\mathbf{\Sigma}_D)_{(\beta,\beta)} > 0$,

$$(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} + (\mathbf{\Sigma}_D)_{(\beta,\beta)} \geq \left| (\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} - (\mathbf{\Sigma}_D)_{(\beta,\beta)} \right| \geq \delta.$$

Using this in (13), for any β s.t. $(\mathbf{\Sigma}_D)_{(\beta,\beta)} > 0$,

$$\begin{aligned} O_D(\epsilon) &= \sum_{\boldsymbol{\alpha}:(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\mathbf{\Sigma}_D)_{(\beta,\beta)}} \left((\mathbf{\Sigma}_{\boldsymbol{C}})^2_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} - (\mathbf{\Sigma}_D)^2_{(\beta,\beta)} \right)^2 \left(\widehat{A}_{(\boldsymbol{\alpha},\beta)} \right)^2 \\ &\geq \sum_{\boldsymbol{\alpha}:(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\mathbf{\Sigma}_D)_{(\beta,\beta)}} \delta^2 \left(\widehat{A}_{(\boldsymbol{\alpha},\beta)} \right)^2. \end{aligned}$$

If there exists $d \in \mathbb{N}$ that depends only on C and D (and not on n) such that $(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\mathbf{\Sigma}_D)_{(\beta,\beta)}$ whenever $\deg(\boldsymbol{\alpha}) > d$, by the above bound and (12),

$$\sum_{\boldsymbol{\alpha}:\deg(\boldsymbol{\alpha})>d} (\widehat{\boldsymbol{a}_\beta})^2_{\boldsymbol{\alpha}} = \frac{1}{\delta^2} \sum_{\boldsymbol{\alpha}:\deg(\boldsymbol{\alpha})>d} \delta^2 \left(\widehat{A}_{(\boldsymbol{\alpha},\beta)} \right)^2 = O_D(\epsilon)\,.$$

Hence, it is sufficient to show that such a $d \in \mathbb{N}$ exists. By Lemma 2, when C is common-information free, there exists $\lambda < 1$ such that $(\mathbf{\Sigma}_C)_{(\alpha,\alpha)} \leq \lambda$ for all $1 \leq \alpha < |\mathcal{X}|$. Hence, for $\boldsymbol{\alpha} \in [\![|\mathcal{X}|]\!]^{[n]}$, (recalling $\mathbf{\Sigma}_{\boldsymbol{C}} = \mathbf{\Sigma}_{C^{\otimes n}} = \mathbf{\Sigma}_C^{\otimes n}$),

$$(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} = \prod_{i\in[n]} (\mathbf{\Sigma}_C)_{(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)} \leq \prod_{\substack{i\in[n] \\ \boldsymbol{\alpha}_i \neq 0}} \lambda \leq \lambda^{\deg(\boldsymbol{\alpha})}.$$

Choose d such that $(\mathbf{\Sigma}_D)_{(\beta,\beta)} \geq \lambda^d$ for all β s.t. $(\mathbf{\Sigma}_D)_{(\beta,\beta)} > 0$. Then, $(\mathbf{\Sigma}_{\boldsymbol{C}})_{(\boldsymbol{\alpha},\boldsymbol{\alpha})} \neq (\mathbf{\Sigma}_D)_{(\beta,\beta)}$ whenever $\deg(\boldsymbol{\alpha}) > d$. This concludes the proof. □

Theorem 2 (Generalized Junta Theorem). *Let $(\Omega, \boldsymbol{\pi})$ be a finite probability space, $|\Omega| = m \geq 2$, in which every outcome has probability at least λ. Let $\mathcal{T}$ be a finite set and let $d \geq 1$. If $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ is a $\mathcal{T}$-valued function such that $\|\boldsymbol{f}^{>d}\|^2 = \epsilon$, then there exists a $\mathcal{T}$-valued degree d function $\boldsymbol{h} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$, such that $\Pr[\boldsymbol{f} \neq \boldsymbol{h}] = O(\epsilon)$, and $\boldsymbol{h}$ depends on $O(1)$ coordinates.*

Lemma 8. *Let D be a non-redundant correlation over $\mathcal{U} \times \mathcal{V}$ and C be a correlation over $\mathcal{X} \times \mathcal{Y}$. Suppose, for each $i \in \mathbb{N}$, there is an ϵ_i-SNIR (A_i, B_i) from D to $C^{\otimes n_i}$ such that $\epsilon_i \to 0$ as $i \to \infty$. For each i, suppose there exists $\mathcal{S}_i \subset [n_i], |\mathcal{S}_i| = \ell$ and a deterministic matrix $\tilde{A}_i$ such that*

$$(\tilde{A}_i)_{(\boldsymbol{x},\cdot)} = (\tilde{A}_i)_{(\boldsymbol{x}',\cdot)} \text{ for all } \boldsymbol{x}, \boldsymbol{x}' \text{ s.t. } \boldsymbol{x}_j = \boldsymbol{x}'_j \text{ for all } j \in \mathcal{S}_i,$$

and, when $\boldsymbol{\pi}_i = C^{\otimes n_i}\mathbf{1}$, it holds that

$$\mathrm{P}_{\boldsymbol{x}\sim\boldsymbol{\pi}_i} \left[(\tilde{A}_i)_{(\boldsymbol{x},\cdot)} \neq (A_i)_{(\boldsymbol{x},\cdot)} \right] \leq \epsilon_i. \tag{14}$$

Then, D has a perfect SNIR to $C^{\otimes \ell}$.

Proof: Fix $i \in \mathbb{N}$. By Proposition 2, there are stochastic matrices U_i and V_i such that

$$\|A_i^\intercal C^{\otimes n} - DV_i\|_{1,1} \le \epsilon_i \qquad \|C^{\otimes n} B_i - U_i^\intercal D\|_{1,1} \le \epsilon_i.$$

Since i is fixed, we will drop the subscript i and denote $n_i, \boldsymbol{\pi}_i, A_i, \tilde{A}_i$ by $n, \boldsymbol{\pi}, A, \tilde{A}$, and so on. Also, we will denote $C^{\otimes n}$ by $\boldsymbol{C}$. We have,

$$\begin{aligned}\|A^\intercal \boldsymbol{C} - \tilde{A}^\intercal \boldsymbol{C}\|_{1,1} &= \mathbf{1}^\intercal |(A - \tilde{A})^\intercal \boldsymbol{C}| \mathbf{1} \\ &\le \mathbf{1}^\intercal |(A - \tilde{A})^\intercal | \boldsymbol{C}\mathbf{1} = \mathbf{1}^\intercal |(A - \tilde{A})^\intercal | \boldsymbol{\pi}.\end{aligned}$$

