Noname manuscript No.
(will be inserted by the editor)

On Constructing Pseudorandom Involutions: Feistel variants using a single round function

Chun Guo · Meiqin Wang · Weijia Wang

Received: date / Accepted: date

Abstract An involution is a permutation that is the inverse of itself. Involutions have attracted plenty attentions in cryptographic community due to their advantage regarding hardware implementations. In this paper, we reconsider constructing *pseudorandom involutions*. We demonstrate two constructions.

(i) First, the 4-round Feistel network *using the same random function (Feistel-SF) in every round* is a pseudorandom involution. This shows the Feistel-SF construction still provides non-trivial cryptographic strength. To complement, we also show insecurity of 3-round Feistel-SF by exhibiting an attack.
(ii) Second, a "mirrored" variant of the Naor-Reingold construction with component reusing yields a pseudorandom involution.

Keywords involution · indistinguishability · Feistel · Naor-Reingold

This paper was originally rejected from ToSC 2023 Issue 3. Then, in 30/08/2023, it was submitted to Designs, Codes and Cryptography, but no notification has been received.

Chun Guo
School of Cyber Science and Technology, Shandong University, Qingdao, China
Key Laboratory of Cryptologic Technology and Information Security of Ministry of Education, Shandong University, Qingdao, Shandong, 266237, China,
Shandong Research Institute of Industrial Technology, Jinan, Shandong, 250102, China
E-mail: chun.guo@sdu.edu.cn

Meiqin Wang
School of Cyber Science and Technology, Shandong University, Qingdao, China
Key Laboratory of Cryptologic Technology and Information Security of Ministry of Education, Shandong University, Qingdao, Shandong, 266237, China,
Quan Cheng Laboratory250103Jinan, Shandong, China
E-mail: mqwang@sdu.edu.cn

Weijia Wang
School of Cyber Science and Technology, Shandong University, Qingdao, China
Quan Cheng Laboratory250103Jinan, Shandong, China
Key Laboratory of Cryptologic Technology and Information Security of Ministry of Education, Shandong University, Qingdao, Shandong, 266237, China,
E-mail: wjwang@sdu.edu.cn

Mathematics Subject Classification (2000) 94A60 · 68P25

1 Introduction

This paper addresses two closely related questions:

(i) **Question 1:** how to construct *pseudorandom involutions* (PRIs).
(ii) **Question 2:** cryptographic strength of Feistel networks built upon a single round function (Feistel-SF).

Below we elaborate on the two issues in detail.

Constructing PRIs. An involution is a permutation that is the inverse of itself. Cryptographic involutions are efficient in terms of hardware areas, and have been adopted in a plenty of blockcipher designs [15,43,12,4,22].

Following the definitions of pseudorandom functions (PRFs) and pseudorandom permutations (PRPs), a *pseudorandom involution* (PRI) is a keyed involution $I : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ that is indistinguishable from a random involution $\mathsf{PRI} : \{0,1\}^n \to \{0,1\}^n$ when instantiated with a secret random key $K \xleftarrow{\$} \mathcal{K}$.

Note that while a PRI provides a somewhat random output $y = I(K,x)$ for the input x, it cannot be straightforwardly used as a blockcipher or enciphering scheme, since the enciphering oracle can be leveraged to decipher any sensitive ciphertext. In fact, a secure blockcipher or enciphering scheme is typically expected to be a *(strong) pseudorandom permutation ((S)PRP)* rather than a PRIs. On the other hand, this does not preclude PRIs from useful tools: the "somewhat" randomness may already suffice in (PRI-based) constructions, while the involutory property saves the cost of implementing I^{-1}. Due to these, several papers have proposed to consider using PRIs to replaced SPRPs in cryptographic constructions, including Feistel variants [24] and Misty network [35,21].[1] By these, PRIs actually deserve their own place alongside PRFs and (S)PRPs, and transformations among the three concepts are important in theory.

Aside from this, constructing PRIs is also of practical value. For example, Nandi [26] proved that enhancing a PRI with a pre-whitening key yields an SPRP, whose inverse has virtually no cost. As will be elaborated in detail later in Sect. 1.2, this could be quite appealing in side-channel and fault protected settings [8] as well as low-latency scenarios [9].

Regarding building PRIs, Naor and Reingold [31] has showed that the composition $P^{-1} \circ \sigma \circ P$ is a PRI when P is a PRP and σ is an (efficient) involution. Namely, P provides cryptographic strength while σ ensures the desired functionality. In theory this provides a PRP-to-PRI transformation, as shown in Fig. 1. However, this has two shortcomings. First, it has to invokes both P and its inverse P^{-1}, and this contradicts using PRIs to reduce

[1] Public truly random involutions have also been considered [20].

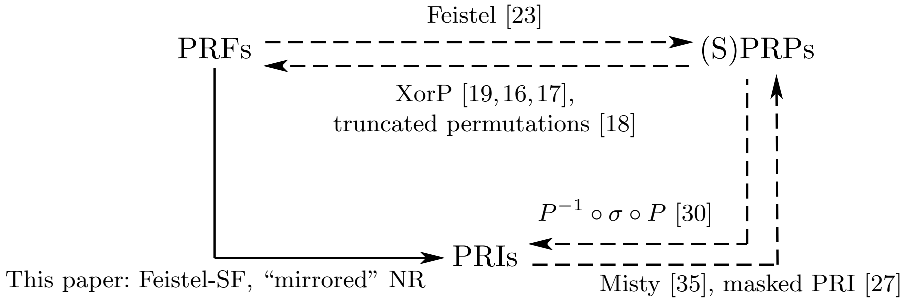
Feistel [23]
PRFs
(S)PRPs
XorP [19,16,17],
truncated permutations [18]
$P^{-1} \circ \sigma \circ P$ [30]
PRIs
This paper: Feistel-SF, "mirrored" NR
Misty [35], masked PRI [27]

Fig. 1: Relations between PRFs, (S)PRPs and PRIs.

implementation costs. Second, its cost roughly doubles the cost of P, whereas we certainly expect PRI to be as efficient as PRPs. Nandi provided another construction named Hash-Counter Involution, which still invokes both E and its inverse E^{-1} for the underlying blockcipher E. It is thus natural to ask if there are inverse-free constructions for PRIs (which could also bridge the gap of PRF-to-PRI transformations).

Feistel-SF. A Feistel permutation $\Psi^F_{K_i}(A\|B) := B\|(A \oplus F_{K_i}(B))$ applies a keyed function $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ with a subkey K_i to a half of the input, and then swap the two halves. It captures the underlying structure of a large proposition of blockciphers [1] including the DES [41]. A t-round Feistel $\Psi^F_t[K_1, K_2, ..., K_t](L\|R) := \Psi^F_{K_t} \circ ... \circ \Psi^F_{K_2} \circ \Psi^F_{K_1}(L\|R)$ is the t-composition of Feistel permutations. Let $\mathsf{swap}(A\|B) := B\|A$ be the swap, then we further denote $\overline{\Psi}^F_t[K_1, K_2, ..., K_t] := \mathsf{swap} \circ \Psi^F_t[K_1, K_2, ..., K_t]$ the variant of Ψ^F_t *without swap in the final round*, as shown in Fig. 2 (left).

Under the condition that F is a PRF, the seminal result of Luby and Rackoff [23] stated that $\Psi^F_3[K_1, K_2, K_3]$ instantiates a PRP and $\Psi^F_4[K_1, K_2, K_3, K_4]$ instantiates an SPRP. It is then natural to simplify this construction w.r.t. the number of keys. In this respect, the minimal Feistel variant $\overline{\Psi}_t[F, ..., F]$, i.e., *Feistel network using the Same random Function (Feistel-SF)*, was originally proposed by Schnorr [40] as a *PRP candidate*. But this expectation was soon broken by Rueppel [36] and further extended by Nandi [28]. Though, this attack did not distinguish Feistel-SF from *random involutions*—actually, they used the fact that Feistel-SF is involutory to distinguish it from *random permutations*. The negative result was later generalized to break PRP security of $\Psi_3[F^{i_1}, F^{i_2}, F^{i_3}]$ using compositions of F as round function [44] and SPRP security of some Feistel variants with round function reusing [38]. Meanwhile, a number of positive (S)PRP security result were later proven for various (strengthened) variants of Feistel-SF [34,32,37,29].

While Feistel-SF does not yield (S)PRPs, it still enjoys some cryptographic strength, and being involutory seems to be the "mere" weakness. In particular,

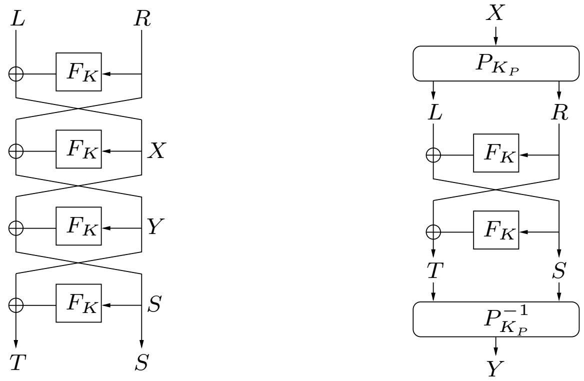

Fig. 2: (Left) The 4-round Feistel $\overline{\Psi}_4^F[K, K, K, K]$ built upon a single random function and without the final swap. (Right) The "mirror" Naor-Reingold construction $\mathsf{NR}^F[K_P, K, K, K_P]$.

it is a natural candidate for the aforementioned PRF-to-PRI transformation or inverse-free PRI constructions.

