Naor-Reingold Goes Public: The Complexity of Known-Key Security

Pratik Soni(✉) and Stefano Tessaro

University of California, Santa Barbara, USA
{pratik_soni,tessaro}@cs.ucsb.edu

Abstract. We study the complexity of building secure block ciphers in the setting where the key is known to the attacker. In particular, we consider two security notions with useful implications, namely public-seed pseudorandom permutations (or psPRPs, for short) (Soni and Tessaro, EUROCRYPT '17) and correlation-intractable ciphers (Knudsen and Rijmen, ASIACRYPT '07; Mandal, Seurin, and Patarin, TCC '12).

For both these notions, we exhibit constructions which make only *two* calls to an underlying non-invertible primitive, matching the complexity of building a pseudorandom permutation in the secret-key setting. Our psPRP result instantiates the round functions in the Naor-Reingold (NR) construction with a secure UCE hash function. For correlation intractability, we instead instantiate them from a (public) random function, and replace the pairwise-independent permutations in the NR construction with (almost) $O(k^2)$-wise independent permutations, where k is the arity of the relations for which we want correlation intractability.

Our constructions improve upon the current state of the art, requiring five- and six-round Feistel networks, respectively, to achieve psPRP security and correlation intractability. To do so, we rely on techniques borrowed from Impagliazzo-Rudich-style black-box impossibility proofs for our psPRP result, for which we give what we believe to be the first *constructive* application, and on techniques for studying randomness with limited independence for correlation intractability.

Keywords: Foundations · Known-key security · Pseudorandomness psPRPs · Correlation-intractability · Limited independence

1 Introduction

1.1 Overview and Motivation

Block ciphers are traditionally used within modes of operation where they are instantiated under a *secret* key. Provable security results typically assume them to be good *pseudorandom permutations* (PRPs). This has motivated a large body of theoretical works on building PRPs from weaker or less structured components, e.g., through the Feistel construction and its variants [24,28,33].

Block ciphers are however also frequently used in settings where the key is fixed, or at least *known*. We refer to this as the *known-key setting*. For instance,

J. B. Nielsen and V. Rijmen (Eds.): EUROCRYPT 2018, LNCS 10822, pp. 653–684, 2018.
https://doi.org/10.1007/978-3-319-78372-7_21

it is common to rely on *permutations*[1] or (equivalently) *fixed-key* ciphers to build hash functions [12,35], authenticated encryption [3], PRNGs [13,23], and even more involved objects, such as garbling schemes [8].

As there is no secret key to rely upon, it is less clear what kind of security properties block ciphers should satisfy in this setting. Hence, security proofs typically assume the cipher to behave like an ideal random permutation on each key. A number of design paradigms for block ciphers (cf. e.g. [1,17–19,21,25] to mention a few results) are therefore analyzed in terms of *indifferentiability* [31], an ideal-model property which implies that for single-stage security games, the cipher inherits all properties of an ideal cipher. Still, the resulting proofs are notoriously involved, and the constructions more complex than seemingly necessary for the applications in which they are used. This is in sharp contrast with hash functions, where indifferentiability has helped shaping real-world designs.

Our contributions. The only two exceptions to the above indifferentiability-based approach we are aware of are the notions of a *public-seed pseudorandom permutation* (psPRP) [37] and of *correlation-intractable block ciphers* [27,29]. Block ciphers satisfying variants of both have been shown to be sufficient to instantiate several schemes that otherwise only enjoyed security proofs assuming the cipher is ideal. Yet, the complexity of actually building these primitives from simpler objects is not understood.

In this work, we present constructions for each of the notions which only make *two* calls to an underlying non-invertible round function. All of our constructions are instantiations of the *Naor-Reingold construction* [33], which is the most efficient known approach to build a secure PRP, and we thus show that it retains meaningful properties when the seed is made *public* under appropriate assumptions on the round functions. The previously known best constructions require Feistel networks with five rounds (for psPRPs) [37] and six rounds (for correlation intractability) [29], and in both cases the security proofs relied indirectly on (weakened) forms of indifferentiability, inheriting seemingly unnecessary complexity. Here, we introduce substantially different techniques to bypass limitations of existing proofs, borrowing from areas such as black-box separations and applications of limited-independence.

We stress that our focus here is on foundations, and more specifically, breaking complexity barriers. While we follow the good practice of giving concrete bounds, we make no claims that these are suitable for practical applications. We hope however to spur quantitative research in this direction.

1.2 Public-Seed Pseudorandomness via the NR Construction

We start with our results on *public-seed pseudorandom permutations* (or psPRPs, for short), a notion recently introduced in [37], which considers a family of permutations E on n-bit strings, indexed by a seed s. (This could be obtained from a block cipher.) Ideally, we would like $\mathsf{E}_s(\cdot)$ and $\mathsf{E}_s^{-1}(\cdot)$ to be indistinguishable

[1] Permutations, as in the sponge construction, correspond to the extreme case where there is only one possible key to choose from.

from ρ and ρ^{-1} (for a random permutation ρ), *even* if the seed s is known to the distinguisher. This is obviously impossible, yet an approach to get around this borrowed from the UCE framework [6] is to *split* the distinguisher into two stages. A first stage, called the *source* S, gets access to either $(\mathsf{E}_s, \mathsf{E}_s^{-1})$ or (ρ, ρ^{-1}), but does *not* know s, and then passes on some leakage L to a second-stage PPT D, the *distinguisher*, which learns s, but has no access to the oracle any more. If E is indeed secure, D will not be able to guess which one of the two oracles S had access to. This is very similar to the security definition of a UCE H, the only difference is that there the source accesses either H_s or a random *function* f.

Clearly, nothing is gained if L is unrestricted, and thus restrictions on S are necessary. Two classes of sources were in particular considered in [37], *unpredictable* and *reset-secure* sources, inspired by analogous notions for UCEs. The former demands that when the source S accesses ρ and ρ^{-1}, an (unbounded)[2] predictor P given the leakage L cannot then guess any of S's queries (and their inverses). In contrast, the latter notion demands that a computationally unbounded distinguisher R, given L cannot tell apart whether it is given oracle access to the *same* permutation ρ, or an independent one, within a polynomial number of queries. Being a psPRP for all unpredictable sources is a potentially weaker assumption than being a psPRP for reset-secure sources, since every unpredictable source is reset-secure, but not vice versa.

Applications. PsPRPs for such restricted source classes are a versatile notion. For example, a psPRP for all reset-secure sources can be used to instantiate the permutation within permutation-based hash functions admitting indifferentiability-based security proofs, such as the sponge construction [12] (which underlies SHA-3), turning them into a UCE-secure hash function sufficient for a number of applications, studied in multiple works [5,6,11,20,30]. Also, [37] shows that psPRPs for unpredictable sources are sufficient to instantiate garbling schemes obtained from fixed-key blocks ciphers [8].

Constructing psPRPs: Previous work. But do psPRPs exist at all? Soni and Tessaro [37] show that they are implied by sufficiently strong UCEs:

Theorem (Informal) [37]**.** *The five-round Feistel construction, with round functions instantiated from a UCE* H *for reset-secure sources, is a psPRP for reset-secure sources.*

This left two obvious questions open, however: **(1)** Whether the number of rounds can be reduced, and **(2)** whether the same holds for unpredictable sources, too. The techniques of [37], based on proving a weaker notion of indifferentiability for five-round Feistel, fail to help answering both questions.

Our contributions. We solve both questions, and even more in fact, showing that the Naor-Reingold (NR) construction [33] solves *both* (1) and (2). In

[2] Computational versions of these notions can be defined, but the resulting notions can easily be shown impossible under the assumption that IO exists [14], and are ignored in this paper.

particular, let H be a family of functions from $n+1$ bits to n, and let P be a family of permutations on $2n$ bit strings. Then, the NR construction on seed $\boldsymbol{s} = (s, s^{\mathsf{in}}, s^{\mathsf{out}})$ and input $u \in \{0,1\}^{2n}$, outputs $v \in \{0,1\}^{2n}$, where

$$x_0 \,\|\, x_1 \leftarrow \mathsf{P}_{s^{\mathsf{in}}}(u)\ ,\ x_2 \leftarrow \mathsf{H}_s(0 \,\|\, x_1) \oplus x_0, \\ x_3 \leftarrow \mathsf{H}_s(1 \,\|\, x_2) \oplus x_1\ ,\ v \leftarrow \mathsf{P}^{-1}_{s^{\mathsf{out}}}(x_3 \,\|\, x_2).$$

The key point here is that P only needs to satisfy a weak non-cryptographic property, namely that for a random s and for any distinct $u \neq u'$, the right halves of $\mathsf{P}_s(u)$ and $\mathsf{P}_s(u')$ only collide with negligible probability. Therefore, only two calls to a "cryptographically hard" round function H are made. Naor and Reingold [33] showed that NR is a (strong) PRP whenever H is a pseudorandom function. Here, we show the following public-seed counterparts:

Theorems 1 and 2 (Informal). *The NR construction, with round functions instantiated with a UCE H for X-sources, is a psPRP for X-sources, where $X \in \{$reset-secure, unpredictable$\}$.*

A detailed overview of our techniques is given below in Sect. 1.4. We remark here that such UCEs are of course strong, and the question of basing these on simpler assumptions is wide open. Still, we believe such results to be very important: First off, they show relations among notions, and getting a UCE (without any injectivity structure) is possibly simpler in practice than in theory (i.e., using the compression function of SHA-256). Second, even if we instantiate H from a random oracle (which gives a good UCE [6]), the result *is* useful, as this would give us a simple instantiation of a (seeded) permutation in applications which are not even known to follow from full-fledged indifferentiability, as discussed by Mittelbach [32].

1.3 Correlation Intractability

The notion of *correlation intractability* (CI) of *hash functions* was introduced by Canetti et al. [15] as a weakening of a random oracle. CI naturally extends to permutations and block ciphers [27,29]: Given the seed s, an adversary should not be able to find an input-output pair (u, v) such that $\mathsf{E}_s(u) = v$ and such that $(u, v) \in R$, where R is a hard-to-satisfy relation for a truly random permutation, a so-called *evasive* relation. This, in turn, can be generalized to k-ary relations, where k input-output pairs are to be provided. CI is well-known not to hold in the standard model for arbitrary evasive relations.[3] Therefore, here, we target constructions in ideal models.

Applications. CI has important applications – for example, let E be a permutation family on $2n$-bit strings. Then, for $n < m < 2n$, consider the function family H from m bits to n bits such that

$$\mathsf{H}_s(x) = \mathsf{E}_s(x \,\|\, 0^{2n-m})[1 \ldots n],$$

[3] Though, of course, it could be true for specific interesting relations.

i.e., this outputs the first n bits of $\mathsf{E}_s(x \,\|\, 0^{2n-m})$. Then, it is not too hard to show that if E satisfies CI for evasive binary relations, then H is collision resistant – indeed, a collision yields two distinct pairs (u_1, v_1), (u_2, v_2) where $u_1[m+1 \ldots 2n] = u_2[m+1 \ldots 2n] = 0^{2n-m}$, whereas $v_1[1 \ldots n] = v_2[1 \ldots n]$. Along similar lines, one can prove that H can be used to instantiate the Fiat-Shamir transform [22] whenever E satisfies CI for *unary* relations. And so on.

Correlation intractability for Feistel networks. Indifferentiability is easily seen to imply CI, and therefore, by [19], the Feistel construction with 8 rounds is correlation intractable. In fact, Mandal et al. [29] observed that a weaker notion of indifferentiability, called sequential indifferentiability, is sufficient for CI, and could show that 6 rounds are enough. It is known that the 5-round Feistel construction is *not* correlation intractable for evasive 4-ary relations (an attack was given in [29]). Other weaker notions of indifferentiability are known to imply CI, but do no appear to lead to any complexity improvements for constructions from non-invertible primitives [2,16].

Our results. We study the correlation intractability of the NR construction described above where the two calls to H are replaced by two calls to (seedless) public random function f from $n+1$ to n bits, and the seed of the construction only consists of the seeds for P. (This is similar to the model of Ramzan and Reyzin [34], although they consider PRP security, and secret seeds.)

In general, this basic form of the NR construction cannot be correlation intractable – indeed, P can be instantiated by one-round Feistel with a pairwise-independent round function, and generic attacks against the correlation intractability of four-round Feistel would still apply. We show however the following result:

Theorem 3 (Informal). *For any* constant $k = O(1)$*, if* P^{-1} *is an almost* $O(k^2)$*-wise independent permutation, then the NR construction is correlation intractable for every* k*-ary evasive relation.*

For unary relations (i.e., $k = 1$), we can in fact show that instantiating P with one-round Feistel using a 10-wise independent round function suffices. As this is effectively a four-round Feistel network, this confirms that no generic attacks exist for unary relations. Our result extends to non-constant k, however under some restrictions on the class of evasive relations for which we can prove correlation intractability. We believe an important part of our result is the technique we use, which gives a surprising paradigm to amplify CI unconditionally which we discuss below in Sect. 1.5.

Limitations. In contrast to existing 6-round results, our result is weaker in that it only covers evasive relations fully if $k = O(1)$. We are not aware of counter examples showing attacks for larger k's, but our proof inherently fails. We note however that most applications of correlation intractability only require constant arity, and we leave the question of assessing whether six calls to a random function are necessary for arbitrary arity for future work.

1.4 Technical Overview – psPRPs

Let us briefly recall the setting: For some PPT source S, which queries a permutation oracle on $2n$-bit strings to produce a leakage L, we need to show that any PPT distinguisher D which learns L *and* $\boldsymbol{s} = (s, s^{\mathsf{in}}, s^{\mathsf{out}})$ cannot tell apart whether S was accessing NR using a UCE H (with seed $\boldsymbol{s}$) or a truly random permutation. We assume S is either (statistically) unpredictable (in Theorem 1) or reset-secure (in Theorem 2).

The source $\overline{S}$. The natural approach we follow is to build another source, $\overline{S}$ from S, for which H should be a secure UCE. This source thus accesses an oracle $\mathcal{O}$ implementing a function from $n+1$ to n bits. It first samples seeds $s^{\mathsf{in}}, s^{\mathsf{out}}$ for P, and then simulates an execution of S. The oracle calls by the latter are processed by evaluating the NR construction using $s^{\mathsf{in}}, s^{\mathsf{out}}$, and the oracle $\mathcal{O}(\cdot)$ in lieu of $\mathsf{H}(s, \cdot)$. Finally, when S produces its output L, $\overline{S}$ outputs $(L, s^{\mathsf{in}}, s^{\mathsf{out}})$. We will show the following two facts:

- Fact 1. If S is unpredictable (w.r.t. the psPRP notion), then $\overline{S}$ is unpredictable (w.r.t. the UCE notion).
- Fact 2. If S is reset-secure (w.r.t. the psPRP notion), then $\overline{S}$ is reset-secure (w.r.t. the UCE notion).