The final equality used the definition $\boldsymbol{\pi}_i = C^{\otimes n_i}\mathbf{1}$.

$$\begin{aligned}\mathbf{1}^\intercal |(A - \tilde{A})^\intercal | \boldsymbol{\pi} &= \sum_{\boldsymbol{x}: A_{(\boldsymbol{x},\cdot)} \neq \tilde{A}_{(\boldsymbol{x},\cdot)}} \mathbf{1}^\intercal \left(|A - \tilde{A}|_{(\boldsymbol{x},\cdot)}\right)^\intercal \boldsymbol{\pi}_{\boldsymbol{x}} \\ &\overset{(a)}{\le} 2 \sum_{\boldsymbol{x}: A_{(\boldsymbol{x},\cdot)} \neq \tilde{A}_{(\boldsymbol{x},\cdot)}} \boldsymbol{\pi}_{\boldsymbol{x}} \\ &= 2\mathrm{P}_{\boldsymbol{x} \sim \boldsymbol{\pi}} \left[(\tilde{A}_i)_{(\boldsymbol{x},\cdot)} \neq (A_i)_{(\boldsymbol{x},\cdot)}\right] \overset{(b)}{\le} 2\epsilon_i.\end{aligned}$$

Here, (a) used the fact $\mathbf{1}^\intercal \left(|A - \tilde{A}|_{(\boldsymbol{x},\cdot)}\right)^\intercal = 2$ whenever $A_{(\boldsymbol{x},\cdot)} \neq \tilde{A}_{(\boldsymbol{x},\cdot)}$ since A and $\tilde{A}$ are stochastic matrices; (b) used (14). Thus, we have argued that $\|A^\intercal \boldsymbol{C} - \tilde{A}^\intercal \boldsymbol{C}\|_{1,1} \le 2\epsilon_i$. Since $B\mathbf{1} = \mathbf{1}$, this further implies that $\|A^\intercal \boldsymbol{C} B - \tilde{A}^\intercal \boldsymbol{C} B\|_{1,1} \le 2\epsilon_i$. But then,

$$\begin{aligned}\|\tilde{A}^\intercal \boldsymbol{C} B - D\|_{1,1} &\le \|\tilde{A}^\intercal \boldsymbol{C} B - A^\intercal \boldsymbol{C} B\|_{1,1} + \|A^\intercal \boldsymbol{C} B - D\|_{1,1} \le 3\epsilon_i, \\ \|\tilde{A}^\intercal \boldsymbol{C} - DV\|_{1,1} &\le \|\tilde{A}^\intercal \boldsymbol{C} - A^\intercal \boldsymbol{C}\|_{1,1} + \|A^\intercal \boldsymbol{C} - D\|_{1,1} \le 3\epsilon_i, \\ \|\boldsymbol{C} B - U^\intercal D\|_{1,1} &\le \epsilon_i.\end{aligned}$$

Thus, $(\tilde{A}, B)$ is a $3\epsilon_i$-SNIR from D to $\boldsymbol{C}$.

From $(\tilde{A}, B)$, we derive an SNIR that uses only ℓ copies C but retains the same security guarantees as the original reduction. We argue this part from a cryptographic perspective: Consider the protocols $\tilde{\mathfrak{A}} : \mathcal{X}^n \to \mathcal{U}$ and $\mathfrak{B} : \mathcal{Y}^n \to \mathcal{V}$ corresponding to the stochastic matrices $\tilde{A}$ and B, respectively. If $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}^n$ are such that $\boldsymbol{x}_i = \boldsymbol{x}'_i$ for all $i \in \mathcal{S}$, then $\tilde{A}_{(\boldsymbol{x},\cdot)} = \tilde{A}_{(\boldsymbol{x}',\cdot)}$. Equivalently, $\tilde{\mathfrak{A}}(\boldsymbol{x})$ and $\tilde{\mathfrak{A}}(\boldsymbol{x}')$ are identically distributed for such $\boldsymbol{x}, \boldsymbol{x}'$. In other words, $\tilde{\mathfrak{A}}$ depends only on $\mathcal{X}^{\mathcal{S}}$. If we remove the copies of C that are ignored by $\mathfrak{A}$ and have $\mathfrak{B}$ sample its side of C for these copies from the marginal distribution, we obtain a protocol that depends only on $|\mathcal{S}| = \ell$ many copies of C and is at least as secure as the original SNIR. Let the deterministic protocol obtained by restricting $\tilde{\mathfrak{A}}$ to $\mathcal{X}^{\mathcal{S}}$ be called $\mathfrak{A}'$, and the stochastic (not necessarily deterministic) protocol obtained by restricting $\mathfrak{B}$ to $\mathcal{Y}^{\mathcal{S}}$ be called $\mathfrak{B}'$. Then, (A', B') is a $3\epsilon_i$-SNIR from D to $C^{\otimes \ell}$.

For each $i \in \mathbb{N}$, we constructed $3\epsilon_i$-SNIR (A'_i, B'_i) from D to $C^{\otimes \ell}$. But then, for each $i \in \mathbb{N}$, by Lemma 1, there exist deterministic matrices $\bar{A}_i \in \{0,1\}^{\mathcal{X}^\ell \times \mathcal{U}}$ and $\bar{B}_i \in \{0,1\}^{\mathcal{Y}^\ell \times \mathcal{V}}$ such that $(\bar{A}_i, \bar{B}_i)$ is a $O_D\left(\sqrt{\epsilon_i}\right)$-SNIR from D to $C^{\otimes \ell}$. Since ℓ is a constant, there exist only a finite number of choices of $\bar{A}_i \in \{0,1\}^{\mathcal{X}^\ell \times \mathcal{U}}$ and $\bar{B}_i \in \{0,1\}^{\mathcal{Y}^\ell \times \mathcal{V}}$, and hence there exist deterministic matrices A^* and B^* such that (A^*, B^*) is a perfect SNIR from D to $C^{\otimes \ell}$. □

Proof of Theorem 1. If D has a statistical SNIR to C, there is a sequence of protocols $(A_i, B_i)_{i \in \mathbb{N}}$ such that, for each $i \in \mathbb{N}$, (A_i, B_i) is an ϵ_i-SNIR from D to $C^{\otimes n_i}$ and $\epsilon_i \to 0$ as $i \to \infty$.