1.1 Our Results

By the above discussion, to seek for PRF-to-PRI transformations or inverse-free PRI constructions and to understand the security of weaker Feistel variants, we analyze the PRI security of the Feistel-SF construction, as well as another weaker variant of the popular Naor-Reingold construction.

1.1.1 Our first construction: Feistel-SF without final swap

We first show that the 4-round $\overline{\Psi}_4^F[K, K, K, K]$, as shown in Fig. 2 (left), is a PRI as long as $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ is a PRF. Security is ensured up to the birthday $2^{n/2}$ adversarial queries. To complement, we also exhibit an attack against 3-round $\overline{\Psi}_3^F[K, K, K]$. This attack simply follows the well-known CCA idea on 3-round $\overline{\Psi}_3^F[K_1, K_2, K_3]$, and our novelty is to provide a rigorous analysis of the attack advantage in the setting of PRIs.

1.1.2 Our second construction: "mirrored" Naor-Reingold

By the above positive result, using the Feistel-SF scheme, it is sufficient and necessary to use 4 PRF-calls per input. To further reduce the number of calls, it is natural to consider the Naor-Reingold construction [30], which is the most efficient known approach to build a secure SPRP. In detail, the Naor-Reingold construction

$$\mathsf{NR}^F[K_{P,1}, K_1, K_2, K_{P,2}](X) := P\big(K_{P,2}, \overline{\Psi}_2^F[K_1, K_2]\big(P(K_{P,1}, X)\big)\big) \qquad (1)$$

is built by sandwiching two *universal permutations* $P_{K_{P,1}}, P_{K_{P,2}} : \{0,1\}^{2n} \to \{0,1\}^{2n}$ (we refer to Sect. 4 for its definition) with a 2-round Feistel network $\overline{\Psi}_2^F[K_1, K_2]$. This may be viewed as a generalization of the Feistel network.

Naor and Reingold [30] proved CCA security for $\mathsf{NR}^F[K_{P,1}, K_1, K_2, K_{P,2}]$ using independent keys $K_{P,1}, K_{P,2} \xleftarrow{\$} \mathcal{K}_P$ and $K_1, K_2 \xleftarrow{\$} \mathcal{K}$, and Soni and Tessaro [42] strengthens the result by partially revealing secrets $K_{P,1}, K_{P,2}, K_1, K_2$ to the adversary. To have an involution, we consider a "mirrored" variant of the Naor-Reingold defined as

$$\mathsf{NR}^F[K_P, K, K, K_P](X) := P^{-1}\big(K_P, \overline{\Psi}_2^F[K, K]\big(P(K_P, X)\big)\big). \qquad (2)$$

I.e., it uses the same key in the two Feistel permutations and the same universal permutation P (and its inverse) at the beginning and end. See Fig. 2 (right) for illustration. We show that $\mathsf{NR}^F[K_P, K, K, K_P]$ instantiates a PRI up to $2^{n/2}$ adversarial queries (when P is good enough), yielding a PRI construction with complexity comparable with the best known SPRP constructions.

1.2 Discussion

Our constructions $\overline{\Psi}_4^F[K, K, K, K]$ and $\mathsf{NR}^F[K_P, K, K, K_P]$ provide inverse-free constructions for PRIs (which are the first, to our knowledge). For this, note that the right-universal permutation P can be instantiated using a Feistel round and a universal hash function $H : \mathcal{K}_P \times \{0,1\}^n \to \{0,1\}^n$. Namely, setting $P(K_P, X) = \mathsf{left}_n(X) \| \big(\mathsf{right}_n(X) \oplus H(K_P, \mathsf{left}_n(X))\big)$ suffices. In this case, P itself is an involution, and $\mathsf{NR}^F[K_P, K, K, K_P]$ remains inverse-free despite invoking P^{-1}.

In addition, the round complexity of $\overline{\Psi}_4^F[K, K, K, K]$ matches the rounds needed for a normal Feistel to be SPRP secure. In this sense, it means *the complexity of PRIs may be comparable with SPRPs.*

PRIs could find several applications. For example, to improve side-channel security of blockcipher-based authenticated encryption modes, Berti et al. [8] proposed a variant of the Hash-then-SPRP MAC scheme $t = E_K(H(m))$ that invokes the blockcipher inverse E^{-1} in verification to avoid leaking critical information. Unfortunately, this inverse E^{-1} is typically an inefficient side-channel (and fault) protected module in such designs [8,6,5,7,39]. Meanwhile, this use of inverse E^{-1} seems unavoidable under certain assumptions [5], and it thus seems crucial to reduce the cost of implementing the protected E^{-1}.

PRIs could be helpful in this setting. Though, it is insufficient to simply replace the SPRP with a PRI: the plain Hash-then-PRI $t = I_K(H(m))$ is insecure. The adversary picks distinct messages m, m' and computes $h \leftarrow$

$H(m)$. It then queries $\mathsf{LVrfy}(m', h)$ to obtain $I_K^{-1}(h) = u$ from the leakage. Since I is an involution, this also means $I_K(h) = u$, and the adversary outputs (m, u) as a valid forgery.

In this respect, Nandi [27] has proved that masking a PRI with another key, i.e., $E_{K_1,K_2}(x) = I_{K_2}(K_1 \oplus x)$, yields an SPRP. Thus, the leakage-resilient MAC can be instantiated with $t = E_{K_1,K_2}(H(m))$. In this case, the construction $E_{K_1,K_2}(x) = I_{K_2}(K_1 \oplus x)$ needs additional side-channel protections to avoid leaking K_1 and $K_1 \oplus x$. Though, since the XOR is linear, the added computations are minor.

Another potential application is to construct low-latency SPRPs (i.e., blockciphers or enciphering schemes). Concretely, low-latency scenarios such as disk and memory encryption may enforce using (very heavy) fully unrolled hardware implementations of the SPRP [9,2,10,3]. It is highly desirable to "recycle" the enciphering circuit for deciphering. To this end, new structures admitting special properties such as the α-reflection [9] have been introduced. The masked PRI construction of Nandi [27] provides another potential approach.

1.3 Organization

Sect. 2 provides notations and definitions. Then, Sect. 3 discusses insecurity of 3-round $\overline{\Psi}_3^F[K, K, K]$ and PRI security of 4-round $\overline{\Psi}_4^F[K, K, K, K]$, while Sect. 4 proves PRI security of the "mirrored" Naor-Reingold construction $\mathsf{NR}^F[P, K, K, P^{-1}]$. We finally conclude in Sect. 5.

2 Preliminaries

Notation. In all the following, we fix an integer $n \geq 1$, and denote by $\mathcal{F}(n, n)$ the set of all functions from $\{0,1\}^n$ to $\{0,1\}^n$. For any positive integer m, we denote by $\mathcal{P}(m)$ the set of all permutations on $\{0,1\}^m$. For integers $1 \leq \ell \leq m$, we write $(m)_\ell = m(m-\ell)\cdots(m-\ell+1)$ and $(m)_0 = 1$ by convention. When two sets A and B are disjoint, we denote $A \sqcup B$ their (disjoint) union. Given an m-bit string x and $a \leq n$, denote by $\mathsf{left}_a(x)$ (resp., $\mathsf{right}_a(x)$) the a leftmost (resp., rightmost) a bits of x.

Involution. The identity map on $\{0,1\}^m$ will be denoted ι. An involution $\sigma \in \mathcal{P}(m)$ is a permutation such that $\sigma \circ \sigma = \iota$. Let $\mathcal{I}(m)$ denote the set of all involutions in $\mathcal{P}(m)$ and let $\mathcal{I}_0(m)$ denote the set of involutions without any fixed point. Also, let $T_0(2^m) = |\mathcal{I}_0(m)|$ and $T(2^m) = |\mathcal{I}(m)|$. If an involution on $\{0,1\}^m$ has no fixed point, then we would have a perfect matching on $\{0,1\}^m$. The number of all possible perfect matchings on $\{0,1\}^m$ is

$$T_0(2^m) = \frac{1}{2^{\frac{2^m}{2}}} \binom{2^m}{\frac{2^m}{2}} \left(\frac{2^m}{2}\right)!, \tag{3}$$

and hence

$$\frac{T_0(2^m-2)}{T_0(2^m)} = \frac{1}{2^m-1}. \tag{4}$$

For $T(2^m)$, we will use the following recursion formula [13].

$$\frac{1}{\sqrt{2^m}+1} \leq \frac{T(2^m-1)}{T(2^m)} \leq \frac{1}{\sqrt{2^m}}. \tag{5}$$

We need to study the interaction between a distinguisher D^I and an involution oracle $I \in \mathcal{I}(m)$ (that instantiates either a PRI construction or a random involution). Following [20], the transcript of the already issued queries and responses should be $\mathcal{Q}_I = \big(\{X_1, Y_1\}, \{X_2, Y_2\}, ...\big)$, where the X_i's are pairwise distinct m-bit strings and the Y_i's are pairwise distinct m-bit strings, and where $\{X, Y\} \in \mathcal{Q}_I$ implies $\mathbf{I}(X) = Y$ or equivalently $\mathbf{I}(Y) = X$. Given such a transcript $\mathcal{Q}_I$ and an involution $I \in \mathcal{I}(m)$, we say that I *extends* $\mathcal{Q}_I$, denoted $I \vdash \mathcal{Q}_I$, if $I(X) = Y$ for all $\{X, Y\} \in \mathcal{Q}_I$.