Theorems 1 and 2 follow from Facts 1 and 2, respectively, by a fairly straightforward application of the (classical) indistinguishability of the NR construction with *random* round functions.[4]

The unpredictable case. Our approach to establish Fact 1 is inspired by an elegant proof of secure domain extension for UCEs via Wegman-Carter MACs [7]. (The case of reset-secure sources will be more involved and use new techniques.)

Assume, towards a contradiction, that $\overline{S}$ is *not* unpredictable; then there exists a strategy (not necessarily efficient) that given L and s^{in} and s^{out}, guesses one of the inner oracle queries of $\overline{S}$ with non-negligible probability ε, when $\overline{S}$'s oracle is a random function from $n+1$ to n bits. Imagine now that given $(L, s^{\mathsf{in}}, s^{\mathsf{out}})$ from $\overline{S}$, we *resample* an execution of $\overline{S}$ (which in particular means re-sampling the oracle used by it) consistent with outputting $(L, s^{\mathsf{in}}, s^{\mathsf{out}})$, and look at the inner oracle queries in this virtual, re-sampled execution. Then, one can show that the real and the virtual executions are likely to share an oracle query, with probability roughly at least ε^2, for our strategy to guess a query must be equally successful on the virtual execution.

We exploit this idea to build a predictor for the original source S, contradicting our hypothesis it is unpredictable. Note that S runs with a random permutation as its oracle, and produces leakage L. Imagine now we sample fresh seeds s^{in}, s^{out} for P, and for each permutation query by S defining an input-output pair (u, v), we define "fake" inputs x_0, x_1 from $x_0 \,\|\, x_1 = \mathsf{P}_{s^{\mathsf{in}}}(u)$ and

[4] A minor caveat is that we need indistinguishability even when s^{in} and s^{out} are revealed at the end of the interaction. We will show this to be true.

$x_3 \| x_2 = \mathsf{P}_{s^{\mathsf{out}}}(v)$. Then, the indistinguishability of the NR construction from a random permutation, and the construction of $\overline{S}$, implies that if we resample a virtual execution of S consistent with leakage L, and compute the resulting fake inputs using s^{in} and s^{out}, then the real and the re-sampled execution will share a fake input with probability approx. ε^2. The properties imposed on P then imply that with probability roughly ε^2 the real and the re-sampled execution must share the input (or output) of a permutation query. This leads naturally to a predictor that just re-samples an execution consistent with the leakage, and picks them as its prediction.

Reset-security. The case of reset-security is somewhat harder. Here we start from the premise that $\overline{S}$ is not reset-secure: Hence, there exists an adversary $\overline{R}$ which receives $L, s^{\mathsf{in}}, s^{\mathsf{out}}$ from $\overline{S}$, and can distinguish (with non-negligible advantage ε) being given access to the same random $f : \{0,1\}^{n+1} \to \{0,1\}^n$ used by $\overline{S}$ from being given access to an independent f'. From this, we would like to build an adversary R which receives L from S, and can distinguish the setting where R and S are given access to the same random permutation ρ, from a setting where they access independent permutations ρ, ρ'.

The challenge here is that we want to simulate $\overline{R}$ correctly, by using a permutation oracle ρ/ρ' rather than f/f'. To better see why this is tricky, say S is the source that queries its permutation oracle on a random $2n$-bit string u, obtaining output v, and leaks $L = (u, v)$. (This defines the corresponding $\overline{S}$.)[5] A clever $\overline{R}$ on input $(L = (u, v), s^{\mathsf{in}}, s^{\mathsf{out}})$ could do the following: It computes $x_0 \| x_1 \leftarrow \mathsf{P}(s^{\mathsf{in}}, u)$ and $x_3 \| x_2 \leftarrow \mathsf{P}(s^{\mathsf{out}}, v)$. Then, it queries x_1 to its oracle, and outputs 1 iff the output equals $x_0 \oplus x_2$. This should always be true when $\overline{R}$ accesses f, and almost never when it accesses f'.

The natural proof approach would now attempt to build R which runs $\overline{R}$ accessing a simulated oracle consistent with the NR construction on the permutation queries made by S. However, the problem is that generically R does not know which queries S has made. Previous work [37] handled this by requiring the construction to satisfy a weaker notion of indifferentiability, called CP-sequential indifferentiability, which essentially implies that there exists a simulator that can simulate f consistently by accessing ρ and ρ^{-1} only, and only needs to know the queries $\overline{R}$ makes to f. This would not work with NR and our $\overline{R}$, as the query x_1 is actually uniformly random, and the simulator would likely fail to set $x_0 \oplus x_2$ as the right output. This is why the approach of [37] ends up using the 5-round Feistel construction, as here $\overline{R}$'s attempt to evaluate the construction are readily detected, and answered consistently.

Our proof strategy via heavy-query sampling. Our main observation is that indifferentiability is an overkill in this setting. There is no reason $\overline{R}$ should act adversarially to the simulator. Even more so, we can use everything $\overline{R}$ knows, namely L, to our advantage! To do this, we use techniques borrowed

[5] The reader should not be confused: $\overline{S}$ is clearly not reset-secure, but remember we are in the setting of a proof by contradiction, so the reduction must work here, too.

from impossibility proofs in the random oracle model [4,26]. Namely, R, on input L from S, first performs a number of permutation queries which are meant to include all of S's likely queries to its oracle, at least when R and S are run with the same permutation oracle ρ. To do this, R samples executions of S consistent with L, and the partial knowledge of the oracle ρ acquired so far. Each time such a partial execution is sampled, all queries contained in it are made to ρ, and the process is repeated a number of times polynomial in $1/\varepsilon$. Then, R samples $s^{\mathsf{in}}, s^{\mathsf{out}}$, and internally defines an oracle $f : \{0,1\}^{n+1} \rightarrow \{0,1\}^n$ that will be used to simulate an execution of $\overline{R}^f(L, s^{\mathsf{in}}, s^{\mathsf{out}})$. To do this, R goes through all input-output pairs (u, v) for queries to ρ it has done while simulating executions of S,[6] and defines

$$f(0 \,\|\, x_1) \leftarrow x_0 \oplus x_2 \ , f(1 \,\|\, x_2) \leftarrow x_1 \oplus x_3,$$

where $x_0 \,\|\, x_1 \leftarrow \mathsf{P}_{s^{\mathsf{in}}}(u)$ and $x_3 \,\|\, x_2 \leftarrow \mathsf{P}_{s^{\mathsf{out}}}(v)$. Then, f is defined to be random on every other input (this can be simulated via lazy sampling). The final output of the simulated $\overline{R}$ is then R's final output.

The core of our proof will show that when S and R share access to ρ, then the probability that R's output is one is similar to that of $\overline{R}$ outputting one when it accesses the same oracle as $\overline{S}$. This will combine properties of the NR construction (allowing us to switch between f and ρ), and similar arguments as those used in [26] to prove that R ensures consistency on all queries that matter.[7]

1.5 Technical Overview – Correlation Intractability

Our approach towards achieving CI is based on the following blueprint. Let R be a relation which is evasive for permutations on $2n$-bit strings, and let π, σ be permutations sampled from some given distribution (this will be meant to be instantiated unconditionally below). Then, we create a new relation $R_{\pi,\sigma}$ such that $(u, v) \in R$ iff $(\pi(u), \sigma(v)) \in R_{\pi,\sigma}$. The hope is to show that *if* R is an evasive relation, then $R_{\pi,\sigma}$ is hard to satisfy for a given construction E which is only correlation-intractable for a subset of all evasive relations. Then, a new composed construction E' which outputs $\sigma^{-1}(E_s(\pi(u)))$ on input u would be correlation intractable for all evasive relations, since satisfying R for E' implies satisfying $R_{\pi,\sigma}$ for E.

Two-round Feistel. In our context, we instantiate E from a two-round Feistel network. That is, on input $x = x_0 \,\|\, x_1$, the two-round Feistel construction outputs $x_2 \,\|\, x_3$, where $x_2 \leftarrow f(0 \,\|\, x_1) \oplus x_0$ and $x_3 \leftarrow f(1 \,\|\, x_2) \oplus x_1$. In a model (as the one where we consider) where f is a random function to which the adversary

[6] The actual simulation will be slightly more involved, for the benefit of simplifying the analysis.

[7] We believe we could adapt our proof to use the better strategy of [4] to get slightly better concrete parameters, yet we found adapting it to our setting not immediate.

is given access, this construction is *not* correlation intractable. For instance, take the (unary) relation which is satisfied by all input-output pairs $(x_0 \,\|\, x_1, x_2 \,\|\, x_3)$ where $x_1 = x_2$. This is clearly evasive, but trivial to satisfy for two-round Feistel. Worse is possible with k-ary relations.

However, many relations *are* hard, even for two-round Feistel. Take for instance any relation R with the property that for all x_0, x_1, x_2, x_3, the number of x^*'s such that $(x^* \,\|\, x_1, x_2 \,\|\, x_3) \in R$ or $(x_0 \,\|\, x_1, x_2 \,\|\, x^*) \in R$ is at most $\delta \cdot 2^n$, for some negligible function δ. No adversary A making a polynomial number of queries to f will satisfy R, except with negligible probability. Indeed, when A queries (say) $y_2 \leftarrow f(1 \,\|\, x_2)$ for some x_2, the only chance to produce a pair that satisfies R is $y_1 \leftarrow f(0 \,\|\, x_1)$ was previously queried for some x_1, and additionally, $(x_2 \oplus y_1 \,\|\, x_1,\ x_2 \,\|\, x_1 \oplus y_2) \in R$. But because y_2 is being set randomly, $x_1 \oplus y_2$ is also random, this can only hold with probability at most δ by our assumption on δ. Thus, the probability that this pair satisfies R is negligible, and the union bound over all pairs of queries shows A is unlikely to *ever* satisfy R.

Where does amplification come from? Let R be a unary evasive relation. Now, imagine, again for the sake of an oversimplified illustration, that π and σ are random permutations. Then, we want to show that with high probability over the choice of π and σ, the relation $R_{\pi,\sigma}$ is hard for two-round Feistel, even if the adversary learns the entire description of π and σ. Indeed, find any x_0, x_1, x_2, fix π, and fix $u = \pi^{-1}(x_0 \,\|\, x_1)$. Because R is evasive, there exists at most $\delta \cdot 2^{2n}$ v's (for some negligible δ) such that $(u, v) \in R$ – call the set of such v's R_u. Because σ is random, the probability that $\sigma^{-1}(x_2 \,\|\, z) \in R_u$ is at most δ, and thus the *expected* number of z such that $(x_0 \,\|\, x_1, x_2 \,\|\, z) \in R_{\pi,\sigma}$ is at most $\delta \cdot 2^n$, by linearity of expectation. A concentration bound will show that the probability that we are far from this expectation is indeed small, say at most 2^{-4n}. Taking a union bound over all x_0, x_1, x_2 shows that the probability this is true for any x_0, x_1, x_2 is at most 2^{-n}. (The symmetric argument when fixing x_1, x_2, x_3 can be handled analogously.) Thus, we have just argued that with high probability over the choice of π and σ, $R_{\pi,\sigma}$ is hard for two-round Feistel!

Challenges. But obviously, this is not very useful– random permutations π, σ are inefficient to sample and describe. Also, the above result only holds for unary relations, and it is interesting to extend this to k-ary relations.

Our first insight is that the above argument only requires a bounded degree of randomness, and that (almost) t-wise independent permutations for a sufficiently small t are sufficient. We prove this using techniques for bounding sums of random variables with bounded independence [10,36], though this will require significant adaptation because almost t-wise independent permutations do not quite produce outputs which are t-wise independent, as they are required to be distinct, and also, only approximate a random permutation. We will instantiate these by using Feistel networks with sufficiently many rounds, and t-wise independent round functions, using bounds from [24]. In fact, for the case of unary relations, we will show that we can instantiate these permutations from one single Feistel round with a 10-wise independent round function.

Moving on to k-ary relations presents even more challenges. Our approach is inherently combinatorial, whereas evasiveness is defined indirectly through the inability of an adversary to win a security game. For this reason, our result will only deal with relations R that satisfy a more structured notion of evasiveness, which we refer to as *strongly* evasive. Most relations of interest that we are aware of are strongly evasive, but evasiveness does not always imply strong evasiveness. *However*, as a result of independent interest, we show that strong evasiveness and evasiveness are related, and asymptotically equivalent when k is a constant.

2 Preliminaries

NOTATIONAL PRELIMINARIES. Throughout this paper, we denote by $\mathsf{Funcs}(X, Y)$ the set of functions $X \to Y$, and in particular use the shorthand $\mathsf{Funcs}(m, n)$ whenever $X = \{0,1\}^m$ and $Y = \{0,1\}^n$. We also denote by $\mathsf{Perms}(X)$ the set of permutations on the set X, and analogously, $\mathsf{Perms}(n)$ denotes the special case where $X = \{0,1\}^n$. For $n \in \mathbb{N}$, we let $[n]$ denote the set $\{1, \ldots, n\}$.

Our security definitions and proofs will often use games, as formalized by Bellare and Rogaway [9]. Typically, our games will have boolean outputs – that is, either true or false – and we use the shorthand $\mathsf{Pr}[\mathsf{G}]$ to denote the probability that a certain game outputs the value true, or occasionally 1 (when the output is binary, rather than boolean). Most results in this paper will be concrete, but natural asymptotic statements can be made by allowing all parameters to be functions of the security parameter.