Fix $i \in \mathbb{N}$; we drop the subscript from $n_i, A_i, B_i, \epsilon_i$ and simply use n, A, B, ϵ instead. We denote $C^{\otimes n_i}$ by $\boldsymbol{C}$ and $\boldsymbol{C}\mathbf{1}$ by $\boldsymbol{\pi}$. Consider the normed vector space $L^2(\mathcal{X}^n, \boldsymbol{\pi})$. Suppose $(\boldsymbol{\Sigma}_D)_{(\beta,\beta)} > 0$ if and only if $\beta \in [\![k]\!] \subseteq [\![|\mathcal{U}|]\!]$. Define

$$\mathcal{T}_D = \left\{ \left(\boldsymbol{F}_D^{-1}\right)_{(u,\beta)} : u \in \mathcal{U}, \beta \in [\![k]\!] \right\}.$$

For each $\beta \in [\![k]\!]$, define $\boldsymbol{a}_\beta \in \mathcal{T}_D^{\mathcal{X}^n}$ as

$$\boldsymbol{a}_\beta = \left(A\boldsymbol{F}_D^{-1}\right)_{(\cdot,\beta)} \in L^2(\mathcal{X}^n, \boldsymbol{\pi}).$$

By Lemma 7, there exists d that depends only on D and C (and not on n) such that $\|\boldsymbol{a}_u^{>d}\|_2 \leq O_D(\epsilon)$. By Theorem 2, for each $\beta \in [\![k]\!]$, there exists $\widetilde{\boldsymbol{a}_\beta} \in \mathcal{T}_D^{\mathcal{X}^n}$ and $\mathcal{S}_\beta \subset [n], |\mathcal{S}_\beta| = l$ where l depends only on d, C and $\mathcal{T}_D$ such that, $(\widetilde{\boldsymbol{a}_\beta})_{\boldsymbol{x}} = (\widetilde{\boldsymbol{a}_\beta})_{\boldsymbol{x}'}$, for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}^n$ such that $\boldsymbol{x}_i = \boldsymbol{x}'_i$ for all $i \in \mathcal{S}$, and

$$\mathrm{P}_{\boldsymbol{x} \sim \boldsymbol{\pi}}\left[(\boldsymbol{a}_\beta)_{\boldsymbol{x}} \neq (\widetilde{\boldsymbol{a}_\beta})_{\boldsymbol{x}}\right] = O_D(\epsilon)\,. \tag{15}$$

Since D is non-redundant, by Lemma 5, there exists $\phi : \mathcal{T}_D^{[\![k]\!]} \to \mathbb{R}^{\mathcal{U}}$ such that for all $\boldsymbol{x} \in \mathcal{X}^n$,

$$A_{(\boldsymbol{x},\cdot)} = \phi\left((\boldsymbol{a}_0)_{\boldsymbol{x}}, (\boldsymbol{a}_1)_{\boldsymbol{x}}, \ldots, (\boldsymbol{a}_k)_{\boldsymbol{x}}\right) \tag{16}$$

Hence,

$$\begin{aligned}
\mathrm{P}_{\boldsymbol{x} \sim \boldsymbol{\pi}}\left[A_{(\boldsymbol{x},\cdot)} \neq \phi((\hat{\boldsymbol{a}}_0)_{\boldsymbol{x}}, \ldots, (\hat{\boldsymbol{a}}_k)_{\boldsymbol{x}})\right] &\leq \mathrm{P}_{\boldsymbol{x} \sim \boldsymbol{\pi}}\left[\exists \beta : (\widetilde{\boldsymbol{a}_\beta})_{\boldsymbol{x}} \neq (\boldsymbol{a}_\beta)_{\boldsymbol{x}}\right] \\
&\leq \sum_{\beta \in [\![k]\!]} \mathrm{P}_{\boldsymbol{x} \sim \boldsymbol{\pi}}\left[(\widetilde{\boldsymbol{a}_\beta})_{\boldsymbol{x}} \neq (\boldsymbol{a}_\beta)_{\boldsymbol{x}}\right] \\
&\overset{(a)}{=} O_D(\epsilon)\,,
\end{aligned}$$

where (a) follows from (15). Define the deterministic matrix $\tilde{A} \in \{0,1\}^{\mathcal{X}^n \times \mathcal{U}}$ such that, for an arbitrary $u^* \in \mathcal{U}$ and for all $\boldsymbol{x} \in \mathcal{X}^n$,

$$\tilde{A}_{(\boldsymbol{x},\cdot)} = \begin{cases} \phi\left((\boldsymbol{a}_0)_{\boldsymbol{x}}, \ldots, (\boldsymbol{a}_k)_{\boldsymbol{x}}\right) & \text{if } \phi\left((\boldsymbol{a}_0)_{\boldsymbol{x}}, \ldots, (\boldsymbol{a}_k)_{\boldsymbol{x}}\right) \in \{\boldsymbol{\xi}_u : u \in \mathcal{U}\}, \\ \boldsymbol{\xi}_{u^*} & \text{otherwise.} \end{cases}$$

Since A is a deterministic matrix, i.e., each row of A belongs to $\{\boldsymbol{\xi}_u : u \in \mathcal{U}\}$,

$$\mathrm{P}_{\boldsymbol{x}\sim\boldsymbol{\pi}}\left[A_{(\boldsymbol{x},\cdot)} \neq \tilde{A}_{(\boldsymbol{x},\cdot)}\right] \leq \mathrm{P}_{\boldsymbol{x}\sim\boldsymbol{\pi}}\left[A_{(\boldsymbol{x},\cdot)} \neq \phi((\hat{\boldsymbol{a}}_0)_{\boldsymbol{x}}, \ldots, (\hat{\boldsymbol{a}}_k)_{\boldsymbol{x}})\right] = O_D(\epsilon)\,.$$

Finally, since $\widetilde{\boldsymbol{a}_\beta}$ depends only on $\mathcal{X}^{\mathcal{S}_\beta}$ for each $\beta \in [\![k]\!]$, $\tilde{A}$ depends only on $\mathcal{X}^{\cup_{\beta\in[\![k]\!]}\mathcal{S}_\beta}$.