With the above, we can prove a lemma about the distribution of the "next" response of a random involution $\mathbf{I}$. The proof is somewhat straightforward, but it may be of value to make this lemma explicit.

In detail, consider the interaction between $D^{\mathbf{I}}$ and a random involution $\mathbf{I}$, and let $\mathcal{Q}_I = \big(\{X_1, Y_1\}, \{X_2, Y_2\}, ..., \{X_{\ell-1}, Y_{\ell-1}\}\big)$ be the transcript of the already issued queries and their responses. Then, the probability to obtain Y_ℓ for the next query $\mathbf{I}(X_\ell)$, $X_\ell \notin \{X_1, ..., X_{\ell-1}, Y_1, ..., Y_{\ell-1}\}$, equals $\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big]$. We have bounds as follows.

Lemma 1 *For any transcript $\mathcal{Q}_I = \big(\{X_1, Y_1\}, \{X_2, Y_2\}, ..., \{X_{\ell-1}, Y_{\ell-1}\}\big)$ of an involution and any $X_\ell \in \{0,1\}^m \backslash \{X_1, ..., X_{\ell-1}, Y_1, ..., Y_{\ell-1}\}$, we have conclusions as follows.*

(i) If $\mathbf{I} \xleftarrow{\$} \mathcal{I}(m)$ is sampled from all m-bit involutions, then

$$\frac{1}{\sqrt{2^m-(\ell-1)}+1} \leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(m) : \mathbf{I}(X_\ell) = X_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] \leq \frac{1}{\sqrt{2^m-(2\ell-2)}}, \textit{ and}$$

$$\frac{1}{\sqrt{2^m-(\ell-1)}+1} \times \frac{1}{\sqrt{2^m-\ell}+1} \leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] \leq \frac{1}{\sqrt{2^m-(2\ell-2)}} \times \frac{1}{\sqrt{2^m-(2\ell-1)}} \tag{6}$$

for any $Y_\ell \notin \{X_1, ..., X_{\ell-1}, X_\ell, Y_1, ..., Y_{\ell-1}\}$.

(ii) If $\mathbf{I} \xleftarrow{\$} \mathcal{I}_0(m)$ is sampled from all m-bit involutions without fixed points, then

$$\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}_0(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{1}{2^m-(2\ell-1)} \tag{7}$$

for any $Y_\ell \notin \{X_1, ..., X_{\ell-1}, X_\ell, Y_1, ..., Y_{\ell-1}\}$.

Proof We address the two cases in turn.

The case of $\mathbf{I} \xleftarrow{\$} \mathcal{I}(m)$. Let N_1 be the number of involutions $I \in \mathcal{I}(m)$ with $I \vdash \mathcal{Q}_I$ and N_2 be the number of $I \in \mathcal{I}(m)$ with $I \vdash \mathcal{Q}_I \cup \big\{\{X_\ell, Y_\ell\}\big\}$. It then holds

$$\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{\Pr_{\mathbf{I}}\big[\mathbf{I} \vdash \mathcal{Q}_I \cup \big\{\{X_\ell, Y_\ell\}\big\}\big]}{\Pr_{\mathbf{I}}\big[\mathbf{I} \vdash \mathcal{Q}_I\big]} = \frac{\frac{N_2}{|\mathcal{I}(m)|}}{\frac{N_1}{|\mathcal{I}(m)|}}$$
$$= \frac{N_2}{N_1}.$$

To calculate N_1, we assume that the number of "fixed points" in $\mathcal{Q}_I$, i.e., $\{X, Y\} \in \mathcal{Q}_I$ with $X = Y$, is ω. This means $\{X_1, Y_1, X_2, Y_2, ..., X_{\ell-1}, Y_{\ell-1}\}$ has $2(\ell-1) - \omega$ distinct elements. It then holds

$$N_1 = T(2^m - (2\ell - 2 - \omega)).$$

On the other hand, N_2 depends on the value of Y_ℓ. When $Y_\ell = X_\ell$, we have

$$N_2 = T(2^m - (2\ell - 1 - \omega)),$$
$$\Pr_{\mathbf{I}}\big[\mathbf{I}(X_\ell) = X_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{T(2^m - (2\ell - 1 - \omega))}{T(2^m - (2\ell - 2 - \omega))}.$$

Using Eq. (5) and $0 \le \omega \le \ell - 1$, we reach

$$\frac{1}{\sqrt{2^m - (\ell - 1)} + 1} \le \Pr_{\mathbf{I}}\big[\mathbf{I}(X_\ell) = X_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] \le \frac{1}{\sqrt{2^m - (2\ell - 2)}}.$$

On the other hand, for any $Y_\ell \notin \{X_1, ..., X_{\ell-1}, X_\ell, Y_1, ..., Y_{\ell-1}\}$ we have $N_2 = T(2^m - (2\ell - \omega))$, and further

$$\Pr_{\mathbf{I}}\big[\mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{T(2^m - (2\ell - \omega))}{T(2^m - (2\ell - 2 - \omega))}$$
$$= \frac{T(2^m - (2\ell - 1 - \omega))}{T(2^m - (2\ell - 2 - \omega))} \times \frac{T(2^m - (2\ell - \omega))}{T(2^m - (2\ell - 1 - \omega))}.$$

Using Eq. (5) and $0 \le \omega \le \ell - 1$, we reach

$$\Pr_{\mathbf{I}}\big[\mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] \le \frac{1}{\sqrt{2^m - (2\ell - 2)}} \times \frac{1}{\sqrt{2^m - (2\ell - 1)}},$$
$$\Pr_{\mathbf{I}}\big[\mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] \ge \frac{1}{\sqrt{2^m - (\ell - 1)} + 1} \times \frac{1}{\sqrt{2^m - \ell} + 1}.$$

The case of $\mathbf{I} \xleftarrow{\$} \mathcal{I}_0(m)$. Let N_3 be the number of involutions $I \in \mathcal{I}_0(m)$ with $I \vdash \mathcal{Q}_I$ and N_4 be the number of $I \in \mathcal{I}_0(m)$ with $I \vdash \mathcal{Q}_I \cup \big\{\{X_\ell, Y_\ell\}\big\}$. It then holds

$$\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}_0(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{\frac{N_4}{|\mathcal{I}_0(m)|}}{\frac{N_3}{|\mathcal{I}_0(m)|}} = \frac{N_4}{N_3}.$$

In this case, $\mathcal{Q}_I$ cannot contain fixed points. Therefore, $N_3 = T(2^m - (2\ell - 2))$, and $N_4 = T(2^m - 2\ell)$ for any $Y_\ell \notin \{X_1, ..., X_{\ell-1}, X_\ell, Y_1, ..., Y_{\ell-1}\}$. Using Eq. (5), we reach

$$\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}_0(m) : \mathbf{I}(X_\ell) = Y_\ell \mid \mathbf{I} \vdash \mathcal{Q}_I\big] = \frac{T(2^m - 2\ell)}{T(2^m - (2\ell - 2))} = \frac{1}{2^m - (2\ell - 1)}.$$

These complete the proofs. □

Pseudorandom functions (PRFs) and involutions (PRIs). Consider a keyed function $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$, and denote $F_K(X)$ for $F(K, X)$. A (q,t)-adversary against F is an algorithm D with oracle access to a function from $\{0,1\}^n \to \{0,1\}^n$, making at most q oracle queries, running in time at most t, and outputting a single bit. The advantage of D in breaking the PRF-security of F is defined as

$$\mathbf{Adv}_F^{\mathsf{PRF}}(D) = \big|\Pr\big[K \leftarrow_\$ \mathcal{K} : D^{F_K} = 1\big] - \Pr\big[\mathbf{F} \leftarrow_\$ \mathcal{F}(n,n) : D^{\mathbf{F}} = 1\big]\big|$$

For simplicity, define

$$\mathbf{Adv}_F^{\mathsf{PRF}}(q,t) := \max_D \mathbf{Adv}_F^{\mathsf{PRF}}(D),$$

where the maximum is taken over all (q,t)-adversary D.

Similarly, consider a keyed involution $I : \mathcal{K} \times \{0,1\}^m \to \{0,1\}^m$, and consider a (q,t)-adversary D with oracle access to a involution $I : \{0,1\}^m \to \{0,1\}^m$. Since $I_K^{-1} \equiv I_K$, access to the "forward" oracle I is equivalent with access to both I and I^{-1}. This also means it does not make sense to build "CPA secure" pseudorandom involutions. The advantage of D in breaking the PRI-security of I is defined as

$$\mathbf{Adv}_I^{\mathsf{PRI}}(D) = \big|\Pr\big[K \leftarrow_\$ \mathcal{K} : D^{I_K} = 1\big] - \Pr\big[\mathbf{I} \leftarrow_\$ \mathcal{I}(m) : D^{\mathbf{I}} = 1\big]\big|$$

We also define

$$\mathbf{Adv}_I^{\mathsf{PRI}}(q,t) := \max_D \mathbf{Adv}_I^{\mathsf{PRI}}(D)$$

for simplicity, where the maximum is taken over all (q,t)-adversary D.