A *function family* with input set X and output set Y is a pair of algorithms $\mathsf{F} = (\mathsf{F.Kg}, \mathsf{F.Eval})$, where the randomized *key (or seed) generation algorithm* $\mathsf{F.Kg}$ outputs a seed s, and the deterministic *evaluation algorithm* $\mathsf{F.Eval}$ takes as inputs a valid seed s and an input $x \in X$, and returns $\mathsf{F.Eval}(s, x) \in Y$. If $X = \{0,1\}^m$ and $Y = \{0,1\}^n$, we say that F is a family of functions from m-bits to n-bits. We usually write $\mathsf{F}(s, \cdot) = \mathsf{F.Eval}(s, \cdot)$. A *permutation family* $\mathsf{P} = (\mathsf{P.Kg}, \mathsf{P.Eval})$ on n bits is the special case where $X = \{+, -\} \times \{0,1\}^n$ and $Y = \{0,1\}^n$, and for every s, there exists a permutation π_s such that $\mathsf{P.Eval}(s, (+, x)) = \pi_s(x)$ and $\mathsf{P.Eval}(s, (-, y)) = \pi_s^{-1}(y)$. We usually write $\mathsf{P}(s, \cdot) = \mathsf{P}(s, (+, \cdot))$ and $\mathsf{P}^{-1}(s, \cdot) = \mathsf{P}(s, (-, \cdot))$.

2.1 UCEs and psPRPs

We review the UCE notion introduced in [6], and the psPRP notion [37]. As explained in the latter work, they can be seen as instantiations of a general paradigm. Yet, we consider separate security games for better readability.

Concretely, let H be function family from m-bits to n-bits. Let S be an adversary called the *source* and D an adversary called the *distinguisher*. We associate with them the game $\mathsf{UCE}^{S,D}_{m,n,\mathsf{H}}$ depicted in Fig. 1. For a family E of permutations on n-bits, the psPRP-security game $\mathsf{psPRP}^{S,D}_{n,\mathsf{E}}$ differs in that $\mathcal{O}$

```
MAIN $\boxed{\mathsf{UCE}^{S,D}_{m,n,\mathsf{H}}}$, $\boxed{\mathsf{psPRP}^{S,D}_{n,\mathsf{E}}}$:
$(1^r, t) \xleftarrow{\$} S(\varepsilon)$, $b \xleftarrow{\$} \{0,1\}$
$s_1, \ldots, s_r \xleftarrow{\$} \boxed{\mathsf{H.Kg}}\ \boxed{\mathsf{E.Kg}}$
$\boxed{f_1, \ldots, f_r \xleftarrow{\$} \mathsf{Funcs}(m, n)}$
$\boxed{\rho_1, \ldots, \rho_r \xleftarrow{\$} \mathsf{Perms}(n)}$
$L \xleftarrow{\$} S^{\mathcal{O}}(t)$
$b' \xleftarrow{\$} D(s_1, \ldots, s_r, L)$
return $b' = b$
```

```
ORACLE $\mathcal{O}(i, x)$:    // $\mathsf{UCE}^{S,D}_{m,n,\mathsf{H}}$
if $b = 1$ then return $\mathsf{H}(s_i, x)$
else return $f_i(x)$

ORACLE $\mathcal{O}(i, (\sigma, x))$:    // $\mathsf{psPRP}^{S,D}_{n,\mathsf{E}}$
if $b = 1$ then
    if $\sigma = +$ then return $\mathsf{E}(s_i, x)$
    else return $\mathsf{E}^{-1}(s_i, x)$
else
    if $\sigma = +$ then return $\rho_i(x)$
    else return $\rho_i^{-1}(x)$
```

Fig. 1. Games to define UCE and psPRP security. Here, S is the source and D is the distinguisher. Boxed statements are only executed in the corresponding game.

allows for inverse queries, and the ideal object is a random permutation. The corresponding advantage metrics for an (S, D) are defined as

$$\begin{aligned} \mathsf{Adv}^{\mathsf{uce}}_{m,n,\mathsf{H}}(S, D) &= 2\Pr\left[\mathsf{UCE}^{S,D}_{m,n,\mathsf{H}}\right] - 1 \\ \mathsf{Adv}^{\mathsf{psprp}}_{n,\mathsf{E}}(S, D) &= 2\Pr\left[\mathsf{psPRP}^{S,D}_{n,\mathsf{E}}\right] - 1. \end{aligned} \tag{1}$$

Note that we adopt the multi-key versions of UCE and psPRP security, as they are the most general, and they are not known to follow from the single-key case. Our treatment scales down to the single-key version by forcing the source to always choose $r = 1$.

We say that H is UCE-secure for a class of sources $\mathcal{S}$ if $\mathsf{Adv}^{\mathsf{uce}}_{m,n,\mathsf{H}}(S, D)$ is negligible for all PPT D and all sources $S \in \mathcal{S}$. Similarly, E is psPRP secure for $\mathcal{S}$ if $\mathsf{Adv}^{\mathsf{psprp}}_{n,\mathsf{E}}(S, D)$ is negligible for all PPT D and all sources $S \in \mathcal{S}$ It is known that $\mathcal{S}$ cannot contain all PPT algorithms for security to be attainable. Next, we discuss two important classes of restrictions – unpredictable and reset-secure sources – considered in the literature [6,7,37].

Unpredictable Sources. Let S be a source and P be an adversary called the *predictor*. We associate with them games $\mathsf{f\text{-}Pred}^{P}_{m,n,S}$ and $\mathsf{p\text{-}Pred}^{P}_{n,S}$ of Fig. 2 which capture the fact that P cannot predict any of the queries of S (or their inverses), when the latter interacts with a random function from m bits to n bits, or respectively a random permutation on n-bit strings. The corresponding advantage metrics are

$$\mathsf{Adv}^{\mathsf{f\text{-}pred}}_{m,n,S}(P) = \Pr\left[\mathsf{f\text{-}Pred}^{P}_{m,n,S}\right], \ \mathsf{Adv}^{\mathsf{p\text{-}pred}}_{n,S}(P) = \Pr\left[\mathsf{p\text{-}Pred}^{P}_{n,S}\right]. \tag{2}$$

We say S is *statistically unpredictable* if $\mathsf{Adv}^{\mathsf{f\text{-}pred}}_{m,n,S}(P)$ (respectively, $\mathsf{Adv}^{\mathsf{p\text{-}pred}}_{n,S}(P)$) is negligible for all predictors P outputting a set Q' of polynomial size.

```
MAIN $\boxed{\mathsf{f\text{-}Pred}^P_{m,n,S}}$, $\boxed{\mathsf{p\text{-}Pred}^P_{n,S}}$:
$Q \leftarrow \emptyset$
$(1^r, t) \xleftarrow{\$} S(\varepsilon)$
$\boxed{f_1, \ldots, f_r \xleftarrow{\$} \mathsf{Funcs}(m, n)}$
$\boxed{\rho_1, \ldots, \rho_r \xleftarrow{\$} \mathsf{Perms}(n)}$
$L \xleftarrow{\$} S^{\mathcal{O}}(t)$
$Q' \xleftarrow{\$} P(1^r, L)$
return $(Q \cap Q' \neq \emptyset)$

ORACLE $\mathcal{O}(i, x)$: // $\mathsf{f\text{-}Pred}^P_{m,n,S}$
$Q \leftarrow Q \cup \{(i, x)\}$
return $f_i(x)$

ORACLE $\mathcal{O}(i, (\sigma, x))$: // $\mathsf{p\text{-}Pred}^P_{n,S}$
if $\sigma = +$ then $y \leftarrow \rho_i(x)$
else $y \leftarrow \rho_i^{-1}(x)$
$Q \leftarrow Q \cup \{(i, x), (i, y)\}$
return $y$
```

```
MAIN $\boxed{\mathsf{f\text{-}Reset}^R_{m,n,S}}$,$\boxed{\mathsf{p\text{-}Reset}^R_{n,S}}$:
$\mathsf{done} \leftarrow \mathsf{false}$; $(1^r, t) \xleftarrow{\$} S(\varepsilon)$
$\boxed{f_1^0, f_1^1, \ldots, f_r^0, f_r^1 \xleftarrow{\$} \mathsf{Funcs}(m, n)}$
$\boxed{\rho_1^0, \rho_1^1, \ldots, \rho_r^0, \rho_r^1 \xleftarrow{\$} \mathsf{Perms}(n)}$
$L \xleftarrow{\$} S^{\mathcal{O}}(t)$; $\mathsf{done} \leftarrow \mathsf{true}$
$b \xleftarrow{\$} \{0, 1\}$; $b' \xleftarrow{\$} R^{\mathcal{O}}(1^r, L)$
return $b' = b$

ORACLE $\mathcal{O}(i, x)$: // $\mathsf{f\text{-}Reset}^R_{m,n,S}$
if $\neg\mathsf{done}$ then return $f_i^0(x)$
else return $f_i^b(x)$

ORACLE $\mathcal{O}(i, (\sigma, x))$: // $\mathsf{p\text{-}Reset}^R_{n,S}$
if $\neg\mathsf{done}$ then
    if $\sigma = +$ then return $\rho_i^0(x)$
    else return $\rho_i^{0\,-1}(x)$
else
    if $\sigma = +$ then return $\rho_i^b(x)$
    else return $\rho_i^{b\,-1}(x)$
```

Fig. 2. Games to define unpredictability (left) and reset-security (right) of sources. Here, S is the source, P is the predictor and R is the reset-adversary. Boxed statements are only executed in the corresponding game.

An analogous notion of computational unpredictability can be defined, but it is unachievable if IO exists [14], and is usually not needed for applications. We also note that what we formalize here is the notion of *simple* unpredictability – P is not permitted to query the underlying primitive. The notion was proved equivalent (asymptotically) for UCEs [6] to a version where we give P access to the primitive. A similar proof follows for psPRPs. (We omit it due to lack of space.)

Reset-secure Sources. Let S be a source and R be an adversary called the reset-adversary. We associate to them the games $\mathsf{f\text{-}Reset}^R_{m,n,S}$ and $\mathsf{p\text{-}Reset}^R_{n,S}$ of Fig. 2 which formalize the reset-security of S against a random function and a random permutation, respectively. The idea here is that R should not be able to tell apart whether S is accessing the same set of oracles it accesses, or not. This is captured via the advantage metrics

$$\mathsf{Adv}^{\mathsf{f\text{-}reset}}_{m,n,S}(R) = 2\Pr\left[\mathsf{f\text{-}Reset}^R_{m,n,S}\right] - 1, \quad \mathsf{Adv}^{\mathsf{p\text{-}reset}}_{n,S}(R) = 2\Pr\left[\mathsf{p\text{-}Reset}^R_{n,S}\right] - 1.$$

We say S is *statistically reset-secure* if the corresponding advantage is negligible for all reset-adversaries R making a polynomial number of *queries* to their oracle, but which are otherwise computationally unrestricted. It is known that a (statistically) unpredictable source is (statistically) reset-secure, for both UCEs [6] and psPRPs [37]. The converse is not true – S may query a fixed known input, and let L be the empty string. S is reset-secure in the strongest sense, while being easily predictable.

2.2 Evasive Relations, Correlation Intractability

In the following, a k-ary relation R over $X \times Y$ is a set of subsets $S \subseteq X \times Y$, where $1 \leq |S| \leq k$.[8] We are going to consider relations which are *evasive* with respect to a random permutation.

Evasive relations. Given a relation R over $\{0,1\}^m \times \{0,1\}^m$, we consider the following advantage metric, involving an adversary A:

$$\mathsf{Adv}^{\mathsf{evp}}_{R,m}(A) = \Pr_{\pi}\left[S \xleftarrow{\$} A^{\pi,\pi^{-1}} : S \in R \wedge \forall (u,v) \in S : \pi(u) = v\right],$$

where $\pi \xleftarrow{\$} \mathsf{Perms}(m)$. We say that a relation R is (q,δ)-evasive for a random permutation if $\mathsf{Adv}^{\mathsf{evp}}_{R,m}(A) \leq \delta$ for all adversaries making q queries.

Correlation intractability. Let M^f be a permutation family on m-bits which makes oracle calls to a function f from n bits to ℓ bits, to be modeled as a random function. Let R be a k-ary relation. Let A be any (possibly unbounded) adversary. We associate to A, M and R the following cri-advantage metric:

$$\mathsf{Adv}^{\mathsf{cri}}_{R,\mathsf{M}}(A) = \Pr_{s,f}\left[S \xleftarrow{\$} A^f(s) \ : \ S \in R \ \wedge \ \forall (u,v) \in S : \mathsf{M}^f(s,u) = v\right],$$

where $f \xleftarrow{\$} \mathsf{Funcs}(n,\ell)$ and $s \xleftarrow{\$} \mathsf{M.Kg}$.

3 Public-Seed Pseudorandomness of Naor-Reingold

This section revisits the Naor-Reingold construction [33] in the public-seed setting. We prove that it transforms a UCE into a psPRP, for both unpredictable (Sect. 3.2) and reset-secure sources (Sect. 3.3). Before turning to these results, however, Sect. 3.1 reviews the construction and proves a strong statement about its indistinguishability.

[8] We think of the elements as *sets*, rather than tuples – this is because looking ahead, it only makes sense in the context of correlation intractability to consider symmetric relation, as an adversary can always re-order its outputs.

3.1 The NR Construction and Its Indistinguishability

Let P be a permutation family on the $2n$-bit strings. We say that P is *α-right-universal* if $\mathsf{Pr}_{s \xleftarrow{\$} \mathsf{P.Kg}}[\mathsf{P}_1(s, u) = \mathsf{P}_1(s, u')] \leq \alpha$ for all distinct $u, u' \in \{0,1\}^{2n}$, where P_1 denote the second n-bit half of the output of P. Note that a pairwise-independent permutation is a good candidate of P, but a simpler approach is to employ one-round of Feistel with a pairwise independent hash function H as the round function, i.e., $\mathsf{P}(s, (u_0, u_1)) = (u_1, \mathsf{H}(s, u_1) \oplus u_0)$.