We have shown that, for each $i \in \mathbb{N}$, there exists a deterministic matrix $\tilde{A}_i \in \mathbb{R}^{\mathcal{X}^{n_i}\times\mathcal{U}}$ such that

$$(\tilde{A}_i)_{(\boldsymbol{x},\cdot)} = (\tilde{A}_i)_{(\boldsymbol{x}',\cdot)}\ \forall \boldsymbol{x}, \boldsymbol{x'} \in \mathcal{X}^n \text{ s.t. } \boldsymbol{x}_j = \boldsymbol{x}'_j \text{ for all } j \in \bigcup_{\beta\in[\![k]\!]} \mathcal{S}_{i,\beta},$$

$$\text{and } \mathrm{P}_{\boldsymbol{x}\sim\boldsymbol{\pi}_i}\left[(\tilde{A}_i)_{(\boldsymbol{x},\cdot)} \neq (A_i)_{(\boldsymbol{x},\cdot)}\right] \leq \epsilon_i,$$

where $\mathcal{S}_{i,\beta}$ corresponds to $\mathcal{S}_\beta$ considered for a fixed i in the above discussion. Since $\cup_{\beta\in[\![k]\!]}\mathcal{S}_{i,\beta} \leq (k+1)l = \ell$ for all $i \in \mathbb{N}$, the statement of the theorem follows from Lemma 8. □

4.2 An Algorithm for the SNIR Problem

In this section, we show that the SNIR problem is decidable. Theorem 1 showed that existence of a statistical SNIR implies that of a perfect SNIR when the source correlation is common information free and the target correlation is non-redundant. As previously observed, it is sufficient to study SNIR between non-redundant correlations. Next, we tackle source correlations with non-zero common information.

Dealing with common information. An early work [28] on *non-secure* non-interactive reduction by Witsenhausen characterized correlations with non-zero common information to be the *complete correlations* — correlations that can be used to derive any desired target correlation — for non-secure reductions. However, as intuition suggests, common information does not help when security is required. This was formally established in [1] and restated in this paper as Lemma 4. Using this result, we will show that decidability of SNIR between general correlations reduces to SNIR between common information free correlations.

Definition 11. For positive numbers $0 < \alpha_1 \leq \ldots \leq \alpha_k < 1$ that add up to 1, and common information free correlations $H_1, \ldots, H_k$, consider the correlation

$$H = \begin{bmatrix} \alpha_1 H_1 & 0 & \ldots & 0 \\ 0 & \alpha_2 H_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \alpha_k H_k \end{bmatrix}.$$

The *parallelization* of H, denoted by $H^{\|}$ is defined as

$$H^{\|} = H_1 \otimes H_2 \otimes \ldots \otimes H_k.$$

◁

$H^{\|}$ is a common information free correlation, and when H is non-redundant, $H^{\|}$ is also non-redundant.

Lemma 9. *Let C be a correlation with non-zero common information. A non-redundant correlation D (with or without common information) has a statistical SNIR to C if and only if $D^{\|}$ has a statistical SNIR to $C^{\|}$.*

Proof: Let $C_{\text{coin}} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}$ be the 1-bit common randomness correlation. We will later show that, since C has non-zero common information, C has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$ and vice-versa. By the composability of SNIR protocols, this implies that D has a statistical SNIR to C if and only if it has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$. If D is common information free, then $D^{\|} = D$, and if D has non-zero common information, then D has a statistical SNIR to $C_{\text{coin}} \otimes D^{\|}$ and vice-versa, as in the case of C. Since common randomness can be (securely) sampled using common randomness, $C_{\text{coin}} \otimes D^{\|}$ has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$ if and only if $D^{\|}$ has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$. But, by Lemma 4, $D^{\|}$ has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$ only if $D^{\|}$ has a statistical SNIR to $C^{\|}$. Since the other direction is trivially true, we have established the statement of the lemma.

It remains to show that C has a statistical SNIR to $C_{\text{coin}} \otimes C^{\|}$ and vice-versa. Observe that Alice and Bob can agree on the distribution $\boldsymbol{\pi}$ over $[k]$ such that $(\boldsymbol{\pi})_i = \alpha_i$ for all i with arbitrarily small error using sufficiently many copies of C_{coin}. But then, by sampling i according to $\boldsymbol{\pi}$ and then sampling according to C_i, Alice and Bob have essentially sampled according to C.

To sample $C_{\text{coin}} \otimes C$ using C, Alice and Bob approximate (with arbitrarily small error) the 1-bit common randomness correlation C_{coin} using the sufficiently many copies of $\boldsymbol{\pi}$ distribution that they (implicitly) share. Furthermore, with probability $\prod_{i=1}^{k} \alpha_i$, the distribution $C^{\otimes l}$ is distributed according to $C_1 \otimes \ldots \otimes C_k$. Hence, Alice and Bob can approximately sample (with arbitrarily small error) from $C_{\text{coin}} \otimes C$ using sufficiently many copies of C. It is easily verified that both the above mentioned sampling schemes are secure, concluding the proof. □

Putting things together. Now we can put together our results so far into an algorithm for the SNIR problem.

1. Given a pair of correlations (C, D) as input, proceed as follows.
2. First replace D by its core and proceed (see Proposition 3). In the following we assume D is non-redundant.
3. If C has non-zero common information, then replace C by $C^{\|}$ and D by $D^{\|}$ (see Lemma 9). (Else retain both C and D unchanged.) In the following we assume C has no common information.
4. Compute $\ell \in \mathbb{N}$ associated with C and D, as stated in Theorem 1. Let $\boldsymbol{C} = C^{\otimes \ell}$.
5. For every pair of (deterministic) matrices $A \in \{0,1\}^{\mathcal{X}^\ell \times \mathcal{U}}$ and $B \in \{0,1\}^{\mathcal{Y}^\ell \times \mathcal{V}}$, check if (A, B) is a perfect SNIR from D to $\boldsymbol{C}$, using Proposition 4. That is, compute $V = \boldsymbol{\Delta}_D^{-1} B^\intercal \boldsymbol{\Delta}_{\boldsymbol{C}}$ and $U = \boldsymbol{\Delta}_{D^\intercal}^{-1} A^\intercal \boldsymbol{\Delta}_{\boldsymbol{C}^\intercal}$, and check if A,B,U, and

V satisfy the conditions in (1). If any pair is a perfect SNIR, accept the input and halt.

6. Else, reject the input and halt.

Steps 2 and 3 are justified by Proposition 3 and Lemma 9 respectively. Then, at the end of Step 3, given a non-redundant D and C with no common information, and the rest of the algorithm is justified by Theorem 1. This leads us to the main result of this paper:

Theorem 3. *The SNIR problem is decidable.*

4.3 More Necessary Conditions

As mentioned in Section 2, even given an algorithm for the SNIR problem, there is value in simple necessary conditions for an SNIR to exist. Here we present a new condition, exploiting Theorem 1.