The H-coefficient technique. We use Patarin's H-coefficient technique [33] to prove our results. We provide a quick overview of its main ingredients here. Our presentation borrows heavily from that of [11]. Fix a distinguisher D that makes at most q queries to its oracles. As in the security definition presented above, D's aim is to distinguish between two worlds: a "real world" and an "ideal world". Assume w.l.o.g. that D is deterministic. The execution of D defines a *transcript* that includes the sequence of queries and answers received from its oracles; D's output is a deterministic function of its transcript. Thus, if $T_{\mathrm{re}}, T_{\mathrm{id}}$ denote the probability distributions on transcripts induced by the real and ideal worlds, respectively, then D's distinguishing advantage is upper bounded by the statistical distance

$$\Delta(T_{\mathrm{re}}, T_{\mathrm{id}}) := \frac{1}{2}\sum_{\mathcal{Q}} \Big| \Pr[T_{\mathrm{re}} = \mathcal{Q}] - \Pr[T_{\mathrm{id}} = \mathcal{Q}] \Big|, \tag{8}$$

where the sum is taken over all possible transcripts $\mathcal{Q}$.

Let $\mathcal{T}$ denote the set of all *attainable transcripts*, i.e., $\Pr[T_{\mathrm{id}} = \mathcal{Q}] > 0$ for all $\mathcal{Q} \in \mathcal{T}$. We look for a partition of $\mathcal{T}$ into two sets $\mathcal{T}_{good}$ and $\mathcal{T}_{bad}$ of "good" and "bad" transcripts, respectively, along with a constant $\epsilon_1 \in [0, 1)$ such that

$$\mathcal{Q} \in \mathcal{T}_{good} \Longrightarrow \frac{\Pr\big[T_{\mathrm{re}} = \mathcal{Q}\big]}{\Pr\big[T_{\mathrm{id}} = \mathcal{Q}\big]} \geq 1 - \epsilon_1. \tag{9}$$

It is then possible to show (see [11] for details) that

$$\Delta(T_{\mathrm{re}}, T_{\mathrm{id}}) \leq \epsilon_1 + \Pr\big[T_{\mathrm{id}} \in \mathcal{T}_{bad}\big] \tag{10}$$

is an upper bound on the distinguisher's advantage.

3 Pseudorandom Involution from Feistel-SF

Recall from the Introduction that the Feistel-SF (without final swap) is defined as

$$\overline{\Psi}_t^F[K, K, ..., K](L\|R) := \mathsf{swap} \circ \underbrace{\Psi_K^F \circ ... \circ \Psi_K^F \circ \Psi_K^F}_{t \text{ compositions}}(L\|R), \tag{11}$$

where $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ is a keyed function, $\Psi_{K_i}^F(A\|B) := B\|(A \oplus F_{K_i}(B))$ is the (1-round) Feistel permutation and $\mathsf{swap}(A\|B) := B\|A$ is the swap. We refer to Fig. 2 (left) for a 4-round version. Our proof will also rely on an idealized Feistel-SF construction, which has

$$\Psi_t[\mathbf{F}, \mathbf{F}, ..., \mathbf{F}](L\|R) := \mathsf{swap} \circ \underbrace{\Psi^{\mathbf{F}} \circ ... \circ \Psi^{\mathbf{F}} \circ \Psi^{\mathbf{F}}}_{t \text{ compositions}}(L\|R), \tag{12}$$

for a truly random function $\mathbf{F} : \{0,1\}^n \to \{0,1\}^n$ and an idealized Feistel permutation $\Psi^{\mathbf{F}}(A\|B) := B\|(A \oplus \mathbf{F}(B))$.

We first describe the attack breaking the PRI security of 3 rounds in Sect. 3.1. Then, in Sect. 3.2, we show that the 4-round scheme $\Psi_4^F[K, K, K, K]$ is a PRI up to $2^{n/2}$ adversarial queries.

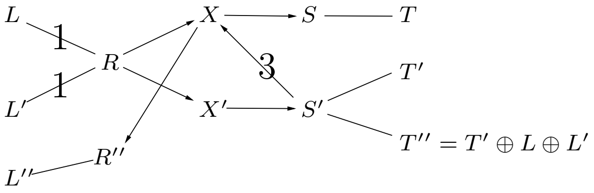

Fig. 3: Chains of values involved in our attack against 3-round Feistel-SF. The numbers 1 and 3 indicate if the chain appears in step 1 or 3.

3.1 3 Rounds do not yield PRIs

Instead of considering the 3-round Feistel-SF $\overline{\Psi}_3^F[K,K,K]$, we consider its random function-based variant. In addition, we consider $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$: if this can be distinguished then $\overline{\Psi}_3[\mathbf{F},\mathbf{F},\mathbf{F}]$ can be distinguished as well.

It is well known that 3-round Feistel $\Psi_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_3]$ suffers from CCA attacks: see e.g. [25, Sect. 2.5.3]. Since the forward and inverse oracle of $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$ are identical, the CCA attack on $\Psi_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_3]$ can be transformed to a CPA attack on $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$. In detail, consider a distinguisher D^{I} interacting with a $2n$-bit involution I that is either $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$ or a random involution $\mathbf{I}$. D^{I} proceeds as follows.

1. D^{I} chooses $L, L', R \in \{0,1\}^n$ with $L \neq L'$, and queries $\mathrm{I}(L\|R) \to T\|S$ and $\mathrm{I}(L'\|R) \to T'\|S'$;
2. If $S' = S$ or $S' = R$ then D^{I} outputs 1.[2] Otherwise, D^{I} proceeds into Step 3.
3. D^{I} sets $T'' \leftarrow T' \oplus L \oplus L'$ and queries $\mathrm{I}(T''\|S') \to L''\|R''$;
4. D^{I} outputs 1 if and only if $R'' = S' \oplus S \oplus R$.

We refer to Fig. 3 for the chains of values involved in the attack.

We first show that D^{I} always outputs 1 when I is $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$. For this, let $X = L \oplus \mathbf{F}_1(R)$ and $X' = L' \oplus \mathbf{F}_1(R)$. It then holds $\mathbf{F}_2(X) = R \oplus S$ and $\mathbf{F}_2(X') = R \oplus S'$. Meanwhile, it holds $X \oplus X' = L \oplus L'$. Then, let $X'' = T'' \oplus \mathbf{F}_1(S')$. It then holds $X'' = T'' \oplus \mathbf{F}_1(S') = (T' \oplus L \oplus L') \oplus (X' \oplus T') = X$ (as shown in Fig. 3), which further implies $R'' = S' \oplus \mathbf{F}_2(X'') = S' \oplus \mathbf{F}_2(X) = S' \oplus R \oplus S$. This means when I is $\overline{\Psi}_3[\mathbf{F}_1,\mathbf{F}_2,\mathbf{F}_1]$, D^{I} always outputs 1 at step 2 or 4.

On the other hand, when I is a random involution $\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n)$, we identify two events in the interaction:

- Event (B-1) occurs, if D finds $S' = S$ or $S' = R$ at step 2;
- Event (B-2) occurs, if D finds $R'' = S' \oplus S \oplus R$ at step 4.

Clearly,

$$\Pr\big[D^{\mathbf{I}} \text{ outputs } 1\big] \le \Pr\big[\text{(B-1)}\big] + \Pr\big[\text{(B-2)} \mid \neg\text{(B-1)}\big]$$

[2] The condition of $S' = R$ may not be necessary, but it simplifies the analysis of attack advantage by excluding the possibilities of $L\|R$ or $L'\|R$ being fixed points of I.

We now consider their probabilities. For (B-1), define

$$\mathcal{S}_1 := \big\{Y \in \{0,1\}^{2n} : \mathsf{right}_n(Y) = S \text{ or } \mathsf{right}_n(Y) = R\big\}.$$

Clearly, $|\mathcal{S}_1| \leq 2 \times 2^n$. By this and using Eq. (6), we reach

$$\begin{aligned}
&\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \text{(B-1)}\big] \\
&\leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(L'R) \in \mathcal{S}_1 \mid \mathbf{I} \vdash \big(\{LR, TS\}\big)\big] \\
&\leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(L'R) = L'R \mid \mathbf{I} \vdash \big(\{LR, TS\}\big)\big] \\
&\quad + \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(L'R) \in \mathcal{S}_1 \backslash \{L'R\} \mid \mathbf{I} \vdash \big(\{LR, TS\}\big)\big] \\
&\leq \frac{1}{\sqrt{2^{2n}-2}} + \frac{2^{n+1}-1}{\sqrt{2^{2n}-2} \times \sqrt{2^{2n}-3}} \leq \frac{3}{\sqrt{2^{2n}-2}}.
\end{aligned}$$

The last inequality follows from $2^{n+1} - 1 \leq \sqrt{2^{2n} - 3}$ as long as $n \geq 2$.