The Naor-Reingold (NR) Construction. Let H be a function family from $n+1$ bits to n bits. We define the permutation family $\mathsf{NR} = \mathsf{NR}[\mathsf{P}, \mathsf{H}]$ on the $2n$-bit strings, where $\mathsf{NR.Kg}$ outputs $(s, s^{\mathsf{in}}, s^{\mathsf{out}})$ such that $s \xleftarrow{\$} \mathsf{H.Kg}$ and $s^{\mathsf{in}}, s^{\mathsf{out}} \xleftarrow{\$} \mathsf{P.Kg}$. Further, forward evaluation proceeds as follows (the inverse is obvious):

```
Proc. $\mathsf{NR}((s, s^{\mathsf{in}}, s^{\mathsf{out}}), U)$:
$x_0 \,\|\, x_1 \leftarrow \mathsf{P}(s^{\mathsf{in}}, U)$, $x_2 \leftarrow \mathsf{H}(s, 0 \,\|\, x_1) \oplus x_0$,
$x_3 \leftarrow \mathsf{H}(s, 1 \,\|\, x_2) \oplus x_1$, $V \leftarrow \mathsf{P}^{-1}(s^{\mathsf{out}}, x_3 \,\|\, x_2)$,
return $V$
```

Naor and Reingold [33] proved that the NR construction with random round functions is indistinguishable from a random permutation under chosen ciphertext attacks. We will need a stronger result, which we prove here, that this is true even when the seed of P is given to the adversary after it stops making queries, and when the adversary can make queries to multiple instances of the construction. It will be convenient to re-use the notation already in place for the psPRP framework, and we denote by $\mathsf{Adv}^{\mathsf{psprp}^+}_{2n,\mathsf{NR}[\mathsf{P},\mathsf{F}]}(S, D)$ the advantage obtained by (S, D) in the $\mathsf{psPRP}^{S,D}_{2n,\mathsf{NR}[\mathsf{P},\mathsf{F}]}$ game, with the modification that D is *not* given the seed for F, only the seeds used by the permutation P.

Proposition 1 (Indistinguishability of the NR construction). *Let* $\mathsf{F} = \mathsf{F}[n+1, n]$ *be the family of all functions from* $n+1$ *to* n *bits, equipped with the uniform distribution. Further, let* P *be* α*-right-universal. For all* S, D*, where* S *makes* q *queries,* $\mathsf{Adv}^{\mathsf{psprp}^+}_{2n,\mathsf{NR}[\mathsf{P},\mathsf{F}]}(S, D) \leq q^2 \cdot \left(2\alpha + \frac{1}{2^{2n}}\right)$.

The proof of Proposition 1 can be found in the full version [38, App. A.1].

3.2 The Case of Unpredictable Sources

We first prove that the NR construction transforms a UCE function family for statistically unpredictable sources into a psPRP for statistically unpredictable sources. Our proof uses a technique inspired from that of Bellare et al. [7], given originally in the setting of UCE domain extension. Concretely, we prove the following.

Theorem 1 (NR security for unpredictable sources). *Let* P *be a* α*-right universal family of permutations on* $2n$*-bit strings. Let* H *be a family of functions* *from $n+1$ bits to n bits. Then, for all distinguishers D and sources S making overall q queries to their oracle, there exists $\overline{D}$ and $\overline{S}$ such that*

$$\mathsf{Adv}^{\mathsf{psprp}}_{2n,\mathsf{NR[P,H]}}(S,D) \leq \mathsf{Adv}^{\mathsf{uce}}_{n+1,n,\mathsf{H}}(\overline{S},\overline{D}) + q^2\left(2\alpha + \frac{1}{2^{2n}}\right). \tag{3}$$

Here, $\overline{D}$ and D are roughly as efficient, and $\overline{S}$ and S are similarly as efficient. In particular, $\overline{S}$ makes $2q$ queries. Moreover, for every predictor $\overline{P}$, there exists a predictor P such that

$$\mathsf{Adv}^{\mathsf{f\text{-}pred}}_{n+1,n,\overline{S}}(\overline{P}) \leq q^2 \cdot \left(2\alpha + \frac{1}{2^{2n}}\right) + p \cdot \sqrt{2q^2\alpha + \mathsf{Adv}^{\mathsf{p\text{-}pred}}_{2n,S}(P)}, \tag{4}$$

where p is a bound on the size of the set output by $\overline{P}$.

The asymptotic interpretation is that if $n = \omega(\log(\lambda))$ and α is negligible, if S is (statistically) unpredictable, then so is $\overline{S}$. Further, if H is a UCE for all unpredictable sources, then NR is a psPRP for all statistically unpredictable sources.

We stress that the predictor P built in the proof does not preserve the efficiency of $\overline{P}$, which is not a problem, as we only consider statistical notions. While we do not elaborate in the proof, it turns out that the running time of P is *exponential* in the length of S's leakage, thus the statement carries over to computational unpredictability if $L = O(\log\lambda)$.

Proof. We first consider three games, $\mathsf{G}_0, \mathsf{G}_1$, and G_2. Game G_0 is the game $\mathsf{psPRP}^{S,D}_{2n,\mathsf{NR[P,F]}}$ in the case $b=1$, and modified to return true if $b'=1$. Game G_2 is the game $\mathsf{psPRP}^{S,D}_{2n,\mathsf{NR[P,F]}}$ in the case $b=0$, and modified to return true if $b'=1$. The intermediate game G_1 is obtained by modifying G_0 as follows: Initially, r random functions $f_1,\ldots,f_r \xleftarrow{\$} \mathsf{Funcs}(n+1,n)$ are sampled, and when evaluating the NR construction within $\mathcal{O}$ queries, the evaluation of $\mathsf{H}(s_i, b\,\|\,x)$ is replaced by an evaluation of the random function $f_i(b\,\|\,x)$. Then,

$$\mathsf{Adv}^{\mathsf{psprp}}_{2n,\mathsf{NR[P,H]}}(S,D) = (\mathsf{Pr}\left[\mathsf{G}_0\right] - \mathsf{Pr}\left[\mathsf{G}_1\right]) + (\mathsf{Pr}\left[\mathsf{G}_1\right] - \mathsf{Pr}\left[\mathsf{G}_2\right]).$$

We can directly get $\mathsf{Pr}\left[\mathsf{G}_1\right] - \mathsf{Pr}\left[\mathsf{G}_2\right] \leq q^2\left(2\alpha + \frac{1}{2^{2n}}\right)$ as a corollary of Proposition 1, since neither of G_1 and G_2 uses the seeds generated by $\mathsf{H.Kg}$.

Going on, let us consider the new source $\overline{S}$ which simulates an execution of S, and uses access to an oracle $\mathcal{O}(i,X)$, implementing for each i a function from $n+1$ bits to n bits, to internally simulate the round functions NR construction used to answer S's queries. A formal description is in Fig. 3. Also consider the distinguisher $\overline{D}$ such that

$$\overline{D}(L' = (L, \boldsymbol{s}^{\mathsf{in}}, \boldsymbol{s}^{\mathsf{out}}), \boldsymbol{s}) = D(L, (\boldsymbol{s}, \boldsymbol{s}^{\mathsf{in}}, \boldsymbol{s}^{\mathsf{out}})),$$

where $\boldsymbol{s} = (s_1,\ldots,s_r)$, $\boldsymbol{s}^{\mathsf{in}} = (s^{\mathsf{in}}_1,\ldots,s^{\mathsf{in}}_r)$, and $\boldsymbol{s}^{\mathsf{out}} = (s^{\mathsf{out}}_1,\ldots,s^{\mathsf{out}}_r)$ Therefore, G_0 and G_1 behave exactly as $\mathsf{UCE}^{\overline{S},\overline{D}}_{n+1,n,\mathsf{H}}$ with challenge bits $b=1$ and $b=0$,

```
Proc. $\overline{S}(\varepsilon)$:
$(1^r, t) \xleftarrow{\$} S(\varepsilon)$
return $(1^r, (1^r, t))$

Proc. $\overline{S}^{\mathcal{O}}(1^r, t)$:
$s_1^{\mathsf{in}}, s_1^{\mathsf{out}}, \ldots, s_r^{\mathsf{in}}, s_r^{\mathsf{out}} \xleftarrow{\$} \mathsf{P.Kg}$
$L \xleftarrow{\$} S^{\overline{\mathcal{O}}}(t)$
return $(L, s_1^{\mathsf{in}}, s_1^{\mathsf{out}}, \ldots, s_r^{\mathsf{in}}, s_r^{\mathsf{out}})$

Proc. $\overline{\mathcal{O}}(i, (\sigma, U))$:
if $\sigma = +$ then
    $x_0 \,\|\, x_1 \leftarrow \mathsf{P}(s_i^{\mathsf{in}}, U)$
    $x_2 \leftarrow \mathcal{O}(i, 0 \,\|\, x_1) \oplus x_0$, $x_3 \leftarrow \mathcal{O}(i, 1 \,\|\, x_2) \oplus x_1$
    $V \leftarrow \mathsf{P}^{-1}(s_i^{\mathsf{out}}, x_3 \,\|\, x_2)$
else
    $x_3 \,\|\, x_2 \leftarrow \mathsf{P}(s_i^{\mathsf{out}}, U)$
    $x_1 \leftarrow \mathcal{O}(i, 1 \,\|\, x_2) \oplus x_3$, $x_0 \leftarrow \mathcal{O}(i, 0 \,\|\, x_1) \oplus x_2$
    $V \leftarrow \mathsf{P}^{-1}(s_i^{\mathsf{in}}, x_0 \,\|\, x_1)$
return $V$
```

Fig. 3. The source $\overline{S}$ in the proof of Theorems 1 and 2.

respectively, with the only difference of outputting true whenever the distinguisher's output is $b' = 1$. Consequently, $\mathsf{Adv}^{\mathsf{uce}}_{n+1,n,\mathsf{H}}(\overline{S}, \overline{D}) = \mathsf{Pr}\left[\mathsf{G}_0\right] - \mathsf{Pr}\left[\mathsf{G}_1\right]$.

The remainder of the proof relates the unpredictability of S and that of $\overline{S}$, establishing (4) in the theorem statement. For lack of space, the argument is deferred to the full version [38, App. A.2]. □

3.3 The Case of Reset-Secure Sources

Theorem 1's importance stems mostly from the fact that it establishes the equivalence of psPRPs and UCEs for the case of (statistically) unpredictable sources. The question was left open in [37]. Many applications (e.g., instantiating the permutation within sponges, or any other indifferentiable hash construction) however require the stronger notion of reset-security. For this, [37] show that the five-round Feistel construction suffices, using a weaker variant of indifferentiability, and left open the question of whether four-rounds suffice.

We do better here: we prove that the NR construction transforms a UCE for statistically reset-secure sources into a psPRP for the same class of sources. The proof starts as the one of Theorem 1, but then shows that the source $\overline{S}$ built therein is in fact statistically reset-secure whenever S is. This step will resort to a variant of the heavy-query sampling method of Impagliazzo and Rudich [26] to simulate a random oracle from the leakage which captures "relevant correlations" with what is learnt by the source.

Theorem 2 (NR security for reset-secure sources). *Let P be a α-right universal family of permutations on $2n$-bit strings, and let H be a function family from $n+1$ bits to n bits. Then, for all distinguishers D and sources S making overall q queries to their oracle, there exists $\overline{D}$ and $\overline{S}$ such that*

$$\mathsf{Adv}^{\mathsf{psprp}}_{2n,\mathsf{NR}[\mathsf{P},\mathsf{H}]}(S, D) \leq \mathsf{Adv}^{\mathsf{uce}}_{n+1,n,\mathsf{H}}(\overline{S}, \overline{D}) + q^2\left(2\alpha + \frac{1}{2^{2n}}\right). \tag{5}$$

Here, $\overline{D}$ and D are roughly as efficient, and $\overline{S}$ and S are similarly as efficient. In particular, $\overline{S}$ makes $2q$ queries. Moreover, for every reset-adversary $\overline{R}$ making p queries, there exists a reset-adversary R such that

$$\mathsf{Adv}^{\mathsf{f\text{-}reset}}_{n+1,n,\overline{S}}(\overline{R}) \leq 2\mathsf{Adv}^{\mathsf{p\text{-}reset}}_{2n,S}(R) + 4\left(q + \frac{8qp^2}{\varepsilon}\ln(4p/\varepsilon)\right)^2\left(2\alpha + \frac{1}{2^{2n}}\right), \quad (6)$$

where $\varepsilon := \mathsf{Adv}^{\mathsf{f\text{-}reset}}_{n+1,n,\overline{S}}(\overline{R})$. In particular, R makes $4qp^2/\varepsilon \cdot \ln(4p/\varepsilon)$ queries to its oracle.

Asymptotically, (6) implies that if $\overline{R}$ exists making $p = \mathsf{poly}(\lambda)$ queries, and achieving non-negligible advantage ε, then R makes also a polynomial number of queries, and achieves non-negligible advantage, as long as α is negligible, and $n = \omega(\log \lambda)$. Thus, reset-security of R yields reset-security of $\overline{R}$.

We also believe that the technique of Barak and Mahmoody [4] can be used to reduce the $8qp^2/\varepsilon$ term to $O(qp)/\varepsilon$. We did not explore this avenue here, as the proof approach of [26] is somewhat easier to adapt to our setting.

Proof. The setup of the proof is identical to that in Theorem 1, in particular the construction of $\overline{S}$ from S (and of $\overline{D}$ from D.) The difference is in relating the reset-security of S and $\overline{S}$. In particular, let

$$\varepsilon := \mathsf{Adv}^{\mathsf{f\text{-}reset}}_{n+1,n,\overline{S}}(\overline{R}) = \Pr\left[\mathsf{f\text{-}Reset}^{\overline{R}}_{n+1,n,\overline{S}} \,\middle|\, b=0\right] - \Pr\left[\neg\mathsf{f\text{-}Reset}^{\overline{R}}_{n+1,n,\overline{S}} \,\middle|\, b=1\right].$$

The RHS is the difference of the probabilities of $\overline{R}$ outputting 0 in the cases $b = 0$ and $b = 1$ respectively. We are going to build a new adversary R against S which satisfies (6). We assume without loss of generality that $\overline{R}$ is deterministic, and makes *exactly* p *distinct* queries to its oracle.

We start the proof with some game transitions that will lead naturally to the definition of the adversary R. Formal descriptions are found in our full version [38, Figs. 7 and 8] – our description here is self-contained.