A concrete example. The motivation for our new condition is the question of whether there is a statistical SNIR of the OT correlation (or more generally, any Oblivious Linear-Function Evaluation (OLE) correlation) to the (string) erasure correlation. We formally define these correlations before formally stating our resolution of the above question (in the negative).

(String) Erasure Correlation. An n-bit string erasure correlation with erasure probability $p \in (0,1)$, denoted by SEC_l^p, is a correlation over $\{0,1\}^l \times (\{0,1\}^l \cup \{\perp\})$ such that, for all $\boldsymbol{x} \in \{0,1\}^l$,

$$(\mathsf{SEC}_l^p)_{(\boldsymbol{x},\boldsymbol{y})} = \begin{cases} \frac{1-p}{2^l} & \text{if } \boldsymbol{y} = \boldsymbol{x}, \\ \frac{p}{2^l} & \text{if } \boldsymbol{y} = \perp. \end{cases}$$

OLE Correlation. The OLE correlation (or Oblivious Linear-Function Evaluation) over a finite field or ring $\mathbb{F}$ is the correlation $\mathsf{OLE}_{\mathbb{F}}$ over the domain $\mathbb{F}^2 \times \mathbb{F}^2$ such that, for all $a, b, x, y \in \mathbb{F}$,

$$(\mathsf{OLE}_{\mathbb{F}})_{((a,b),(x,y))} = \begin{cases} \frac{1}{|\mathbb{F}|^3} & \text{if } a \cdot b = x + y, \\ 0 & \text{otherwise .} \end{cases}$$

A New Impossibility Criterion. We state the following combinatorial criterion to rule out a SNIR.

Lemma 10. *Let C be a correlation over $\mathcal{X} \times \mathcal{Y}$ such that, for some x, $C_{(x,y)} > 0$ for all y. Let D be a non-redundant correlation over $\mathcal{U} \times \mathcal{V}$ such that, for each u, there exists v such that $D_{(u,v)} = 0$. Then, D does not have a statistical SNIR to C.*

Proof: By Theorem 1, D has a statistical SNIR to C only if there is a perfect SNIR $(\mathfrak{A}, \mathfrak{B})$ from D to $C^{\otimes \ell}$ for some $\ell \in \mathbb{N}$. By our assumption, there exists $\boldsymbol{x} \in \mathcal{X}^\ell$ such that $\left(C^{\otimes \ell}\right)_{(\boldsymbol{x},\boldsymbol{y})} > 0$ for all $\boldsymbol{y} \in C^{\otimes \ell}$. Consider any u such that

$\mathrm{P}[\mathfrak{A}(\boldsymbol{x}) = u] > 0$ (indeed $\mathfrak{A}$ is a deterministic protocol, but we do not need this property). It is easy to see that, for all v in the image of $\mathfrak{B}$,

$$\mathrm{P}_{(X,Y)\sim C^{\otimes \ell}}[\mathfrak{A}(X) = u, \mathfrak{B}(Y) = v] > 0.$$

This contradicts our assumption about D. □

The above lemma implies that statistical SNIR of $\mathsf{OLE}_{\mathbb{F}}$ to SEC_l^p is impossible for all $p \in (0,1)$, $l \in \mathbb{N}$ and ring $\mathbb{F}$. This follows from the fact that, when (X,Y) is distributed according to the n-bit string erasure correlation, $\mathrm{P}_{(X,Y)}[x, \perp] > 0$ for all $x \in \{0,1\}^n$. Whereas, when (U,V) is distributed according to SEC_l^p, for any (a,b) such that $a \cdot b \neq x + y$, $\mathrm{P}_{(U,V)}[U = (a,b), V = (x,y)] = 0$.

5 Generalized Junta Theorem

Kindler and Safra showed that if the energy of a function above degree d is small, then it is close to a junta that only depends on $O(d)$ many variables [21,20]. We need a generalized version of this result, Theorem 2, which we will prove in this section. The generalization is in terms of using a *generalized* Fourier transform to define degree and energy for functions over a domain Ω^n rather than $\{0,1\}^n$ (see Section 3). Our statement and proof closely follow the treatment by Filmus [9], which itself gives a generalization of the original result in [21,20] (which was restricted to functions with boolean *outputs* as well as inputs). The proofs of the lemmas are provided in the full version of the paper [6].

5.1 Tools: Influence, Hypercontractivity and Invariance Principle

We will first present some definitions that will be used later in this section. Let $(\Omega, \boldsymbol{\pi})$ be a finite probability space, $|\Omega| = m \geq 2$, in which every outcome has probability at least λ.

Definition 12. For a function $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$, and a position $i \in [n]$, we define the following:

$$\mathbb{E}[\boldsymbol{f}] = \mathop{\mathbb{E}}_{x\sim\boldsymbol{\pi}^{\otimes n}}[\boldsymbol{f}(x)], \qquad \mathsf{Var}[\boldsymbol{f}] = \|\boldsymbol{f} - \mathbb{E}[\boldsymbol{f}]\|^2,$$
$$\mathsf{E}_i\boldsymbol{f} = \mathop{\mathbb{E}}_{x'\sim\boldsymbol{\pi}}[\boldsymbol{f}(x_{[n]\setminus i}||x')] \qquad L_i\boldsymbol{f} = \boldsymbol{f} - \mathsf{E}_i\boldsymbol{f},$$
$$\mathsf{Inf}_i[\boldsymbol{f}] = \|L_i\boldsymbol{f}\|^2 \qquad \mathsf{TotInf}[\boldsymbol{f}] = \sum_{i=1}^{n} \mathsf{Inf}_i[\boldsymbol{f}]$$

where $x_{[n]\setminus i}||x'$ denotes replacing x_i by x' in x. ◁

Note that $L_i\boldsymbol{f}$ is a function associated with $\boldsymbol{f}$, called its *Laplacian*, and it captures the contribution of a particular coordinate i for each point in the domain of $\boldsymbol{f}$. Its energy $\mathsf{Inf}_i[\boldsymbol{f}]$, called the *influence* of a position i, is a quantity that measures this contribution. The total influence $\mathsf{TotInf}[\mathbf{f}]$ simply sums up

the influence from all coordinates. We also note that the expectation $\mathbb{E}$ (and variance Var) for a function $\boldsymbol{f}$ w.r.t. a distribution over its domain are defined as the expectation (and variance) of the random variable corresponding to the output of $\boldsymbol{f}$ when evaluated on an input drawn from the given distribution; these definitions extend to continuous domains as well.