The analysis for (B-3) is similar: define

$$\mathcal{S}_2 := \{X \in \{0,1\}^{2n} : \mathsf{right}_n(X) = S' \oplus S \oplus R\}.$$

Clearly, $|\mathcal{S}_2| \leq 2^n$. Conditioned on that (B-1) did not occur, the query $\mathbf{I}(T''\|S')$ in step 3 is new. By these and using Eq. (6), we reach

$$\begin{aligned}
&\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \text{(B-2)}\big] \\
&\leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(T''\|S') \in \mathcal{S}_2 \mid \mathbf{I} \vdash \big(\{LR, TS\}, \{L'R, T'S'\}\big)\big] \\
&\leq \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(T''\|S') = T''\|S' \mid \mathbf{I} \vdash \big(\{LR, TS\}, \{L'R, T'S'\}\big)\big] \\
&\quad + \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I}(T''\|S') \in \mathcal{S}_1 \backslash \{T''\|S'\} \mid \mathbf{I} \vdash \big(\{LR, TS\}, \{L'R, T'S'\}\big)\big] \\
&\leq \frac{1}{\sqrt{2^{2n}-4}} + \frac{2^n-1}{\sqrt{2^{2n}-4} \times \sqrt{2^{2n}-5}} \leq \frac{3}{\sqrt{2^{2n}-4}}.
\end{aligned}$$

The last inequality follows from $2^n - 1 \leq \sqrt{2^{2n} - 5}$ as long as $n \geq 2$. Therefore,

$$\Pr\big[D^{\mathbf{I}} \text{ outputs } 1\big] \leq \frac{3}{\sqrt{2^{2n}-2}} + \frac{3}{\sqrt{2^{2n}-4}} \leq \frac{6}{\sqrt{2^{2n}-4}},$$

and the distinguishing advantage of D is at least $1 - 6/\sqrt{2^{2n} - 4} \approx 1$.

3.2 PRI from 4 Rounds

The positive result is formally stated as follows.

Theorem 1 (PRI from $\overline{\Psi}_4$) *The following holds for the 4-round Feistel-SF scheme $\overline{\Psi}_4^F[K,K,K,K]$:*

$$\mathbf{Adv}^{\mathsf{PRI}}_{\overline{\Psi}_4^F[K,K,K,K]}(q,t) \leq \mathbf{Adv}^{\mathsf{PRF}}_F\big(4q, t + O(q)\big) + \frac{6q^2 + 2\sqrt{2}q^{3/2}}{2^n}. \qquad (13)$$

We devote to prove Theorem 1 in the remaining of this subsection. As the first step, we modify the construction $\overline{\Psi}_4^F[K,K,K,K]$ and replace the PRF F_K with a true random function $\mathbf{F}:\{0,1\}^n\to\{0,1\}^n$. This yields a random function-based construction $\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]$. By a standard hybrid argument, it holds

$$\mathbf{Adv}^{\mathsf{SPRI}}_{\overline{\Psi}_4^F[K,K,K,K]}(q,t)\le\mathbf{Adv}^{\mathsf{SPRI}}_{\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]}(q)+\mathbf{Adv}^{\mathsf{PRF}}_{F}\big(4q,t+O(q)\big)\tag{14}$$

Our main analysis, which uses the H-coefficient method, focuses on the idealized construction $\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]$. For this, we follow Sect. 2 (or [20]) and write $\mathcal{Q}_I=\big(\{L_1R_1,T_1S_1\},\{L_2R_2,T_2S_2\},...,\{L_qR_q,T_qS_q\}\big)$ for the transcript of adversarial queries and responses, where $\{LR,TS\}\in\mathcal{Q}_I$ implies $I(LR)=TS$ or equivalently $I(TS)=LR$. This means any two distinct records $\{L_iR_i,T_iS_i\}$, $\{L_jR_j,T_jS_j\}\in\mathcal{Q}_I$ have both $L_iR_i\ne L_jR_j$ and $L_iR_i\ne T_jS_j$. W.l.o.g. we assume that D never asks such “redundant” queries, which yields $|\mathcal{Q}_I|=q$. We also assume that the number of “fixed points” in $\mathcal{Q}_I$, i.e., $\{LR,TS\}\in\mathcal{Q}_I$ with $LR=TS$, is ω.

For $\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]$, we will not use bad transcripts. Instead, we simply fix an arbitrary attainable transcript $\mathcal{Q}_I$. Since $\mathcal{Q}_I$ contains exactly ω fixed points, the ideal world probability is

$$\Pr\big[T_{\mathrm{id}}=\mathcal{Q}_I\big]=\Pr\big[\mathbf{I}\xleftarrow{\$}\mathcal{I}(2n):\mathbf{I}\vdash\mathcal{Q}_I\big]=\frac{T(2^{2n}-2q+\omega)}{T(2^{2n})}.\tag{15}$$

We now lower bound the probability

$$\Pr\big[T_{\mathrm{re}}=\mathcal{Q}_I\big]=\Pr\big[\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]\vdash\mathcal{Q}_I\big].$$

To this end, we follow [14] and define a “bad” predicate $\mathsf{Bad}(\mathbf{F})$ on $\mathbf{F}$, such that the event $\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]\vdash\mathcal{Q}_I$ is equivalent with the round function $\mathbf{F}$ satisfying $2q-\omega$ distinct equations, as long as $\mathsf{Bad}(\mathbf{F})$ is not fulfilled. The probability to have these equations is $1/2^{(2q-\omega)n}$ which is sufficiently close to $T(2^{2n}-2q+\omega)/T(2^{2n})$. Meanwhile, using the randomness of $\mathbf{F}$, upper bound of the probability $\Pr_{\mathbf{F}}\big[\mathsf{Bad}(\mathbf{F})\big]$ could be derived. These enable lower bounding $\Pr\big[T_{\mathrm{re}}=\mathcal{Q}_I\big]/\Pr\big[T_{\mathrm{id}}=\mathcal{Q}_I\big]$.

3.2.1 Bad predicate $\mathsf{Bad}(\mathbf{F})$

For any $\mathbf{F}\in\mathcal{F}(n,n)$, the predicate $\mathsf{Bad}(\mathbf{F})$ holds, if one of the following conditions is fulfilled:

- (B-1) There exist distinct records $\{L_iR_i,T_iS_i\},\{L_jR_j,T_jS_j\}\in\mathcal{Q}_I$ such that:
  - (B-11) $L_i\oplus\mathbf{F}(R_i)=L_j\oplus\mathbf{F}(R_j)$, or
  - (B-12) $T_i\oplus\mathbf{F}(S_i)=T_j\oplus\mathbf{F}(S_j)$, or
  - (B-13) $L_i\oplus\mathbf{F}(R_i)=T_j\oplus\mathbf{F}(S_j)$, or
  - (B-14) $L_j\oplus\mathbf{F}(R_j)=T_i\oplus\mathbf{F}(S_i)$.

- (B-2) There exists $\{L_iR_i, T_iS_i\} \in \mathcal{Q}_I$ such that $L_iR_i \neq T_iS_i$, though $L_i \oplus \mathbf{F}(R_i) = T_i \oplus \mathbf{F}(S_i)$;
- (B-3) There exist records $\{L_iR_i, T_iS_i\}, \{L_jR_j, T_jS_j\} \in \mathcal{Q}_I$ (which are not necessarily distinct) such that either $L_i \oplus \mathbf{F}(R_i) \in \{R_j, S_j\}$ or $T_i \oplus \mathbf{F}(S_i) \in \{R_j, S_j\}$.

First, consider any distinct records $\{L_iR_i, T_iS_i\}, \{L_jR_j, T_jS_j\} \in \mathcal{Q}_I$. As remarked in Sect. 2, it holds both $L_iR_i \neq L_jR_j$ and $L_iR_i \neq T_jS_j$: otherwise the two records are not distinct. Therefore,

- if $R_i = R_j$ then $L_i \neq L_j$ and $L_i \oplus \mathbf{F}(R_i) \neq L_j \oplus \mathbf{F}(R_j)$;
- Otherwise, since $\mathbf{F}(R_i)$ and $\mathbf{F}(R_j)$ are uniform and independent, the probability to have $L_i \oplus \mathbf{F}(R_i) = L_j \oplus \mathbf{F}(R_j)$ is $\frac{1}{2^n}$.

Thus, the probability to have $L_i \oplus \mathbf{F}(R_i) = L_j \oplus \mathbf{F}(R_j)$ is at most $1/2^n$. Similarly, the probability to have $T_i \oplus \mathbf{F}(S_i) = T_j \oplus \mathbf{F}(S_j)$ is at most $1/2^n$.

Regarding $L_i \oplus \mathbf{F}(R_i)$ versus $T_j \oplus \mathbf{F}(S_j)$, if $R_i = S_j$ then again $L_i \neq T_j$ since $L_iR_i \neq T_jS_j$, and $L_i \oplus \mathbf{F}(R_i) \neq T_j \oplus \mathbf{F}(S_j)$. Otherwise, $\mathbf{F}(R_i)$ and $\mathbf{F}(S_j)$ become uniform and independent again. The probability to have $L_i \oplus \mathbf{F}(R_i) = T_j \oplus \mathbf{F}(S_j)$ is thus at most $1/2^n$ as well. Similarly, the probability to have $L_j \oplus \mathbf{F}(R_j) = T_i \oplus \mathbf{F}(S_i)$ is at most $1/2^n$. Summing over the subconditions and the $\binom{q}{2}$ pairs of distinct records yields

$$\Pr\big[\text{(B-1)}\big] \leq \frac{4}{2^n} \times \binom{q}{2} \leq \frac{2q(q-1)}{2^n}. \tag{16}$$

The analysis of (B-2) is similar: for every $\{L_iR_i, T_iS_i\} \in \mathcal{Q}_I$ with $L_iR_i \neq T_iS_i$, (i) if $R_i = S_i$ then again $L_i \neq T_i$ and $L_i \oplus \mathbf{F}(R_i) \neq T_i \oplus \mathbf{F}(S_i)$; (ii) otherwise, $\mathbf{F}(R_i)$ and $\mathbf{F}(S_i)$ are uniform and independent, and the probability to have $L_i \oplus \mathbf{F}(R_i) = T_i \oplus \mathbf{F}(S_i)$ is $1/2^n$. Summing over the at most q records yields $\Pr\big[\text{(B-2)}\big] \leq q/2^n$.