The initial game G_1 is simply $\mathsf{f\text{-}Reset}^{\overline{R}}_{\overline{S}}$ with the bit $b = 0$, i.e., $\overline{S}$ and $\overline{R}$ access the *same* functions $f_1, \ldots, f_r$ here. Further, G_1 returns true iff $\overline{R}$ returns 0. Thus, $\Pr[\mathsf{G}_1] = \Pr\left[\mathsf{f\text{-}Reset}^{\overline{R}}_{\overline{S}} \,\middle|\, b=0\right]$. Game G_2 slightly changes G_1: It keeps track (in a set Q_{P}) of the triples (i, U, V) describing $\overline{\mathcal{O}}$ queries made by the simulated S within $\overline{S}$; i.e., either S queried $(i, (+, U))$, and obtained V, or queries $(i, (-, V))$, and obtained U. After $\overline{S}$ terminates with leakage $(L, \boldsymbol{s})$, where $\boldsymbol{s} = (s^{\mathsf{in}}_1, s^{\mathsf{out}}_1, \ldots, s^{\mathsf{in}}_r, s^{\mathsf{out}}_r)$, for every $(i, U, V) \in Q_{\mathsf{P}}$ we compute $x_0 \,\|\, x_1 \leftarrow \mathsf{P}(s^{\mathsf{in}}_i, U)$ and $x_3 \,\|\, x_2 \leftarrow \mathsf{P}(s^{\mathsf{out}}_i, V)$, and define table entries

$$T[i, 0 \,\|\, x_1] \leftarrow x_0 \oplus x_2\ , \quad T[i, 1 \,\|\, x_2] \leftarrow x_1 \oplus x_3\ .$$

For later reference, we denote by X the set of pairs (i, x) for which we set $T[i, x]$ using Q_{P} and $\boldsymbol{s}$. We then run $\overline{R}(L, \boldsymbol{s})$, and answer its oracle queries (i, x) using $T[i, x]$. If the entry is undefined, then we return a random value. (As we assumed all of $\overline{R}$'s queries are distinct, we do not need to remember the output.) As before, G_2 outputs true iff $\overline{R}$ outputs 0.

Note that we always have $T[i, x] = f_i(x)$ for very (i, x) such that $f_i(x)$ was queried by $\overline{S}$, and re-sampling values un-queried by $\overline{S}$ upon $\overline{R}$'s queries does not change the distribution of $\overline{R}$'s output, and hence $\mathsf{Pr}[\mathsf{G}_1] = \mathsf{Pr}[\mathsf{G}_2]$.

The intersection sampler. The game G_3 generates a surrogate for Q_{P}. This is the output of an algorithm Sam which, after $\overline{S}$ terminates with output $(L, \boldsymbol{s})$, takes as input the leakage L (crucially, not $\boldsymbol{s}$!) and an iteration parameter $\eta = 4p/\varepsilon \ln(4p/\varepsilon)$ (we let also $\tau = p \cdot \eta$). Sam queries the very same $\overline{\mathcal{O}}$ implemented by $\overline{S}$ to answer S's queries (which internally simulates the NR construction using $\overline{S}$'s own oracle), and returns a set $\widetilde{Q}_{\mathsf{P}}$ of 4-tuples (i, U, V, j) such that $j \in [p]$, and (i, U, V) is such that $\overline{\mathcal{O}}(i, (+, U))$ would return V (or equivalently $\overline{\mathcal{O}}(i, (-, V))$ would return U). Internally, Sam will make calls to another (randomized) subprocedure $\mathcal{Q}$ which takes as input the leakage L, as well as a set Q of tuples (i, U, V, j) consistent with $\overline{\mathcal{O}}$, and returns a set Δ of at most q tuples (i, U, V), which are not necessarily consistent with $\overline{\mathcal{O}}$. We will specify in detail later below what $\mathcal{Q}$ exactly does, as some further game transitions will come handy to set up proper notation. For now, a generic understanding will suffice. In particular, given such $\mathcal{Q}$, Sam operates as in Fig. 4. As we can see, for each $(i, U, V, j) \in \widetilde{Q}_{\mathsf{P}}$, j indicates the outer iteration in which this query was added to $\widetilde{Q}_{\mathsf{P}}$. Using this information, for every $j \in [p]$, and every 4-tuple (i, U, V, j) we compute $x_0 \,\|\, x_1 \leftarrow \mathsf{P}(s_i^{\mathsf{in}}, U)$ and $x_3 \,\|\, x_2 \leftarrow \mathsf{P}(s_i^{\mathsf{out}}, V)$, define

$$\widetilde{T}[i, 0 \,\|\, x_1] \leftarrow x_0 \oplus x_2 \ , \ \ \widetilde{T}[i, 1 \,\|\, x_2] \leftarrow x_1 \oplus x_3,$$

and add $(i, 0 \,\|\, x_1), (i, 1 \,\|\, x_2)$ to the set $\widetilde{X}^j$. A for now irrelevant caveat is that if one of the entries in $\widetilde{T}$ is already set, then we do not overwrite it.[9]

```
ALGORITHM $\mathsf{Sam}^{\overline{\mathcal{O}}}(L, \tau)$ :
$\widetilde{Q}_{\mathsf{P}} \leftarrow \emptyset$
for $j = 1$ to $p$ do
  for $k = 1$ to $\eta$ do
    $\Delta_{j,k} \leftarrow \mathcal{Q}(L, \widetilde{Q}_{\mathsf{P}})$
    for all $(i, U, V) \in \Delta_{j,k}$ do
      $V' \leftarrow \overline{\mathcal{O}}(i, (+, U))$, $U' \leftarrow \overline{\mathcal{O}}(i, (-, V))$, $\widetilde{Q}_{\mathsf{P}} \overset{\cup}{\leftarrow} \{(i, U, V', j), (i, U', V, j)\}$
return $\widetilde{Q}_{\mathsf{P}}$
```

Fig. 4. Description of algorithm Sam.

[9] This does not matter here, as an entry can only be overwritten with the same value; below, we will change the experiment in a way that overwrites may be inconsistent, and we want to ensure we agree to keep the first value.

Then, after all of this, G_3 resumes by executing $\overline{R}(L, \boldsymbol{s})$. For $\overline{R}$'s j-th query (i, x) we do the following:

1. If $(i, x) \in \widetilde{X}^{j'}$ for some $j' \leq j$, then we respond with $\widetilde{T}[i, x]$.
2. Otherwise, if $(i, x) \in X$, but the first condition was not met, we respond with $T[i, x]$.
3. Finally, if neither of the above is true, we respond randomly.

As before, G_3 outputs $\mathtt{true}$ iff $\overline{R}$ outputs 0. For now, all modifications are syntactical. Indeed, up to the point we start $\overline{R}$, we satisfy the invariant that $T[i, x] = f_i(x)$ or $\widetilde{T}[i, x] = f_i(x)$ whenever these are defined, because $\overline{\mathcal{O}}$ behaves according to the NR construction using $\boldsymbol{s}$. On the other hand, if during the execution $\overline{R}(L, \boldsymbol{s})$ we respond randomly, we know for sure $f_i(x)$ was not queried by $\overline{S}$, and thus we can re-sample it. Thus, $\mathsf{Pr}[\mathsf{G}_3] = \mathsf{Pr}[\mathsf{G}_2] = \mathsf{Pr}[\mathsf{G}_1]$.

Moving to G_4, we now answer $\overline{\mathcal{O}}$ queries by S (within $\overline{S}$) and by Sam using random permutations $\pi_1, \ldots \pi_r$, instead of simulating the NR construction using $f_1, \ldots, f_r$, i.e., $\overline{\mathcal{O}}(i, (+, U)) = \pi_i(U)$ and $\overline{\mathcal{O}}(i, (-, V)) = \pi_i^{-1}(V)$. The seeds $\boldsymbol{s}$ are now independent of $\overline{\mathcal{O}}$. We do not change anything else. We note that the indistinguishability of G_3 and G_4 directly reduces to a suitable distinguisher for Proposition 1, as only Sam and S (within $\overline{S}$) make queries to $\overline{\mathcal{O}}$, but they do not get the keys $\boldsymbol{s}$, which are used only after all queries to $\overline{O}$ have been made to define X and $\widetilde{X}$. Therefore,

$$\mathsf{Pr}[\mathsf{G}_1] = \mathsf{Pr}[\mathsf{G}_3] \leq \mathsf{Pr}[\mathsf{G}_4] + (q + 2q\tau)^2 \left(2\alpha + \frac{1}{2^{2n}}\right), \qquad (7)$$

where we have used the fact that Sam makes $2q\tau = 2pq\eta$ queries.

The final game is G_5 is identical to G_4, *except* that in the process of answering $\overline{R}$'s queries, if case 2 happens, we also set answer randomly. However, should such situation occur, a bad flag is set in G_5, and since up to the point this flag is set, the behavior of G_4 and G_5 is identical,

$$\mathsf{Pr}[\mathsf{G}_4] - \mathsf{Pr}[\mathsf{G}_5] \leq \mathsf{Pr}[\mathsf{G}_5 \text{ sets } \mathsf{bad}].$$

To analyze the probability on the RHS, we need to specify $\mathcal{Q}(L, \widetilde{Q})$ used by Sam here. (Note all statements so far were independent of it.) For a given L which appears with positive probability in G_5, consider the distribution of the input-output pairs Q_{P} defined by the interaction of S with $\overline{\mathcal{O}}$, conditioned on the leakage being L, and $\pi_1, \ldots, \pi_r$ being consistent with the triples defined by $\widetilde{Q}$. Then, $\mathcal{Q}(L, \widetilde{Q})$ outputs a sample of Q_{P} according to this distribution. Using this, we prove the following lemma in our full version [38, App. A.4], which uses ideas similar to those from [26], with some modifications due to the setting (and the fact that $\overline{R}$ makes p queries).

Lemma 1. $\mathsf{Pr}\left[\mathsf{G}_5 \textit{ sets } \mathsf{bad}\right] \leq \varepsilon/2$

```
ADVERSARY $R^{\mathcal{O}}(1^r, L)$ :
$c \leftarrow 0,\ \widetilde{X} \leftarrow \emptyset$
$\boldsymbol{s} = (s_1^{\mathsf{in}}, s_1^{\mathsf{out}}, \ldots, s_r^{\mathsf{in}}, s_r^{\mathsf{out}}) \xleftarrow{\$} \mathsf{P.Kg}$
$\widetilde{Q}_{\mathsf{P}} \xleftarrow{\$} \mathsf{Sam}^{\mathcal{O}}(L)$
for $j = 1$ to $p$ do
 for all $(i, U, V, j) \in \widetilde{Q}$ do
   $x_0 \,\|\, x_1 \leftarrow \mathsf{P}(s_i^{\mathsf{in}}, U),\ x_3 \,\|\, x_2 \leftarrow \mathsf{P}(s_i^{\mathsf{out}}, V)$
   $\widetilde{X}^j \leftarrow \widetilde{X}^j \cup \{(i, 0 \,\|\, x_1), (i, 1 \,\|\, x_2)\}$
   if $\widetilde{T}[i, 0 \,\|\, x_1] = \perp$ then $\widetilde{T}[i, 0 \,\|\, x_1] \leftarrow x_0 \oplus x_2$
   if $\widetilde{T}[i, 1 \,\|\, x_2] = \perp$ then $\widetilde{T}[i, 1 \,\|\, x_2] \leftarrow x_1 \oplus x_3$
$b' \leftarrow \overline{R}^{\mathcal{O}'}(L, \boldsymbol{s})$
return $b'$

Proc. $\mathcal{O}'(i, x)$:
$c \leftarrow c + 1,\ \widetilde{X} \leftarrow \widetilde{X} \cup \widetilde{X}^c$
if $T[i, x] = \perp$ then
   if $(i, x) \in \widetilde{X}$ then
      $T[i, x] \leftarrow \widetilde{T}[i, x]$
   else $T[i, x] \xleftarrow{\$} \{0, 1\}^n$
return $T[i, x]$
```

Fig. 5. Adversary R in the proof of Theorem 2.

Given this, we are now ready to give our adversary R, which we build from $\overline{R}$ and Sam as described in Fig. 5. By a purely syntactical argument,

$$\mathsf{Pr}\left[\mathsf{G}_5\right] = \mathsf{Pr}\left[\mathsf{p\text{-}Reset}_S^R \,\middle|\, b = 0\right], \tag{8}$$

recalling that the case $b = 0$ is the one where both S and R access the same permutations $\pi_1, \ldots, \pi_r$. Therefore, we have established, combining (8), (7), Lemma 1,

$$\mathsf{Pr}\left[\mathsf{p\text{-}Reset}_S^R \,\middle|\, b = 0\right] \geq \mathsf{Pr}\left[\mathsf{f\text{-}Reset}_S^{\overline{R}} \,\middle|\, b = 0\right] - \frac{\varepsilon}{2} - (q + 2q\tau)^2 \left(2\alpha + \frac{1}{2^{2n}}\right). \tag{9}$$

In the full version [38, App. A.5] we also prove formally that in the case $b = 1$, R in the game $\mathsf{p\text{-}Reset}_S^R$ almost perfectly simulates an execution of $\mathsf{f\text{-}Reset}_S^{\overline{R}}$, or more formally,

$$\mathsf{Pr}\left[\neg\mathsf{p\text{-}Reset}_S^R \,\middle|\, b = 1\right] \leq \mathsf{Pr}\left[\neg\mathsf{f\text{-}Reset}_S^{\overline{R}} \,\middle|\, b = 1\right] + (q + 2q\tau)^2 \left(2\alpha + \frac{1}{2^{2n}}\right). \tag{10}$$

We can combine (10) and (9) to obtain, with $\varDelta = 2(q + 2q\tau)^2 \left(2\alpha + \frac{1}{2^{2n}}\right)$,

$$\begin{aligned}\mathsf{Adv}_{2n,S}^{\mathsf{p\text{-}reset}}(R) &= \mathsf{Pr}\left[\mathsf{p\text{-}Reset}_S^R \,\middle|\, b = 0\right] - \mathsf{Pr}\left[\neg\mathsf{p\text{-}Reset}_S^R \,\middle|\, b = 1\right] \\ &\geq \mathsf{Pr}\left[\mathsf{f\text{-}Reset}_S^{\overline{R}} \,\middle|\, b = 0\right] - \mathsf{Pr}\left[\neg\mathsf{f\text{-}Reset}_S^{\overline{R}}) \,\middle|\, b = 1\right] - \frac{\varepsilon}{2} - \varDelta \\ &\geq \varepsilon/2 - \varDelta\,.\end{aligned}$$

This concluses the proof. □

4 Correlation Intractability of Public-Seed Permutations

In this section, we study the correlation intractability (CI) of the NR construction against k-ary evasive relations. Firstly, in Sect. 4.1, we define a stronger notion of evasiveness – *strong evasiveness* – and show that evasiveness and strong evasiveness are asymptotically equivalent when $k = O(1)$. In Sect. 4.2 we study the relations that are hard for two-round Feistel. In Sect. 4.3 we show that for $k = O(1)$ the NR construction where P^{-1} is a family of almost $O(k^2)$-wise independent permutations is correlation intractable against k-ary evasive relations. In the special case of unary evasive relations ($k = 1$), we show that (see [38, App. B.6]) P instead can be instantiated from one-round Feistel with a 10-wise independent round function.