Hypercontractivity. We will need a generalization of the Bonami Lemma in the hypercontractivity type of results. The following version is obtained by substituting $q = 4$ in Theorem 10.21 in [25]. Here we state this lemma restricted to the case when all the variables are coming from the same domain Ω and are distributed identically according to $\boldsymbol{\pi}$.

Lemma 11 (Hypercontractivity [25])**.** *Let $(\Omega, \boldsymbol{\pi})$ be a finite probability space, $|\Omega| = m \geq 2$, in which every outcome has probability at least λ. Let $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ be a function of degree at most d, then*

$$\|\boldsymbol{f}^2\|^2 \leq (9/\lambda)^d \, \|\boldsymbol{f}\|^4.$$

We will use the above lemma to prove a result that is a dichotomy (given below), which will be used as a tool in the main proof. This has been taken verbatim from [9], the only difference being that now we use the generalized version of the Bonami lemma given above. Also, the expectations and probability calculations will now be with respect to the probability distribution given by $\boldsymbol{\pi}^{\otimes n}$ instead of the uniform distribution.

Lemma 12. *Let $\mathcal{S}$ be a finite set and let $d \geq 1$. If $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ is a $\mathcal{S}$-valued function satisfying $\|\boldsymbol{f}^{>d}\|^2 = \epsilon$ then either $\|\boldsymbol{f}\|^2 = O(\epsilon)$ or $\|\boldsymbol{f}^{\leq d}\|^2 = \Omega(1)$.*

Invariance Principle. Given a function in $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$, we define a *polynomial* $\mathsf{P}\boldsymbol{f}$ with $n(|\Omega| - 1)$ variables, obtained by replacing the generalized Fourier basis with a polynomial basis. More precisely, the polynomial $\mathsf{P}\boldsymbol{f}$ with formal variables $\{X_{i,\boldsymbol{\alpha}}\}_{i \in [n], \boldsymbol{\alpha} \in [|\Omega|-1]}$ is defined as

$$\mathsf{P}\boldsymbol{f}(X_{1,1}, \ldots, X_{n,|\Omega|-1}) = \sum_{\boldsymbol{\alpha} \in [\![|\Omega|]\!]^n} \widehat{\boldsymbol{f}}_{\boldsymbol{\alpha}} \prod_{i \in [n] : \alpha_i \neq 0} X_{i,\alpha_i}.$$

We will be using the following variant of the invariance principle to complete our proof of the generalized junta theorem, that is implicit in [25] (obtained from Exercise 11.49(b) followed by an application of the technique used in the proof of Corollary 11.67):

Lemma 13 (Invariance Principle [25])**.** *Let $(\Omega, \boldsymbol{\pi})$ be a finite probability space, $|\Omega| = m \geq 2$, in which every outcome has probability at least λ. Suppose $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ has degree at most d, with $\mathsf{Var}[\boldsymbol{f}] = 1$,[6] and $\mathsf{Inf}_t[\boldsymbol{f}] \leq \epsilon$, for every $t \in [n]$. Then for any $\psi : \mathbb{R} \to \mathbb{R}$ that has a continuous third derivative and satisfies $\|\psi'''\|_\infty \leq c$, we have*

$$\left| \mathbb{E}[\psi \circ \boldsymbol{f}] - \mathop{\mathbb{E}}_{w \sim N(0,1)^{(m-1)n}} [\psi(\mathsf{P}\boldsymbol{f}(w))] \right| \leq \frac{2c}{3} (2\sqrt{2/\lambda})^d d\sqrt{\epsilon}.$$

[6] Actually, this lemma holds even when $\mathsf{Var}[\boldsymbol{f}] \leq 1$, but the unit variance case is sufficient for us to be able to apply the Carbery-Wright theorem later.

Following [25], we will use the above lemma along with Proposition 5 below to get the desired version of the invariance principle, that compares probabilities of the functions taking values less than some threshold.

Proposition 5 (Carbery-Wright Theorem.). *Let $p : \mathbb{R}^{(m-1)n} \to \mathbb{R}$ be a polynomial of degree at most d, let $w \sim N(0,1)^{(m-1)n}$, and assume $\mathbb{E}[p(w)^2] = 1$. Then, for all $\epsilon > 0$,*

$$\Pr[|p(w)| \leq \epsilon] \leq O(d\epsilon^{1/d}),$$

where the $O(\cdot)$ hides a universal constant.

The lemma given below will also be used in our proof of the invariance principal. Its proof has been completed after following exercises 11.40 and 11.41(b) from [25].

Lemma 14. *Fix $u \in \mathbb{R}$, let $\psi(s) = 1_{s \leq u}$, and $0 < \eta < 1/2$. Then there exists a smooth approximation $\widetilde{\psi}_\eta$ of ψ that satisfies the following properties:*

- *$\widetilde{\psi}_\eta$ is a non-increasing function which agrees with the indicator function $\psi(s) = 1_{s \leq u}$ on the intervals $(-\infty, u - \eta]$ and $[u + \eta, \infty)$.*
- *$\widetilde{\psi}_\eta$ is smooth and satisfies $\|\widetilde{\psi}_\eta^{(k)}\|_\infty \leq c_k/\eta^k$ for each $k \in \mathbb{N}$, where c_k only depends on k.*

We prove the below lemma by following a line of reasoning in [25], similar to the way this version is proved for the Berry-Esseen theorem (which considers sums of random variables instead of multilinear polynomials). The smooth approximation function $\widetilde{\psi}_\eta$ for the desired indicator function $\psi(s) = 1_{s \leq u}$ has been taken from Lemma 14. We first apply the basic invariance principle Lemma 13 to this approximation and then use its properties to derive some basic inequalities, concluding in the desired result.