Finally, for (B-3), consider any two records $\{L_iR_i, T_iS_i\}, \{L_jR_j, T_jS_j\} \in \mathcal{Q}_I$ (which may be the same). Since both $\mathbf{F}(R_i)$ and $\mathbf{F}(S_i)$ are uniform, the probability to have $L_i \oplus \mathbf{F}(R_i) \in \{R_j, S_j\}$ or $T_i \oplus \mathbf{F}(S_i) \in \{R_j, S_j\}$ is clearly $4/2^n$. Summing over the q^2 pairs of records yields $\Pr\big[\text{(B-3)}\big] \leq 4q^2/2^n$. Thus,

$$\Pr\big[\mathbf{F} \xleftarrow{\$} \mathbf{F}(n,n) : \mathsf{Bad}(\mathbf{F})\big] \leq \frac{2q(q-1)}{2^n} + \frac{q}{2^n} + \frac{4q^2}{2^n} \leq \frac{6q^2}{2^n}. \tag{17}$$

3.2.2 Further expanding

For any $\mathbf{F} \in \mathcal{F}(n,n)$, we define an "extended transcript" as

$$\mathcal{Q}^{out}(\mathbf{F}) = \big\{\big(R, \mathbf{F}(R)\big), \big(S, \mathbf{F}(S)\big)\big\}_{\{LR,TS\} \in \mathcal{Q}_I}.$$

We further define $\mathcal{T}^{out}$ as the set of all such extended transcripts, i.e.,

$$\mathcal{T}^{out} = \big\{\mathcal{Q}^{out}(\mathbf{F})\big\}_{\mathbf{F} \in \mathcal{F}(n,n)},$$

and a set of "good" extended transcripts based on functions that do not fulfill the bad predicate $\mathsf{Bad}(\mathbf{F})$, i.e.,

$$\mathcal{T}_{good}^{out} = \big\{\mathcal{Q}^{out}(\mathbf{F})\big\}_{\mathbf{F}\in\mathcal{F}(n,n),\ \neg\mathsf{Bad}(\mathbf{F})}.$$

With these, it holds

$$\begin{aligned}\Pr\big[T_{\mathrm{re}} = \mathcal{Q}_I\big] &= \Pr\big[\mathbf{F} \xleftarrow{\$} \mathcal{F}(n,n) : \overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I\big] \\ &= \sum_{\mathcal{Q}^{out}\in\mathcal{T}^{out}} \Pr_{\mathbf{F}}\big[\mathbf{F}\vdash\mathcal{Q}^{out} \wedge \overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I\big] \\ &\geq \sum_{\mathcal{Q}^{out}\in\mathcal{T}_{good}^{out}} \Pr_{\mathbf{F}}\big[\mathbf{F}\vdash\mathcal{Q}^{out} \wedge \overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I\big]. \qquad (18)\end{aligned}$$

We will prove

$$\Pr_{\mathbf{F}}\big[\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I \mid \mathbf{F}\vdash\mathcal{Q}^{out}\big] = \frac{1}{2^{(2q-\omega)n}} \qquad (19)$$

for any "good" extended transcript $\mathcal{Q}^{out}$, with which Eq. (18) further implies

$$\begin{aligned}\Pr\big[T_{\mathrm{re}} = \mathcal{Q}_I\big] &\geq \sum_{\mathcal{Q}^{out}\in\mathcal{T}_{good}^{out}} \Pr_{\mathbf{F}}\big[\mathbf{F}\vdash\mathcal{Q}^{out} \wedge \overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I\big] \\ &\geq \sum_{\mathcal{Q}^{out}\in\mathcal{T}_{good}^{out}} \frac{1}{2^{(2q-\omega)n}} \times \Pr_{\mathbf{F}}\big[\mathbf{F}\vdash\mathcal{Q}^{out}\big] \\ &= \frac{1}{2^{(2q-\omega)n}} \times \Pr_{\mathbf{F}}\big[\neg\mathsf{Bad}(\mathbf{F})\big] \geq \frac{1}{2^{(2q-\omega)n}} \times \Big(1 - \Pr_{\mathbf{F}}\big[\mathsf{Bad}(\mathbf{F})\big]\Big). \qquad (20)\end{aligned}$$

Probability for good functions. We now prove that any "good" extended transcript $\mathcal{Q}^{out} \in \mathcal{T}_{good}^{out}$ has

$$\Pr_{\mathbf{F}}\big[\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I \mid \mathbf{F}\vdash\mathcal{Q}^{out}\big] = \frac{1}{2^{(2q-\omega)n}}. \qquad (21)$$

For this, we list the records in $\mathcal{Q}_I$ as

$$\{L_1R_1, T_1S_1\}, ..., \{L_qR_q, T_qS_q\}, \qquad (22)$$

such that $L_iR_i = T_iS_i$ (to wit, being a "fixed point") if and only if $1 \leq i \leq \omega$.

Given $\mathbf{F}\vdash\mathcal{Q}^{out}$ in arbitrary, let $X_i = L_i \oplus \mathbf{F}(R_i)$ and $Y_i = T_i \oplus \mathbf{F}(S_i)$ for all $\{L_iR_i, T_iS_i\} \in \mathcal{Q}_I$. Actually, the values $X_1, Y_1, ..., X_q, Y_q$ are fixed by the records in $\mathcal{Q}^{out}$. Since $\mathcal{Q}^{out} \in \mathcal{T}_{good}^{out}$, $X_1, Y_1, ..., X_q, Y_q$ satisfy certain properties that will be elaborated.

It can be seen that for each $\{L_iR_i, T_iS_i\} \in \mathcal{Q}_I$ with $L_iR_i = T_iS_i$, it holds $X_i = Y_i$, and

$$\begin{aligned}\Pr\big[\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}](L_iR_i) = T_iS_i\big] &= \Pr\big[\mathbf{F}(X_i) = R_i \oplus Y_i \wedge \mathbf{F}(Y_i) = X_i \oplus S_i\big] \\ &= \Pr\big[\mathbf{F}(X_i) = R_i \oplus X_i\big].\end{aligned}$$

On the other hand, for each $\{L_iR_i, T_iS_i\} \in \mathcal{Q}_I$ with $L_iR_i \neq T_iS_i$, it holds $X_i \neq Y_i$: otherwise, the condition (B-2) is fulfilled and $\mathcal{Q}^{out}$ cannot be in $\mathcal{T}^{out}_{good}$. Thus

$$\Pr\big[\overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}](L_iR_i) = T_iS_i\big] = \Pr\big[\mathbf{F}(X_i) = R_i \oplus Y_i \wedge \mathbf{F}(Y_i) = X_i \oplus S_i\big].$$

By these and following the order fixed in Eq. (22), we reach

$$\begin{aligned}
&\Pr\big[\mathbf{F} \xleftarrow{\$} \mathcal{F}(n,n) : \overline{\Psi}_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}] \vdash \mathcal{Q}_I \mid \mathbf{F} \vdash \mathcal{Q}^{out}\big]\\
&= \Big(\prod_{i=1,\ldots,\omega} \Pr_{\mathbf{F}}\big[\mathbf{F}(X_i) = R_i \oplus X_i \mid \mathbf{F} \vdash \mathcal{Q}^{out} \wedge \mathbf{F}(X_j) = R_j \oplus X_j,\ j = 1,\ldots,i-1\big]\Big)\\
&\quad \times \Big(\prod_{i=\omega+1,\ldots,q} \Pr_{\mathbf{F}}\big[\mathbf{F}(X_i) = R_i \oplus Y_i \wedge \mathbf{F}(Y_i) = X_i \oplus S_i\\
&\qquad\qquad \mid \mathbf{F} \vdash \mathcal{Q}^{out} \wedge \mathbf{F}(X_j) = R_j \oplus X_j,\ j = 1,\ldots,\omega\\
&\qquad\qquad \wedge \mathbf{F}(X_j) = R_j \oplus Y_j \wedge \mathbf{F}(Y_j) = X_j \oplus S_j,\ j = \omega+1,\ldots,i-1\big]\Big). \qquad (23)
\end{aligned}$$

Conditioned on the event $\mathbf{F} \vdash \mathcal{Q}^{out}$ for $\mathcal{Q}^{out} \in \mathcal{T}^{out}_{good}$, the $2q - \omega$ induced $\mathbf{F}$-inputs $X_1, \ldots, X_\omega, X_{\omega+1}, Y_{\omega+1}, \ldots, X_q, Y_q$ appeared in Eq. (23) are distinct:

(i) $X_1, \ldots, X_q$ are distinct, otherwise (B-11) happens and $\mathcal{Q}^{out}$ cannot be in $\mathcal{T}^{out}_{good}$;

(ii) $Y_{\omega+1}, \ldots, Y_q$ are distinct, otherwise (B-12) happens and $\mathcal{Q}^{out}$ cannot be in $\mathcal{T}^{out}_{good}$;

(iii) For any $i \in \{1, \ldots, q\}$ and $j \in \{\omega + 1, q\}$, $i \neq j$, X_i and Y_j are distinct, otherwise either (B-13) happens or (B-14) happens;

(iv) For any $i \in \{\omega + 1, \ldots, q\}$, X_i and Y_i are distinct, otherwise (B-2) happens (since $L_iR_i \neq T_iS_i$ for all $i \in \{\omega + 1, \ldots, q\}$.