4.1 Strong Evasiveness

Evasiveness is defined through the hardness of winning a security game. For our results, we need instead a combinatorial understanding of evasive relations. To this end, we will rely on the following notion of evasiveness, which, as we show below, is generally implied by evasiveness if $k = O(1)$.

Definition 1 (Strongly evasive relations). *Let R be a k-ary relation over $X \times X$ and $\delta \in [0, 1]$. We say that R is δ-*strongly evasive *if the following are true for all $0 \leq j < k' \leq k$:*

- *For all distinct $\mathbf{u}_1, \ldots, \mathbf{u}_{k'} \in X$, all $\mathbf{v}_1, \ldots, \mathbf{v}_j \in X$, we have*

$$|\{(\mathbf{v}_{j+1}, \ldots, \mathbf{v}_{k'}) \, : \, \{(\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'})\} \in R\}| \leq \delta \cdot \prod_{i=j}^{k'-1} (|X| - i).$$

- *For all distinct $\mathbf{v}_1, \ldots, \mathbf{v}_{k'} \in X$, all $\mathbf{u}_1, \ldots, \mathbf{u}_j \in X$, we have*

$$|\{(\mathbf{u}_{j+1}, \ldots, \mathbf{u}_{k'}) \, : \, \{(\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'})\} \in R\}| \leq \delta \cdot \prod_{i=j}^{k'-1} (|X| - i).$$

It is not hard to see that if a relation is δ-strongly evasive, then it is also evasive, in the sense that it is $(q, q^k\delta)$-evasive. In particular, $q^k\delta$ is negligible whenever δ is negligible, q polynomial, and $k = O(1)$.

We remark that there are relations R which are evasive, yet not strongly evasive. Consider for example the relation which contains $\{(0^{2n}, 0^{2n}), (\mathbf{u}, \mathbf{v})\}$ for *all* $\mathbf{u}, \mathbf{v} \neq 0^{2n}$. This relation is obviously evasive to start with – satisfying it requires $\pi(0^{2n}) = 0^{2n}$, which will happen with probability 2^{-2n} only, yet for $\mathbf{u}_1 = \mathbf{v}_1 = 0^{2n}$, and $\mathbf{u}_2 \neq 0^{2n}$, all strings $\mathbf{v}_2$ make $\{(\mathbf{u}_1, \mathbf{v}_1), (\mathbf{u}_2, \mathbf{v}_2)\}$ valid. Still, somehow, the intuition is that the core of R is the relation $R^* = \{\{(0^{2n}, 0^{2n})\}\}$, which *is* strongly evasive, with $\delta = 2^{-2n}$. Indeed, an attacker that satisfies the original relation R, can directly satisfy R^*, thus the fact that R^* is evasive (and in particular, strongly evasive) implies that R is evasive.

The following lemma generalizes this, and implies e.g. that for $\delta = \mathsf{negl}(\lambda)$ and $k = O(1)$, evasiveness and strong evasiveness are (qualitatively) equivalent. The proof is found in the full version [38, App. B.1].

Lemma 2 (Normalization of evasive relations). *Let $\delta > 0$, and let R be a k-ary (k^2, δ^k)-evasive relation on $X \times X$ for random permutations. Then, there exists a relation R^* which is δ-strongly evasive for random permutations, and moreover, for every $S \in R$, there exists $\emptyset \neq S^* \subseteq S$ such that $S^* \in R^*$.*

Lemma 2 now is all we need. Say E is correlation intractable for all k-ary strongly evasive relations (for some negligible δ), where $k = O(1)$. Then, E must be correlation intractable for any (k^2, δ)-evasive relation R, too. Were it not, we could take an adversary A breaking the CI for R with non-negligible advantage, and use it to break CI of R^*. To this end, we simply run A, and when it outputs $S \in R$, we outputs the corresponding $S^* \in R^*$ guaranteed by Lemma 2. (As $k = O(1)$, a random subset of S will do with constant loss in the advantage.) But since R^* is $\sqrt[k]{\delta}$-strongly evasive, this contradicts our assumption on E.

Clearly, the equivalence is merely asymptotic. If one is interested in concrete security, the best approach to use our results below is to directly assess the δ for which a specific relation R is δ-strongly evasive.

4.2 Partial Correlation Intractability of Two-Round Feistel

The two-round Feistel construction Fei_2^f, is a permutation on $2n$-bit strings that makes calls to an oracle $f : \{0,1\}^{n+1} \to \{0,1\}^n$. In particular, on input $\mathbf{x} = x_0 \,\|\, x_1$, where $x_0, x_1 \in \{0,1\}^n$, running $\mathsf{Fei}_2^f(\mathbf{x})$ outputs $\mathbf{y} = x_2 \,\|\, x_3$, where

$$x_2 \leftarrow x_0 \oplus f(0 \,\|\, x_1)\ ; x_3 \leftarrow x_1 \oplus f(1 \,\|\, x_2).$$

Symmetrically, upon an inverse query, $\mathsf{Fei}_2^{f^{-1}}(\mathbf{y} = x_2 \,\|\, x_3)$ simply computes the values backwards, and outputs $\mathbf{x} = x_0 \,\|\, x_1$.

In this section, we discuss relations R on $2n$-bit strings that are hard for two-round Feistel when instantiated with a random function. In particular, we will give a combinatorial characterization which is sufficient to achieve this.

Feistel evasiveness. We first note that in a relation R, certain sets $S \in R$ can never be satisfied by the two-round Feistel construction out of structural constraints. In particular, if we have two input-output pairs $(\mathbf{x}_1[0] \,\|\, \mathbf{x}_1[1], \mathbf{x}_1[2] \,\|\, \mathbf{x}_1[3])$ and $(\mathbf{x}_2[0] \,\|\, \mathbf{x}_2[1], \mathbf{x}_2[2] \,\|\, \mathbf{x}_2[3])$ with $\mathbf{x}_1[2] = \mathbf{x}_2[2]$ in the same set $S \in R$, then we *must* have $\mathbf{x}_1[3] \oplus \mathbf{x}_2[3] = \mathbf{x}_1[1] \oplus \mathbf{x}_2[1]$. Symmetrically, if $\mathbf{x}_1[1] = \mathbf{x}_2[1]$, then we must have $\mathbf{x}_1[0] \oplus \mathbf{x}_2[0] = \mathbf{x}_1[2] \oplus \mathbf{x}_2[2]$. It will thus be convenient to define the following.

Definition 2. *For every k-ary relation R on $2n$-bit strings, we define the relation $\overline{R} \subseteq R$ that only contains $S \in R$ if for every $(\mathbf{x}_1[0] \,\|\, \mathbf{x}_1[1], \mathbf{x}_1[2] \,\|\, \mathbf{x}_1[3])$, $(\mathbf{x}_2[0] \,\|\, \mathbf{x}_2[1], \mathbf{x}_2[2] \,\|\, \mathbf{x}_2[3]) \in S$, the following is true:*

- *If* $\mathbf{x}_1[2] = \mathbf{x}_2[2]$*, then* $\mathbf{x}_1[3] \oplus \mathbf{x}_2[3] = \mathbf{x}_1[1] \oplus \mathbf{x}_2[1]$.
- *If* $\mathbf{x}_1[1] = \mathbf{x}_2[1]$*, then* $\mathbf{x}_1[0] \oplus \mathbf{x}_2[0] = \mathbf{x}_1[2] \oplus \mathbf{x}_2[2]$.

Clearly, the significance of this is that when assessing whether R is correlation intractable for two-round Feistel, it suffices to prove that $\overline{R}$ is correlation intractable, as $S \in R \setminus \overline{R}$ can never be satisfied. We are now ready to state the following combinatorial requirement on relations, which we will prove to be evasive for two-round Feistel below.

Definition 3 (δ-2-Feistel evasive relations). *Let R be a k-ary relation over $\{0,1\}^{2n}$, and $\delta \in [0,1]$. We say that R is* δ-2-Feistel evasive *if the following are true for all $0 \le j < k' \le k$:*

- *For all distinct $\mathbf{x}_1, \dots, \mathbf{x}_{k'} \in \{0,1\}^{2n}$, distinct $\mathbf{y}_1, \dots, \mathbf{y}_j \in \{0,1\}^{2n}$, and $y^* \in \{0,1\}^n$ s.t. $\mathbf{x}_{j+1}[1], \dots, \mathbf{x}_{k'}[1]$ are distinct and $y^* \notin \{\mathbf{y}_1[0], \dots, \mathbf{y}_j[0]\}$,*

$$\Big|\{(y_{j+1}, \dots, y_{k'})\} \;:\; \{(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_j, \mathbf{y}_j), (\mathbf{x}_{j+1}, y^* \,\|\, y_{j+1}), \dots, (\mathbf{x}_{k'}, y^* \,\|\, y_{k'})\} \in \overline{R'}\Big| \le \delta' \cdot 2^n \,. \tag{11}$$

- *For all distinct $\mathbf{y}_1, \dots, \mathbf{y}_{k'} \in \{0,1\}^{2n}$, distinct $\mathbf{x}_1, \dots, \mathbf{x}_j \in \{0,1\}^{2n}$, and $x^* \in \{0,1\}^n$ s.t. $\mathbf{y}_{j+1}[0], \dots, \mathbf{y}_{k'}[0]$ are distinct and $x^* \notin \{\mathbf{x}_1[1], \dots, \mathbf{x}_j[1]\}$,*

$$\Big|\{(x_{j+1}, \dots, x_{k'})\} \;:\; \{(\mathbf{x}_1, \mathbf{y}_1), \dots, (\mathbf{x}_j, \mathbf{y}_j), (x_{j+1} \,\|\, x^*, \mathbf{y}_{j+1}), \dots, (x_{k'} \,\|\, x^*, \mathbf{y}_{k'})\} \in \overline{R'}\Big| \le \delta' \cdot 2^n \,. \tag{12}$$

Also, we let $\mathsf{FEv}(k, \delta)$ denote the set of all k-ary δ-2-Feistel evasive relations.

Feistel correlation intractability. We now prove that for all relations satisfying Definition 3, two-round Feistel is indeed correlation intractable in the model where both round functions are independent random functions, to which the adversary is given oracle access.

Proposition 2 (CI of Two-round Feistel). *For $\delta \in [0,1]$ and $k \ge 1$ be an integer, let $R \in \mathsf{FEv}(k, \delta)$. For any (unbounded) adversary A making at most q queries to f, $\mathsf{Adv}^{\mathsf{cri}}_{R,\mathsf{Fei}_2^f}(A) \le 2k\delta \cdot q^{2k+1}$.*

The proof of Proposition 2 can be found in the full version [38, App. B.2].

Remark 1. For the special case of $k = 1$, that is, unary relations, one can adapt the above proof and show that $\mathsf{Adv}^{\mathsf{cri}}_{R,\mathsf{Fei}_2^f}(A) \le \delta \cdot q^2$ where A makes q queries to f and $R \in \mathsf{FEv}(1, \delta)$.

4.3 Correlation Intractability of the NR Construction

In this section we view the NR construction as a family $\mathsf{NR}^f[\mathsf{P}]$ that makes oracle calls to $f : \{0,1\}^{n+1} \to \{0,1\}^n$ and P is a family of permutations on $2n$-bits. The key generation algorithm $\mathsf{NR.Kg}$ now just outputs a tuple $(s^{\mathsf{in}}, s^{\mathsf{out}})$ where $s^{\mathsf{in}}, s^{\mathsf{out}} \xleftarrow{\$} \mathsf{P.Kg}$ and the evaluation algorithm $\mathsf{NR.Eval}$ proceeds as before but instead makes calls to f for evaluating the round function.

We show that $\mathsf{NR}^f[\mathsf{P}]$, where P^{-1} is a family of almost $O(k^2)$-wise independent permutations, is correlation intractable against strongly evasive k-ary relations when the adversary is given the seed $(s^{\mathsf{in}}, s^{\mathsf{out}})$ of the NR construction and only oracle access to f. The proof of CI proceeds by showing that P transforms a strongly evasive relation R into a 2-Feistel evasive relation $R_{\pi,\sigma}$ (see Fig. 6) and hence for the adversary to break the CI of $\mathsf{NR}^f[\mathsf{P}]$ it needs to break the CI of two-round Feistel against $R_{\pi,\sigma}$ which was studied in Sect. 4.2.

p-WISE INDEPENDENT PERMUTATIONS. For any $\varepsilon \in [0,1]$ and $p \geq 1$, we say that a family of permutations P on m-bit strings is (ε, p)*-wise independent* if for all distinct $u_1, \dots, u_p \in \{0,1\}^m$, the distributions of $\mathsf{P}(s, u_1), \dots, \mathsf{P}(s, u_p)$ (for $s \xleftarrow{\$} \mathsf{P.Kg}$) and of $\rho(u_1), \dots, \rho(u_p)$ (for $\rho \xleftarrow{\$} \mathsf{Perms}(m)$) are at most ε-apart in statistical distance.

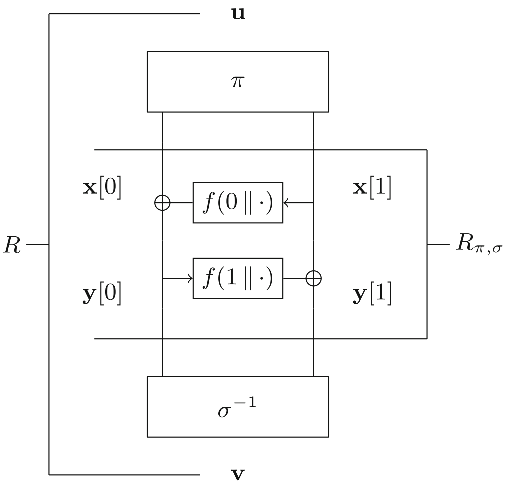

Fig. 6. The NR construction instantiated with a permutation family P on $2n$-bits such that P^{-1} is $(\varepsilon, k \cdot t)$-wise independent where $\pi = \mathsf{P}(s^{\mathsf{in}}, \cdot)$ and $\sigma = \theta(\mathsf{P}(s^{\mathsf{out}}, \cdot))$ for $s^{\mathsf{in}}, s^{\mathsf{out}} \leftarrow \mathsf{P.Kg}$. Here, θ is a permutation on $2n$-bits such that for all $\mathbf{x} \in \{0,1\}^{2n}$ we have $\theta(\mathbf{x} = x_0||x_1) = x_1||x_0$. For some k-ary strongly-evasive relation R, we construct a 2-Feistel evasive relation $R_{\pi,\sigma}$ by transforming every $(\mathbf{u}, \mathbf{v}) \in S$ where $S \in R$, by applying π to $\mathbf{u}$ and σ to $\mathbf{v}$.