Lemma 15. *Let $(\Omega, \boldsymbol{\pi})$ be a finite probability space, $|\Omega| = m \geq 2$, in which every outcome has probability at least λ. Suppose $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ has degree at most d, with $\mathsf{Var}[\boldsymbol{f}] = 1$, and $\mathsf{Inf}_t[\boldsymbol{f}] \leq \epsilon$, for every $t \in [n]$. Then, for all $u \in \mathbb{R}$,*

$$\left| \Pr_{x \sim \boldsymbol{\pi}^{\otimes n}}[\boldsymbol{f}(x) \leq u] - \Pr_{w \sim N(0,1)^{(m-1)n}}[\mathsf{P}\boldsymbol{f}(w) \leq u] \right| \leq O(d\epsilon^{\frac{1}{3d+1}} \lambda^{\frac{-1}{3}}).$$

5.2 Main Proof

In this section, we give our proof for the junta theorem. Lemma 16 states that for every function whose high-degree energy is small, there is a set J of small number of co-ordinates s.t. all other positions have low influence on the function $\boldsymbol{f}_z$ obtained by fixing these co-ordinates to the value $z \in \Omega^J$. The remaining lemmas then basically try to show that low influence implies low variance. To this end, the invariance principle is used inside Lemma 18 to claim that the function $\boldsymbol{f}_z$ has low probability on every value in the domain. Lemma 19 then shows that if we consider a restriction of the function $\boldsymbol{f}$ to J and average out the

rest, we get a good approximation $\boldsymbol{g}$ for $\boldsymbol{f}$. From there, one only needs to round $\boldsymbol{g}$ to the nearest values in the set $\mathcal{T}$ to get the final approximation $\boldsymbol{h}$. Most of the following description has been taken verbatim from [9], with the expectations and probabilities now being calculated w.r.t. the general distribution given by products of π instead of a uniform distribution. Whenever not mentioned, the variable $\boldsymbol{\alpha}$ is coming from $[\![|\Omega|]\!]^n$.

There are some updates to the proof of the following lemma in the generalized setting as compared to [9]. First, the Laplacian functions $L_i\boldsymbol{f}$, for $i \in [n]$, have a different range defined with respect to π, whereas it is a much simpler description in the uniform setting. Secondly, the constant $2^{|J|}$ in an inequality comparing the conditional expectation of $(L_i\boldsymbol{f})^2$ when the input at positions J has been set to z and the normal expectation of $(L_i\boldsymbol{f})^2$, must be replaced by $(1/\lambda)^{|J|}$.

Lemma 16. *Let $\mathcal{T}$ be a finite set and let $d \geq 1$. If $\boldsymbol{f} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$ is a $\mathcal{T}$-valued function such that $\|\boldsymbol{f}^{>d}\|^2 = \epsilon$ then we can find a set J of $O(1)$ coordinates such that for each $z \in \Omega^J$, the function $[\boldsymbol{f}_z]$ on $\Omega^{\bar{J}}$ obtained by substituting $x|_J = z$ satisfies*

$$\mathsf{Inf}_i[\boldsymbol{f}_z] = O(\epsilon) \text{ for every } i \in \bar{J}.$$

There is a similar difference to the following lemma from the version in [9], that the constant in an inequality comparing norms of $\boldsymbol{f}_z$ and $\boldsymbol{f}$ must be replaced from $2^{|J|}$ to $(1/\lambda)^{|J|}$.

Lemma 17. *Assuming the setting of Lemma 16, then for every $z \in \Omega^J$, either $\mathsf{Var}[\boldsymbol{f}_z] = O(\epsilon)$ or $\mathsf{Var}[\boldsymbol{f}_z] = \Omega(1)$.*

The following lemma finally concludes that variance of $\boldsymbol{f}_z$ is small, for every $z \in \Omega^J$. The key ingredient used here is our variant of the generalized invariance principle (Lemma 15) that we have proved earlier. This is used to first show that the difference in the (appropriately-defined) probabilities that the functions $\boldsymbol{g}$ and $\mathsf{P}\boldsymbol{g}$ take values in the set $(u-\gamma, u]$ (for any u) is small, where $\boldsymbol{g} = \boldsymbol{f}_z/\mathsf{Var}[\boldsymbol{f}_z]$. Since the variable obtained by substituting a Gaussian distribution to $\mathsf{P}\boldsymbol{g}$ is continuous, this probability goes to 0 as we take the limit γ going to 0.

Lemma 18. *Assuming the setting of Lemma 16, for every $z \in \Omega^J$, we have $\mathsf{Var}[\boldsymbol{f}_z] = O(\epsilon)$.*

Lemma 19. *Assuming the setting of Lemma 16, there is a function $\boldsymbol{g} \in L^2(\Omega^n, \boldsymbol{\pi}^{\otimes n})$, depending only on the co-ordinates in J, such that $\|\boldsymbol{f} - \boldsymbol{g}\|^2 = O(\epsilon)$.*

We will now finish the proof of the generalized junta theorem.

Proof of Theorem 2: Lemma 19 gives a function $\boldsymbol{g}$, depending on $O(1)$ co-ordinates, such that $\|\boldsymbol{f} - \boldsymbol{g}\|^2 = O(\epsilon)$. Let $\boldsymbol{h}(x)$ be obtained by rounding $\boldsymbol{g}(x)$ to the closest element of $\mathcal{T}$. For every x we have $|\boldsymbol{h}(x) - \boldsymbol{g}(x)| \leq |\boldsymbol{f}(x) - \boldsymbol{g}(x)|$ and so $|\boldsymbol{h}(x) - \boldsymbol{f}(x)| \leq |\boldsymbol{h}(x) - \boldsymbol{g}(x)| + |\boldsymbol{g}(x) - \boldsymbol{f}(x)| \leq 2|\boldsymbol{f}(x) - \boldsymbol{g}(x)|$. Consequently, $\|\boldsymbol{h} - \boldsymbol{f}\|^2 \leq 4\|\boldsymbol{g} - \boldsymbol{f}\|^2 = O(\epsilon)$.

Since $\boldsymbol{f}$ and $\boldsymbol{h}$ are both $\mathcal{T}$-valued, for all x either $\boldsymbol{f}(x) = \boldsymbol{h}(x)$ or $|\boldsymbol{h}(x) - \boldsymbol{f}(x)| = \Omega(1)$. Consequently, $\mathbb{E}_{x \sim \boldsymbol{\pi}^{\otimes n}}[(\boldsymbol{h} - \boldsymbol{f})^2] = \Omega(\Pr[\boldsymbol{h} \neq \boldsymbol{f}])$, and so $\Pr[\boldsymbol{h} \neq \boldsymbol{f}] = O(\|\boldsymbol{h} - \boldsymbol{f}\|^2) = O(\epsilon)$.