Conditioned on $\mathbf{F} \vdash \mathcal{Q}^{out}$, the $2q-\omega$ entries $\mathbf{F}(X_1), \ldots, \mathbf{F}(X_q), \mathbf{F}(Y_{\omega+1}), \ldots, \mathbf{F}(Y_q)$ remains fresh and uniformly distributed, since the condition $\mathbf{F} \vdash \mathcal{Q}^{out}$ only fixes the entries $\mathbf{F}(R_1), \mathbf{F}(S_1), \ldots, \mathbf{F}(R_q), \mathbf{F}(S_q)$ which do not overlap with $\mathbf{F}(X_1), \ldots, \mathbf{F}(Y_q)$ by $\neg$(B-3). Therefore, we have Eq. (23)= $1/2^{(2q-\omega)n}$.

3.2.3 Concluding

Gathering Eqs. (15), (17) and (18) yields

$$\begin{aligned}
\frac{\Pr\big[T_{\text{re}} = \mathcal{Q}\big]}{\Pr\big[T_{\text{id}} = \mathcal{Q}\big]} &\geq \left(1 - \frac{6q^2}{2^n}\right) \times \left(\frac{1}{2^n}\right)^{2q-\omega} \Big/ \frac{T(2^{2n} - 2q + \omega)}{T(2^{2n})}\\
&= \left(1 - \frac{6q^2}{2^n}\right) \times \left(\prod_{\ell=0}^{2q-\omega-1} \frac{T(2^{2n} - \ell)}{2^n \times T(2^{2n} - \ell - 1)}\right)\\
&\geq \left(1 - \frac{6q^2}{2^n}\right) \times \left(\prod_{\ell=0}^{2q-1} \frac{\sqrt{2^{2n} - \ell}}{2^n}\right)\\
&\geq \left(1 - \frac{6q^2}{2^n}\right) \times \left(\prod_{\ell=0}^{2q-1} \frac{\sqrt{2^{2n} - 2q}}{2^n}\right).
\end{aligned}$$

When $\ell \le 2q \le 2^{2n}$, it can be proven

$$\sqrt{2^{2n} - 2q} \ge 2^n - \sqrt{2q},$$

which yields

$$\begin{aligned}\frac{\Pr\big[T_{\mathrm{re}} = \mathcal{Q}\big]}{\Pr\big[T_{\mathrm{id}} = \mathcal{Q}\big]} &\ge \left(1 - \frac{6q^2}{2^n}\right) \times \left(\prod_{\ell=0}^{2q-1} \frac{2^n - \sqrt{2q}}{2^n}\right) \\ &\ge \left(1 - \frac{6q^2}{2^n}\right) \times \left(1 - \frac{2q\sqrt{2q}}{2^n}\right) \ge 1 - \frac{6q^2 + 2\sqrt{2}q^{3/2}}{2^n}.\end{aligned}$$

By this and by Eq. (9) and (10), we have $\mathbf{Adv}^{\mathsf{PRI}}_{\Psi_4[\mathbf{F},\mathbf{F},\mathbf{F},\mathbf{F}]}(q) \le \frac{6q^2+2\sqrt{2}q^{3/2}}{2^n}$, which plus Eq. (14) yield Eq. (13).

4 Pseudorandom Involution from Naor-Reingold

This section proves PRI security for $\mathsf{NR}^F[K_P, K, K, K_P]$, the "mirrored" variant of the Naor-Reingold construction [30]. Recall from the Introduction that this variant is defined as

$$\mathsf{NR}^F[K_P, K, K, K_P](X) := P^{-1}\big(K_P, \overline{\Psi}_2^F[K, K]\big(P(K_P, X)\big)\big),$$

where $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ is a PRF and $P : \mathcal{K}_P \times \{0,1\}^{2n} \to \{0,1\}^{2n}$ is a ε-right-universal permutation family on the $2n$-bit strings (this notion was due to [42]), meaning that

$$\Pr_{K_P \xleftarrow{\$} \mathcal{K}_P}\big[\mathsf{right}_n\big(P(K_P, X)\big) = \mathsf{right}_n\big(P(K_P, X')\big)\big] \le \varepsilon$$

for all distinct $X, X' \in \{0,1\}^{2n}$. Candidates for P include pairwise-independent permutations and 1-round Feistel built upon a pairwise independent hash function $H : \mathcal{K}_P \times \{0,1\}^n \to \{0,1\}^n$, i.e., $P\big(K_P, (L, R)\big) = \big(R, H(K_P, R) \oplus L\big)$.

Our second result focuses on the "mirrored" variant $\mathsf{NR}^F[K_P, K, K, K_P]$.

Theorem 2 (PRI from NR) *Let P be a ε-right-universal permutation family and $F : \mathcal{K} \times \{0,1\}^n \to \{0,1\}^n$ be a PRF. Then, for the "mirrored" Naor-Reingold $\mathsf{NR}^F[K_P, K, K, K_P]$, it holds*

$$\mathbf{Adv}^{\mathsf{PRI}}_{\mathsf{NR}^F[K_P,K,K,K_P]}(q, t) \le \mathbf{Adv}^{\mathsf{PRF}}_F\big(2q, t + O(q)\big) + 2\varepsilon q^2 + \frac{2\sqrt{2}q^{3/2}}{2^n}. \quad (24)$$

Proof We also begin by replacing the PRF F_K with a random function $\mathbf{F} : \{0,1\}^n \to \{0,1\}^n$ to yield the idealized Naor-Reingold $\mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P]$. It also holds

$$\begin{aligned}\mathbf{Adv}^{\mathsf{PRI}}_{\mathsf{NR}^F[K_P,K,K,K_P]}(q, t) \le{}& \mathbf{Adv}^{\mathsf{PRF}}_F\big(2q, t + O(q)\big) \\ &+ \mathbf{Adv}^{\mathsf{PRI}}_{\mathsf{NR}[K_P,\mathbf{F},\mathbf{F},K_P]}(q). \quad (25)\end{aligned}$$

We thus focus on the idealized $\mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P]$ and use the H-coefficient method. For this, we follow Sect. 2 and write $\mathcal{Q}_I = \big(\{X_1, Y_1\}, ..., \{X_q, Y_q\}\big)$ for the transcript of adversarial queries and responses. We also assume that D never asks such "redundant" queries, i.e., $|\mathcal{Q}_I| = q$. Meanwhile, the number of fixed points in $\mathcal{Q}_I$, i.e., $\{X, Y\} \in \mathcal{Q}_I$ with $X = Y$, is ω.

We follow [11] and provide the distinguisher D, at the end of its interaction, with the actual right-universal permutation key K_P when it is interacting with $\mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P]$, or with a dummy key K_P selected uniformly from $\mathcal{K}_P$ when it is interacting with $\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n)$. This is without loss of generality since the distinguisher is free to ignore this additional information. Therefore, the transcript is $\mathcal{Q} = (\mathcal{Q}_I, K_P)$.

Bad transcripts. An attainable transcript $\mathcal{Q} = (\mathcal{Q}_I, K_P)$ is *bad*, if one of the following conditions is fulfilled:

- (C-1) There exist distinct records $\{X_i, Y_i\}, \{X_j, Y_j\} \in \mathcal{Q}_I$ such that:
  - (C-11) $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, X_j)\big)$, or
  - (C-12) $\mathsf{right}_n\big(P(K_P, Y_i)\big) = \mathsf{right}_n\big(P(K_P, Y_j)\big)$, or
  - (C-13) $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, Y_j)\big)$, or
  - (C-14) $\mathsf{right}_n\big(P(K_P, Y_i)\big) = \mathsf{right}_n\big(P(K_P, X_j)\big)$.
- (C-2) There exists $\{X_i, Y_i\} \in \mathcal{Q}_I$ with $X_i \neq Y_i$, but $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, Y_i)\big)$.