From Strongly Evasive Relation to 2-Feistel-Evasive Relation. Let R be a k-ary relation over $\{0,1\}^{2n} \times \{0,1\}^{2n}$. For $s^{\mathsf{in}}, s^{\mathsf{out}} \xleftarrow{\$} \mathsf{P.Kg}$, let $\pi = \mathsf{P}(s^{\mathsf{in}}, \cdot)$ and $\sigma = \theta(\mathsf{P}(s^{\mathsf{out}}, \cdot))$, where θ is a permutation on $2n$-bits such that for $\mathbf{x} \in \{0,1\}^{2n}$ we have $\theta(\mathbf{x} = x_0||x_1) = x_1||x_0$[10]. We define a relation $R_{\pi,\sigma}$ which is a result of transforming $\{(\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'}))\} \in R$ via π and σ in the following way,

$$R_{\pi,\sigma} = \{\{(\pi(\mathbf{u}_1), \sigma(\mathbf{v}_1)), \ldots, (\pi(\mathbf{u}_{k'}), \sigma(\mathbf{v}_{k'}))\} \mid \{(\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'})\} \in R\}.$$

Then, for every δ-strongly evasive k-ary relation R we show that $R_{\pi,\sigma} \in \mathsf{FEv}(k, \delta')$ for some δ' larger than δ, except with small probability, where the probability is taken over the random choice of (π, σ). This is more formally captured in the following:

Proposition 3 (CI Amplification). *For $\delta \in [0,1)$ and an integer $k \geq 1$, let R be a k-ary δ-strongly evasive relation over $\{0,1\}^{2n}$. For an even integer $t \geq 2$, let P be a family of permutations such that P^{-1} is $(\varepsilon, k \cdot t)$-wise independent. Then, for $\delta' \in [0,1]$ such that $\delta' > \delta$,*

$$\Pr_{\pi,\sigma}[R_{\pi,\sigma} \notin \mathsf{FEv}(k, \delta')] \leq 12k^2 \left(\frac{1}{\delta' - \delta}\right)^t 2^{(4k-1)n} \left[C_t \cdot \left(\frac{4\delta^* t}{2^n}\right)^{t/2} + \varepsilon \cdot (1+\delta)^t\right],$$

where $C_t = 2e^{1/6t}\ \sqrt{\pi t}\ \left(\frac{5}{2e}\right)^{t/2}$, $\delta^* = \max\left(\delta, \frac{t \cdot 2^k}{2^n}\right)$, $\pi = \mathsf{P}(s^{\mathsf{in}}, \cdot)$ and $\sigma = \theta(\mathsf{P}(s^{\mathsf{out}}, \cdot))$ for $s^{\mathsf{in}}, s^{\mathsf{out}} \xleftarrow{\$} \mathsf{P.Kg}$.

Now, Proposition 3 can be combined with Proposition 2 to establish the correlation intractability of NR against strongly evasive relations (Theorem 3).

Theorem 3 (CI of NR). *For $\delta \in [0,1)$ and an integer $k \geq 1$, let R be a k-ary δ-strongly evasive relation over $\{0,1\}^{2n}$. Further, let P be a family of permutations on $2n$-bits such that P^{-1} is $(\varepsilon, 10k^2)$-wise independent where $\varepsilon \leq 1/2^{5kn}$. Then for any (potentially unbounded) adversary A making q queries,*

$$\mathsf{Adv}^{\mathsf{cri}}_{R,\mathsf{NR}^f[\mathsf{P}]}(A) \leq \frac{24k^2 \cdot (40k\delta^*)^{5k} + 12k^2 \cdot (1+\delta)^{10k}}{2^{4kn/9}} + 2k\delta' \cdot q^{2k+1}, \tag{13}$$

where $\delta' = \delta + 2^{-n/18}$ and $\delta^* = \max\left(\delta, \frac{10k \cdot 2^k}{2^n}\right)$.

The proof of Theorem 3 can be found in the full version [38, App. B.3]. The asymptotic interpretation of Eq. (13) is that when $n = \omega(\log \lambda)$, $k = O(1)$, $\delta = \mathsf{negl}(\lambda)$ and $q = \mathsf{poly}(\lambda)$, $\mathsf{NR}^f[\mathsf{P}]$ is correlation intractable for k-ary strongly evasive relations. Combining this with Lemma 2, the CI then extends to any k-ary evasive relation. We also remark that Theorem 3 extends to the setting of

[10] It is easy to see that $\theta = \theta^{-1}$ hence $\sigma^{-1} = \mathsf{P}^{-1}(s^{\mathsf{out}}, \theta(\cdot))$. We note that θ is introduced to ensure consistency with the definition of the NR construction as $\mathsf{P}^{-1}(s^{\mathsf{out}})$ operates on $\mathbf{y}[1]||\mathbf{y}[0]$ where $\mathbf{y}$ is the output of the underlying two-round feistel.

multi-key correlation intractability introduced in [16], but to avoid notational overhead we limit ourselves to the single-key setting for this version.

On instantiating P^{-1} from Theorem 3. We detail the construction of (ε, p)-wise permutations in the full version [38, App. B.5] and show that an $O(k)$-round Feistel construction with $10k^2$-wise independent round functions can instantiate the permutation family P^{-1}. We refer the reader to the full version for more details.

4.4 Proof of Proposition 3

We will show that for every $0 \leq j < k' \leq k$ the following hold,

1. For all distinct $\mathbf{x}_1, \ldots, \mathbf{x}_{k'}$, distinct $\mathbf{y}_1, \ldots, \mathbf{y}_j$ and $y^* \in \{0,1\}^n$ such that $\mathbf{x}_{j+1}[1], \ldots, \mathbf{x}_{k'}[1]$ are distinct and $y^* \notin \{\mathbf{y}_1[0], \ldots, \mathbf{y}_j[0]\}$,
$$\mathsf{Pr}[\text{Eq. (11) does not hold for } R_{\pi,\sigma}] \leq e_1(k', j) + \varepsilon \cdot e_2(k', j). \tag{14}$$
2. For all distinct $\mathbf{y}_1, \ldots, \mathbf{y}_{k'}$, all $\mathbf{x}_1, \ldots, \mathbf{x}_j$ and all $x^* \in \{0,1\}^n$ such that $\mathbf{y}_{j+1}[0], \ldots, \mathbf{y}_{k'}[0]$ are distinct and $x^* \notin \{\mathbf{x}_1[1], \ldots, \mathbf{x}_j[1]\}$,
$$\mathsf{Pr}[\text{Eq. (12) does not hold for } R_{\pi,\sigma}] \leq e_1(k', j) + \varepsilon \cdot e_2(k', j), \tag{15}$$

where the probability is taken over the random choice of (π, σ) and

$$e_1(k', j) = 3 \cdot C_t \left(\frac{1}{\delta' - \delta}\right)^t \left(\frac{2\delta^* t}{2^n}\right)^{t/2} 2^{(k'-j)(t/2+1)},$$
$$e_2(k', j) = 2 \left(\frac{1+\delta}{\delta' - \delta}\right)^t 2^{k'-j}.$$

Given that the above hold, we then take appropriate union bounds (for Eq. (14)) over all $\mathbf{y}_1, \ldots, \mathbf{y}_{k'}$, $\mathbf{x}_1, \ldots, \mathbf{x}_j$ and x^* and then over all j, k'. Symmetrically, we take union bounds (for Eq. (15)). Then the following holds and this concludes the proof of Theorem 3.

$$\begin{aligned} &\mathsf{Pr}[R_{\pi,\sigma} \notin \mathsf{FEv}(k, \delta')] \\ &\leq 2 \sum_{k'=1}^{k} \sum_{j=0}^{k'-1} 2^{n(2k'+2j+1)} \cdot (e_1(j, k') + \varepsilon \cdot e_2(j, k')) \\ &\leq 12k^2 \left(\frac{1}{\delta' - \delta}\right)^t 2^{(4k-1)n} \left[C_t \cdot \left(\frac{4\delta^* t}{2^n}\right)^{t/2} + \varepsilon \cdot (1+\delta)^t \right]. \end{aligned}$$

From now on we focus on showing Eq. (15) and the analysis for Eq. (14) is symmetrical.

Establishing Equation (15). Let us fix some arbitrary k' and j such that $0 \leq j < k' \leq k$. Let us also fix some distinct $\mathbf{y}_1, \ldots, \mathbf{y}_{k'} \in \{0,1\}^{2n}$, distinct $\mathbf{x}_1, \ldots, \mathbf{x}_j$ and $x^* \in \{0,1\}^n$ such that $\mathbf{y}_{j+1}[0], \ldots, \mathbf{y}_{k'}[0]$ are distinct and $x^* \notin \{\mathbf{x}_1[0], \ldots, \mathbf{x}_j[0]\}$. Then, we are interested in counting the number of tuples $((\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'}))$ in R that on applying π and σ transform to $((\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_j, \mathbf{y}_j), (\cdot \,\|\, x^*, \mathbf{y}_{j+1}), \ldots, (\cdot \,\|\, x^*, \mathbf{y}_{k'}))$. Let us fix σ and this defines $\mathbf{v}_i = \sigma^{-1}(\mathbf{y}_i)$ for every $i \in [k']$ allowing us to focus only on the following set $\mathcal{U}$ of tuples.

$$\mathcal{U} = \{(\mathbf{u}_1, \ldots, \mathbf{u}_{k'}) \mid \{(\mathbf{u}_1, \mathbf{v}_1), \ldots, (\mathbf{u}_{k'}, \mathbf{v}_{k'})\} \in R\}.$$

Then, we are interested in counting the number of tuples $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_{k'})$ in $\mathcal{U}$ that satisfy,

1. $\pi(\mathbf{u}_1) = \mathbf{x}_1,\ \pi(\mathbf{u}_2) = \mathbf{x}_2,\ \ldots,\ \pi(\mathbf{u}_j) = \mathbf{x}_j$.
2. $\pi_1(\mathbf{u}_{j+1}) = \pi_1(\mathbf{u}_{j+2}) \ldots = \pi_1(\mathbf{u}_{k'}) = x^*$.
3. For every $i \in \{j+1, \ldots, k'\}$, $\pi_0(\mathbf{u}_i) \oplus \pi_0(\mathbf{u}_{j+1}) = \Delta_i$, where $\Delta_i = \mathbf{y}_{j+1}[0] \oplus \mathbf{y}_i[0]$[11],

where $\pi_0(\mathbf{u})$ and $\pi_1(\mathbf{u})$ denote the first n-bits and last n-bits of $\pi(\mathbf{u})$. Or equivalently, count the number of $\mathbf{U}$'s such that $\pi(\mathbf{U})$[12] falls in $\mathcal{X}$ where,

$$\mathcal{X} = \{(\mathbf{x}_1, \ldots, \mathbf{x}_j, x \oplus \Delta_{j+1} \,\|\, x^*, \ldots, x \oplus \Delta_{k'} \,\|\, x^*) \mid x \in \{0,1\}^n\}.$$

Note that every element $\mathbf{X}$ of $\mathcal{X}$ is completely described by an n-bit string x. Now for $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_{k'}) \in \mathcal{U}$, let $I_{\mathbf{U}}$ be an indicator random variable,

$$I_{\mathbf{U}} = \begin{cases} 1 & \text{if } (\pi(\mathbf{u}_1), \ldots, \pi(\mathbf{u}_{k'})) \in \mathcal{X}, \\ 0 & \text{otherwise.} \end{cases}$$

Then it suffices to prove that,

$$\Pr_{\pi}\left[\sum_{\mathbf{U} \in \mathcal{U}} I_{\mathbf{U}} > \delta' \cdot 2^n\right] \leq e_1(k', j) + \varepsilon \cdot e_2(k', j).$$

Instead of looking at the sum $\sum I_{\mathbf{U}}$, we look at an equivalent sum $\sum I_x$ of, albeit, different indicator random variables I_x's, which will be convenient to analyse. For $x \in \{0,1\}^n$ we define an indicator random variable I_x which is 1 if π^{-1} transforms $\mathbf{X} \in \mathcal{X}$ (that corresponds to x) into some $\mathbf{U} \in \mathcal{U}$. More formally,

$$I_x = \begin{cases} 1 & \text{if } (\pi^{-1}(\mathbf{x}_1), \ldots, \pi^{-1}(x \oplus \Delta_{j+1} || x^*), \ldots, \pi^{-1}(x \oplus \Delta_{k'} || x^*)) \in \mathcal{U}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

[11] As the definition of δ'-2-Feistel evasiveness concerns itself with $\overline{R_{\pi,\sigma}}$.

[12] By $\pi(\mathbf{U})$ we mean the tuple $(\pi(\mathbf{u}_1), \ldots, \pi(\mathbf{u}_{k'}))$.

Then, it is easy to see that counting the number of $\mathbf{U} \in \mathcal{U}$ such that $I_{\mathbf{U}} = 1$ (or $\pi(\mathbf{U}) \in \mathcal{X}$) is the same as counting the number of $x \in \{0,1\}^n$ such that $I_x = 1$ (or $\pi^{-1}(\mathbf{X}) \in \mathcal{U}$). Therefore, $\sum\limits_{\mathbf{U}\in\mathcal{U}} I_{\mathbf{U}} = \sum\limits_{x\in\{0,1\}^n} I_x$ and we aim to show that,

$$\Pr_{\pi}\left[\sum_{x\in\{0,1\}^n} I_x > \delta' \cdot 2^n\right] \leq e_1(k', j) + \varepsilon \cdot e_2(k', j). \tag{17}$$

Partitioning $\{0,1\}^n$. We would like to use concentration bounds for the sum of random variables I_x's. But note that they are not independent as they may depend on the output of π^{-1} on the same input. Therefore, as a first step towards constructing independent random variables, we partition $\{0,1\}^n$ into subsets which will allow us to break the sum $\sum I_x$ into sums over these subsets.