Finally, suppose that $\boldsymbol{h}$ does not have degree d. Then $\widehat{\boldsymbol{h}}_{\boldsymbol{\alpha}} \neq 0$ for some $|\boldsymbol{\alpha}| > d$. Since $\boldsymbol{h}$ depends on $M = O(1)$ coordinates, $\widehat{\boldsymbol{h}}_{\boldsymbol{\alpha}} = \mathbb{E}[\boldsymbol{h} \cdot \boldsymbol{\gamma}_{\boldsymbol{\alpha}}]$ is a non-zero value which is the average of m^M elements, and consequently $(\widehat{\boldsymbol{h}}_{\boldsymbol{\alpha}})^2 = \Omega(1)$, implying that $\|\boldsymbol{h}^{>d}\|^2 = \Omega(1)$. On the other hand,

$$\|\boldsymbol{h}^{>d}\|^2 \leq 2\|\boldsymbol{f}^{>d}\|^2 + 2\|\boldsymbol{h}^{>d} - \boldsymbol{f}^{>d}\|^2 = 2\epsilon + 2\|(\boldsymbol{h}-\boldsymbol{f})^{>d}\|^2 \leq 2\epsilon + \|\boldsymbol{h}-\boldsymbol{f}\|^2 = O(\epsilon).$$

This shows that $\epsilon = \Omega(1)$. In such a setting, we can just ignore all the above analysis and pick some constant function $\boldsymbol{h}$ as the approximation. This function has degree 0, it depends on no co-ordinates, while the probability that it differs from $\boldsymbol{f}$ is still less than 1, and hence $O(\epsilon)$. □

References

1. Pratyush Agarwal, Varun Narayanan, Shreya Pathak, Manoj Prabhakaran, Vinod M. Prabhakaran, and Mohammad Ali Rehan. Secure non-interactive reduction and spectral analysis of correlations. In *EUROCRYPT 2022*, pages 797–827, 2022.
2. Hamidreza Amini Khorasgani, Hemanta K. Maji, and Hai H. Nguyen. Secure non-interactive simulation: Feasibility and rate. In *EUROCRYPT 2022*, pages 767–796, 2022.
3. Venkat Anantharam, Amin Aminzadeh Gohari, Sudeep Kamath, and Chandra Nair. On maximal correlation, hypercontractivity, and the data processing inequality studied by erkip and cover. *CoRR*, abs/1304.6133, 2013.
4. Amos Beimel, Yuval Ishai, Ranjit Kumaresan, and Eyal Kushilevitz. On the cryptographic complexity of the worst functions. In *TCC*, pages 317–342, 2014.
5. Amos Beimel and Tal Malkin. A quantitative approach to reductions in secure computation. In *TCC*, pages 238–257, 2004.
6. Kaartik Bhushan, Ankit Kumar Misra, Varun Narayanan, and Manoj Prabhakaran. Secure non-interactive reducibility is decidable. Cryptology ePrint Archive, 2022.
7. Anindya De, Elchanan Mossel, and Joe Neeman. Non interactive simulation of correlated distributions is decidable. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18, page 2728–2746, USA, 2018. Society for Industrial and Applied Mathematics.
8. Uriel Feige, Joe Kilian, and Moni Naor. A minimal model for secure computation (extended abstract). In *STOC*, pages 554–563, 1994.
9. Yuval Filmus. A simple proof of the Kindler–Safra theorem. https://yuvalfilmus.cs.technion.ac.il/Manuscripts/KindlerSafra.pdf, 2022.
10. P. Gács and J. Körner. Common information is far less than mutual information. *Problems of Control and Information Theory*, 2(2):149–162, 1973.
11. Badih Ghazi, Pritish Kamath, and Madhu Sudan. Decidability of non-interactive simulation of joint distributions. In *FOCS*, pages 545–554. IEEE Computer Society, 2016.
12. Oded Goldreich. *Foundations of Cryptography: Basic Applications*. Cambridge University Press, 2004.
13. Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play ANY mental game. In *STOC*, pages 218–229, 1987. See [12, Chap. 7] for more details.

14. Oded Goldreich and Ronen Vainish. How to solve any protocol problem - an efficiency improvement. In *CRYPTO*, pages 73–86, 1987.
15. Saumya Goyal, Varun Narayanan, and Manoj Prabhakaran. Oblivious-transfer complexity of noisy coin-toss via secure zero communication reductions. In these proceedings, 2022.
16. Yuval Ishai, Manoj Prabhakaran, and Amit Sahai. Founding cryptography on oblivious transfer - efficiently. In *CRYPTO*, pages 572–591, 2008.
17. Sudeep Kamath and Venkat Anantharam. On non-interactive simulation of joint distributions. *IEEE Trans. Inf. Theory*, 62(6):3419–3435, 2016.
18. Hamidreza Amini Khorasgani, Hemanta K. Maji, and Hai H. Nguyen. Decidability of secure non-interactive simulation of doubly symmetric binary source. Cryptology ePrint Archive, Report 2021/190, 2021. https://eprint.iacr.org/2021/190.
19. Joe Kilian. Founding cryptography on oblivious transfer. In *STOC*, pages 20–31, 1988.
20. Guy Kindler. *Property Testing, PCP, and juntas*. PhD thesis, Tel Aviv University, 2002.
21. Guy Kindler and Shmuel Safra. Noise-resistant boolean functions are juntas. Manuscript available from http://www.math.tau.ac.il/~safra/PapersAndTalks/nibfj.ps, 2002.
22. Daniel Kraschewski, Hemanta Maji, Manoj Prabhakaran, and Amit Sahai. A full characterization of completeness for two-party randomized function evaluation. In *EUROCRYPT*, 2014.
23. Hemanta Maji, Manoj Prabhakaran, and Mike Rosulek. *Complexity of Multi-Party Computation Functionalities*, volume 10 of *Cryptology and Information Security Series*, pages 249 – 283. IOS Press, Amsterdam, 2013.
24. Varun Narayanan, Manoj Prabhakaran, and Vinod M. Prabhakaran. Zero-communication reductions. In *TCC*, volume 12552, pages 274–304. Springer, 2020.
25. Ryan O'Donnell. Analysis of boolean functions. *CoRR*, abs/2105.10386, 2021.
26. Madhu Sudan, Himanshu Tyagi, and Shun Watanabe. Communication for generating correlation: A unifying survey. *IEEE Trans. Inf. Theory*, 66(1):5–37, 2020.
27. H. S. Witsenhausen. On sequences of pairs of dependent random variables. *SIAM Journal on Applied Mathematics*, 28(1):100–113, 1975.
28. Hans S. Witsenhausen. The zero-error side information problem and chromatic numbers (corresp.). *IEEE Transactions on Information Theory*, 22(5):592–593, 1976.
29. A. D. Wyner. The common information of two dependent random variables. *IEEE Transactions on Information Theory*, 21(2):163–179, 1975.