Let us consider the following relation on $\{0,1\}^n \times \{0,1\}^n$. For any $x, x' \in \{0,1\}^n$, we say that x is related to x' (denoted as $x \sim x'$) if there exists an index set $\mathcal{B} \subseteq \{j+1, \ldots, k'\}$ where such that,

$$x = x' \oplus \bigoplus_{i\in\mathcal{B}} \Delta_i.$$

It is easy to see that the relation $\sim$ is an equivalence relation. Then, for any $x \in \{0,1\}^n$, let EQ_x denote its equivalence class, that is, $\mathsf{EQ}_x = \{x' \in \{0,1\}^n \,| x \sim x'\}$. Let $|\mathsf{EQ}_x| = l$ and it is easy to see that $l \leq 2^{k'-j}$. Let $\{\mathsf{EQ}_i\}_{i=1}^M$ be the M equivalence classes of $\sim$ where $|\mathsf{EQ}_i| = l$ and $M \cdot l = 2^n$. Furthermore, let $\mathsf{EQ}_i = \{x_1^i, x_2^i, \ldots, x_l^i\}$ be an enumeration of EQ_i where x_q^i is the qth member of the ith equivalence class EQ_i. Then, we can break the sum of I_x's into,

$$\sum_{x\in\{0,1\}^n} I_x = \sum_{i=1}^{M}\sum_{q=1}^{l} I_{x_q^i} = \sum_{q=1}^{l}\sum_{i=1}^{M} I_{x_q^i}.$$

For $q \in [l]$, let $X_q = \sum\limits_{i=1}^{M} I_{x_q^i}$. In other words, X_q is the sum of qth member of each equivalence class EQ_i. We are going to show that for every $q \in [l]$,

$$\Pr[X_q > \delta' \cdot M] \leq 3C_t \cdot \frac{1}{(\delta' - \delta)^t}\left(\frac{2t\delta^*}{2^n}\right)^{t/2} l^{t/2} + 2\varepsilon \cdot \left(\frac{1+\delta}{\delta' - \delta}\right)^t \tag{18}$$

Taking union bound over all $q \in [l]$ and using $l \leq 2^{k'-j}$, we have that Eq. (17) holds.

Bounding the subsum X_q. From now on, we will focus on analysing one of the subsums X_q and the other subsums can be analogously handled. Fix some q and let $X = X_q$. Let the corresponding set of x's be $\{x^1, \ldots, x^M\}$ where each x^i comes from a different equivalence class EQ_i. For every $i_1 \neq i_2 \in [M]$,

- Firstly, $\Delta_{j+1}, \ldots, \Delta_{k'}$ are distinct as $\mathbf{y}_{j+1}[0], \ldots, \mathbf{y}_{k'}[0]$ are distinct. Therefore, $x^{i_1} \oplus \Delta_{j+1}, \ldots, x^{i_1} \oplus \Delta_{k'}$ are distinct.
- Secondly for any index set $\mathcal{B} \subseteq \{j+1, \ldots, k'\}$,

$$x^{i_1} \neq x^{i_2} \oplus \bigoplus_{i \in \mathcal{B}} \Delta_i. \tag{19}$$

This implies that for any $I_{x^{i_1}}$ and $I_{x^{i_2}}$, $\{x^{i_1}_{j+1} \oplus \Delta_{j+1}, \ldots, x^{i_1}_{k'} \oplus \Delta_{k'}\}$ and $\{x^{i_2}_{j+1} \oplus \Delta_{j+1}, \ldots, x^{i_2}_{k'} \oplus \Delta_{k'}\}$ are disjoint. Therefore, except the first j (values that correspond to $\pi^{-1}(\mathbf{x}_i)$ for $i \in [j]$), the remaining set of values in the output of π^{-1} that each $I_{x^{i_1}}$ and $I_{x^{i_2}}$ depend on are disjoint.

We will crucially exploit these two properties of I_{x^i}'s to show that the following:

Lemma 3. *For X (as defined above), there exists a random variable Z with expectation $\mu = \mathbb{E}[Z] \leq \delta \cdot M$ where Z is a sum of M independent indicator random variables, such that for any integer $a > 0$,*

$$\mathsf{Pr}[|X - \mu| > a] \leq \frac{3 \cdot \mathbb{E}[(Z-\mu)^t]}{a^t} + 2\varepsilon \cdot \frac{(M+\mu)^t}{a^t}.$$

For each indicator random variable I_{x^i}, we first define another indicator random variable $I^{\rho}_{x^i}$ where the only difference is that we replace the $k \cdot t$- wise independent permutation π^{-1} with a random permutation ρ. Note that the resulting $I^{\rho}_{x^i}$ are still not independent as they depend on the output of ρ. So, we then define a sequence of random variable $I^*_{x^i}$ that have the same marginal distribution as that of $I^{\rho}_{x^i}$ but are independent. Then, we show a domination argument that relates the t-th moment of $(Y - \mu)$ with the t-th moment of $(Z - \mu)$ where Y and Z are the sum of $I^{\rho}_{x^i}$ and $I^*_{x^i}$ respectively. The proof of Lemma 3 can be found in the full version [38, App. B.4]. Next, we apply the following concentration bound due to [10] to the random variable Z.

Lemma 4 (A.4. from [10]). *Let $t \geq 2$ be an even integer. Suppose $Z_1, \ldots, Z_n$ are independent random variables taking values in $[0,1]$. Let $Z = Z_1 + \ldots + Z_n$ and $\mu = \mathbb{E}[Z]$. Then,*

$$\mathbb{E}\left[(Z-\mu)^t\right] \leq C_t \cdot (t\mu + t^2)^{t/2}.$$

Then as $\mu \leq \delta \cdot M$ we have,

$$\mathsf{Pr}[X > \delta \cdot M + a] \leq \mathsf{Pr}[X > \mu + a] \leq 3C_t \cdot \left(\frac{t\mu + t^2}{a^2}\right)^{t/2} + 2\varepsilon \cdot \left(\frac{M+\mu}{a}\right)^t,$$

Now let $a = (\delta' - \delta) \cdot M$ and using $M \cdot l = 2^n$ and $\delta^* = \max(\delta, t \cdot 2^k/2^n)$, we have

$$\mathsf{Pr}[X > \delta' \cdot M] \leq 3C_t \cdot \frac{1}{(\delta' - \delta)^t} \left(\frac{2t\delta^*}{2^n}\right)^{t/2} l^{t/2} + 2\varepsilon \cdot \left(\frac{1+\delta}{\delta' - \delta}\right)^t$$

which establishes that Eq. (18) holds (which establishes that Eq. (15) holds) and thereby concludes the proof of Proposition 3. □

Acknowledgments. The authors were supported by NSF grants CNS-1553758 (CAREER), CNS-1423566, CNS-1719146, CNS-1528178, and IIS-1528041, and by a Sloan Research Fellowship.

References

1. Andreeva, E., Bogdanov, A., Dodis, Y., Mennink, B., Steinberger, J.P.: On the indifferentiability of key-alternating ciphers. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013. LNCS, vol. 8042, pp. 531–550. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40041-4_29
2. Andreeva, E., Bogdanov, A., Mennink, B.: Towards understanding the known-key security of block ciphers. In: Moriai, S. (ed.) FSE 2013. LNCS, vol. 8424, pp. 348–366. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-43933-3_18
3. Aumasson, J.-P., Jovanovic, P., Neves, S.: NORX8 and NORX16: authenticated encryption for low-end systems. Cryptology ePrint Archive, Report 2015/1154 (2015). http://eprint.iacr.org/2015/1154
4. Barak, B., Mahmoody-Ghidary, M.: Merkle puzzles are optimal — an $O(n^2)$-query attack on any key exchange from a random oracle. In: Halevi, S. (ed.) CRYPTO 2009. LNCS, vol. 5677, pp. 374–390. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03356-8_22
5. Bellare, M., Hoang, V.T.: Resisting randomness subversion: fast deterministic and hedged public-key encryption in the standard model. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 627–656. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46803-6_21
6. Bellare, M., Hoang, V.T., Keelveedhi, S.: Instantiating random oracles via UCEs. In: Canetti, R., Garay, J.A. (eds.) CRYPTO 2013. LNCS, vol. 8043, pp. 398–415. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40084-1_23
7. Bellare, M., Hoang, V.T., Keelveedhi, S.: Cryptography from compression functions: the UCE bridge to the ROM. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8616, pp. 169–187. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44371-2_10
8. Bellare, M., Hoang, V.T., Keelveedhi, S., Rogaway, P.: Efficient garbling from a fixed-key blockcipher. In: 2013 IEEE Symposium on Security and Privacy, pp. 478–492. IEEE Computer Society Press, May 2013
9. Bellare, M., Rogaway, P.: The security of triple encryption and a framework forcode-based game-playing proofs. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 409–426. Springer, Heidelberg (2006). https://doi.org/10.1007/11761679_25
10. Bellare, M., Rompel, J.: Randomness-efficient oblivious sampling. In: 35th FOCS, pp. 276–287. IEEE Computer Society Press, November 1994
11. Bellare, M., Stepanovs, I.: Point-function obfuscation: a framework and generic constructions. In: Kushilevitz, E., Malkin, T. (eds.) TCC 2016. LNCS, vol. 9563, pp. 565–594. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49099-0_21
12. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: On the indifferentiability of the sponge construction. In: Smart, N. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 181–197. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78967-3_11

13. Bertoni, G., Daemen, J., Peeters, M., Van Assche, G.: Sponge-based pseudo-random number generators. In: Mangard, S., Standaert, F.-X. (eds.) CHES 2010. LNCS, vol. 6225, pp. 33–47. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15031-9_3
14. Brzuska, C., Farshim, P., Mittelbach, A.: Indistinguishability obfuscation and UCEs: the case of computationally unpredictable sources. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014. LNCS, vol. 8616, pp. 188–205. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44371-2_11
15. Canetti, R., Goldreich, O., Halevi, S.: The random oracle methodology, revisited (preliminary version). In: 30th ACM STOC, pp. 209–218. ACM Press, May 1998
16. Cogliati, B., Seurin, Y.: Strengthening the known-key security notion for block ciphers. In: Peyrin, T. (ed.) FSE 2016. LNCS, vol. 9783, pp. 494–513. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-52993-5_25
17. Coron, J.-S., Patarin, J., Seurin, Y.: The random oracle model and the ideal cipher model are equivalent. In: Wagner, D. (ed.) CRYPTO 2008. LNCS, vol. 5157, pp. 1–20. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85174-5_1
18. Dachman-Soled, D., Katz, J., Thiruvengadam, A.: 10-Round feistel is indifferentiable from an ideal cipher. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9666, pp. 649–678. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49896-5_23
19. Dai, Y., Steinberger, J.: Indifferentiability of 8-round feistel networks. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9814, pp. 95–120. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53018-4_4
20. Dodis, Y., Ganesh, C., Golovnev, A., Juels, A., Ristenpart, T.: A formal treatment of backdoored pseudorandom generators. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9056, pp. 101–126. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46800-5_5
21. Dodis, Y., Stam, M., Steinberger, J., Liu, T.: Indifferentiability of confusion-diffusion networks. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9666, pp. 679–704. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49896-5_24
22. Fiat, A., Shamir, A.: How to prove yourself: practical solutions to identification and signature problems. In: Odlyzko, A.M. (ed.) CRYPTO 1986. LNCS, vol. 263, pp. 186–194. Springer, Heidelberg (1987). https://doi.org/10.1007/3-540-47721-7_12
23. Gaži, P., Tessaro, S.: Provably robust sponge-based PRNGs and KDFs. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9665, pp. 87–116. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49890-3_4
24. Hoang, V.T., Rogaway, P.: On generalized feistel networks. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 613–630. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14623-7_33
25. Holenstein, T., Künzler, R., Tessaro, S.: The equivalence of the random oracle model and the ideal cipher model, revisited. In: Fortnow, L., Vadhan, S.P. (eds.) 43rd ACM STOC, pp. 89–98. ACM Press, June 2011
26. Impagliazzo, R., Rudich, S.: Limits on the provable consequences of one-way permutations. In: 21st ACM STOC, pp. 44–61. ACM Press, May 1989
27. Knudsen, L.R., Rijmen, V.: Known-key distinguishers for some block ciphers. In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 315–324. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-76900-2_19
28. Luby, M., Rackoff, C.: How to construct pseudorandom permutations from pseudorandom functions. SIAM J. Comput. **17**(2), 373–386 (1988)

29. Mandal, A., Patarin, J., Seurin, Y.: On the public indifferentiability and correlation intractability of the 6-round feistel construction. In: Cramer, R. (ed.) TCC 2012. LNCS, vol. 7194, pp. 285–302. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28914-9_16
30. Matsuda, T., Hanaoka, G.: Chosen ciphertext security via UCE. In: Krawczyk, H. (ed.) PKC 2014. LNCS, vol. 8383, pp. 56–76. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-54631-0_4
31. Maurer, U., Renner, R., Holenstein, C.: Indifferentiability, impossibility results on reductions, and applications to the random oracle methodology. In: Naor, M. (ed.) TCC 2004. LNCS, vol. 2951, pp. 21–39. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24638-1_2
32. Mittelbach, A.: Salvaging indifferentiability in a multi-stage setting. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 603–621. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-642-55220-5_33
33. Naor, M., Reingold, O.: On the construction of pseudo-random permutations: Luby-Rackoff revisited (extended abstract). In: 29th ACM STOC, pp. 189–199. ACM Press, May 1997
34. Ramzan, Z., Reyzin, L.: On the round security of symmetric-key cryptographic primitives. In: Bellare, M. (ed.) CRYPTO 2000. LNCS, vol. 1880, pp. 376–393. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44598-6_24
35. Rogaway, P., Steinberger, J.: Security/efficiency tradeoffs for permutation-based hashing. In: Smart, N. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 220–236. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78967-3_13
36. Schmidt, J.P., Siegel, A., Srinivasan, A.: Chernoff-hoeffding bounds for applications with limited independence. In: Ramachandran, V. (ed.), 4th SODA, pp. 331–340. ACM-SIAM, January 1993
37. Soni, P., Tessaro, S.: Public-seed pseudorandom permutations. In: Coron, J.-S., Nielsen, J.B. (eds.) EUROCRYPT 2017. LNCS, vol. 10211, pp. 412–441. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-56614-6_14
38. Soni, P., Tessaro, S.: Naor-reingold goes public: The complexity of known-key security. Cryptology ePrint Archive, Report 2018/137 (2018). https://eprint.iacr.org/2018/137