First, consider any two distinct records $\{X_i, Y_i\}, \{X_j, Y_j\} \in \mathcal{Q}_I$. As per our convention, it holds both $X_i \neq X_j$ and $X_i \neq Y_j$. Therefore, due to the ε-right-universality of P, the probability to have one of the following four equalities

- $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, X_j)\big)$,
- $\mathsf{right}_n\big(P(K_P, Y_i)\big) = \mathsf{right}_n\big(P(K_P, Y_j)\big)$,
- $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, Y_j)\big)$, and
- $\mathsf{right}_n\big(P(K_P, Y_i)\big) = \mathsf{right}_n\big(P(K_P, X_j)\big)$

is at most 4ε. Summing over the subconditions and the $\binom{q}{2}$ pairs of distinct records yields

$$\Pr\big[\text{(C-1)}\big] \leq 4\varepsilon \times \binom{q}{2} \leq 2\varepsilon q(q-1). \tag{26}$$

Regarding (C-2), for every $\{X_i, Y_i\} \in \mathcal{Q}_I$ with $X_i \neq Y_i$, the probability to have $\mathsf{right}_n\big(P(K_P, X_i)\big) = \mathsf{right}_n\big(P(K_P, Y_i)\big)$ is ε. Thus $\Pr\big[\text{(C-2)}\big] \leq \varepsilon q$. Thus,

$$\Pr\big[T_{\text{id}} \in \mathcal{T}_{bad}\big] \leq 2\varepsilon q(q-1) + \varepsilon q \leq 2\varepsilon q^2. \tag{27}$$

Probabilities of good transcripts. Fix a good transcript $\mathcal{Q} = (\mathcal{Q}_I, K_P)$. Since $\mathcal{Q}_I$ contains ω fixed points, the ideal world probability is bounded by

$$\Pr\big[T_{\text{id}} = \mathcal{Q}\big] = \Pr\big[K_P' \xleftarrow{\$} \mathcal{K}_P : K_P' = K_P\big] \times \Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I} \vdash \mathcal{Q}_I\big]. \tag{28}$$

We now lower bound the probability

$$\begin{aligned}\Pr\big[T_{\mathrm{re}} = \mathcal{Q}\big] &= \Pr\big[K'_P \xleftarrow{\$} \mathcal{K}_P : K'_P = K_P\big] \\ &\quad \times \Pr\big[\mathbf{F} \xleftarrow{\$} \mathcal{F}(n,n) : \mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P] \vdash \mathcal{Q}_I\big].\end{aligned}$$

Let's list the records in $\mathcal{Q}_I$ as

$$\{X_1, Y_1\}, ..., \{X_q, Y_q\}, \tag{29}$$

such that $X_i = Y_i$ if and only if $1 \le i \le \omega$.

Given $\mathbf{F} \in \mathcal{F}(n,n)$ in arbitrary, let $L_i \| R_i = P(K_P, X_i)$ and $T_i \| S_i = P(K_P, Y_i)$ for all $\{X_i, Y_i\} \in \mathcal{Q}_I$. It can be seen that for each $\{X_i, Y_i\} \in \mathcal{Q}_I$ with $X_i = Y_i$, it holds $R_i = S_i$ and $L_i = T_i$, and

$$\begin{aligned}\Pr\big[\mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P](X_i) = Y_i\big] &= \Pr\big[\mathbf{F}(R_i) = L_i \oplus S_i \wedge \mathbf{F}(S_i) = R_i \oplus T_i\big] \\ &= \Pr\big[\mathbf{F}(R_i) = L_i \oplus S_i\big].\end{aligned}$$

On the other hand, for each $\{X_i, Y_i\} \in \mathcal{Q}_I$ with $X_i \neq Y_i$, it holds $R_i \neq S_i$: otherwise, the condition (C-2) is fulfilled and $\mathcal{Q} = (\mathcal{Q}_I, P)$ is not good. Thus

$$\Pr\big[\mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P](X_i) = Y_i\big] = \Pr\big[\mathbf{F}(R_i) = L_i \oplus S_i \wedge \mathbf{F}(S_i) = R_i \oplus T_i\big].$$

By these and following the order fixed in Eq. (29), we reach

$$\begin{aligned}&\Pr\big[\mathbf{F} \xleftarrow{\$} \mathcal{F}(n,n) : \mathsf{NR}[P, \mathbf{F}, \mathbf{F}, P^{-1}] \vdash \mathcal{Q}_I\big] \\ &= \Bigg(\prod_{i=1,\ldots,\omega} \Pr_F\big[\mathbf{F}(R_i) = L_i \oplus R_i \mid \mathbf{F}(R_j) = L_j \oplus R_j,\ j = 1, ..., i-1\big]\Bigg) \\ &\quad \times \Bigg(\prod_{i=\omega+1,\ldots,q} \Pr\big[\mathbf{F}(R_i) = L_i \oplus S_i \wedge \mathbf{F}(S_i) = R_i \oplus T_i \\ &\qquad \mid \mathbf{F}(R_j) = L_j \oplus R_j,\ j = 1, ..., \omega \\ &\qquad \wedge \mathbf{F}(R_j) = L_j \oplus S_j \wedge \mathbf{F}(S_j) = R_j \oplus T_j,\ j = \omega+1, ..., i-1\big]\Bigg).\end{aligned} \tag{30}$$

Since $\mathcal{Q} = (\mathcal{Q}_I, K_P)$ is good, the induced $2q-\omega$ $\mathbf{F}$-inputs $R_1, ..., R_\omega, R_{\omega+1}, S_{\omega+1}, ..., R_q, S_q$ appeared in Eq. (30) are distinct:

(i) $R_1, ..., R_q$ are distinct, otherwise (C-11) happens and $\mathcal{Q}$ is not good;
(ii) $S_{\omega+1}, ..., S_q$ are distinct, otherwise (C-12) happens and $\mathcal{Q}$ is not good;
(iii) For any $i \in \{1, ..., q\}$ and $j \in \{\omega+1, q\}$, $i \neq j$, R_i and S_j are distinct, otherwise either (C-13) happens or (C-14) happens;
(iv) For any $i \in \{\omega+1, q\}$, R_i and S_i are distinct, otherwise (C-2) happens.

Therefore, we have Eq. (30)$= 1/2^{(2q-\omega)n}$, and further

$$\begin{aligned}\frac{\Pr\big[T_{\mathrm{re}} = \mathcal{Q}\big]}{\Pr\big[T_{\mathrm{id}} = \mathcal{Q}\big]} &= \frac{\Pr\big[\mathbf{F} \xleftarrow{\$} \mathcal{F}(n,n) : \mathsf{NR}[K_P, \mathbf{F}, \mathbf{F}, K_P] \vdash \mathcal{Q}_I\big]}{\Pr\big[\mathbf{I} \xleftarrow{\$} \mathcal{I}(2n) : \mathbf{I} \vdash \mathcal{Q}_I\big]} \\ &\ge \left(\frac{1}{2^n}\right)^{2q-\omega} \times \frac{T(2^{2n})}{T(2^{2n} - 2q + \omega)} \ \ge\ 1 - \frac{2q\sqrt{2q}}{2^n},\end{aligned} \tag{31}$$

where the inequality has been proven in Sect. 3.2. Gathering Eqs. (25), (27) and (31) and using Eq. (9) and (10) yield Eq. (24). □

5 Conclusion

We prove pseudorandom involution (PRIs) security for two Feistel variants with function reusing: the 4-round Feistel $\overline{\Psi}_4^F[K, K, K, K]$ using a single key, and the "mirrored" Naor-Reingold construction $\mathsf{NR}^F[P, K, K, P^{-1}]$. To complement, we also exhibit a simple attack breaking the PRIs security of 3-round Feistel $\overline{\Psi}_3$. Besides characterizing cryptographic strength of the two Feistel variants, this also exhibits the first PRF-to-PRI transformations.

An intriguing direction is to design pseudorandom involution "from the scratch", i.e., "involutory blockciphers". Regarding provable security, it is natural to ask if the "mirrored" 5- and 6-round Feistel, i.e., $\overline{\Psi}_5^F[K_1, K_2, K_3, K_2, K_1]$, $\overline{\Psi}_6^F[K_1, K_2, K_3, K_3, K_2, K_1]$ and their further simplifications, yield PRIs with beyond-birthday security. In addition, unlike a PRP, a PRI cannot instantiate a PRF. It thus remains open to construct PRFs from PRIs directly, as shown in Fig. 1. In particular, would XOR of PRIs and truncated PRIs yield PRFs?

Acknowledgments

We thank the anonymous reviewers for insightful comments that help us significantly improve the quality of our paper.

Data Deposition Information: No datasets have been used

Conflict of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article.

Funding: The authors are supported by the National Natural Science Foundation of China (Grant No. 62002202) and the Taishan Scholars Program (for Young Scientists) of Shandong.

References

1. Avanzi, R.: A salad of block ciphers. Cryptology ePrint Archive (2016)
2. Avanzi, R.: The QARMA block cipher family. IACR Trans. Symm. Cryptol. 2017(1), 4–44 (2017)
3. Banik, S., Isobe, T., Liu, F., Minematsu, K., Sakamoto, K.: Orthros: A low-latency PRF. IACR Trans. Symm. Cryptol. 2021(1), 37–77 (2021)
4. Beierle, C., Leander, G., Moradi, A., Rasoolzadeh, S.: CRAFT: Lightweight tweakable block cipher with efficient protection against DFA attacks. IACR Trans. Symm. Cryptol. 2019(1), 5–45 (2019)
5. Bellizia, D., Bronchain, O., Cassiers, G., Grosso, V., Guo, C., Momin, C., Pereira, O., Peters, T., Standaert, F.X.: Mode-level vs. implementation-level physical security in symmetric cryptography - A practical guide through the leakage-resistance jungle. In: Shacham, H., Boldyreva, A. (eds.) CRYPTO 2020, Part I. pp. 369–400. LNCS, Springer, Heidelberg (Aug 2020)
6. Berti, F., Guo, C., Pereira, O., Peters, T., Standaert, F.X.: TEDT: a leakage-resistant AEAD mode. IACR TCHES 2020(1), 256–320 (2019), `https://tches.iacr.org/index.php/TCHES/article/view/8400`

7. Berti, F., Guo, C., Peters, T., Shen, Y., Standaert, F.: Secure message authentication in the presence of leakage and faults. IACR Trans. Symmetric Cryptol. 2023(1), 288–315 (2023), https://doi.org/10.46586/tosc.v2023.i1.288-315
8. Berti, F., Pereira, O., Peters, T., Standaert, F.X.: On leakage-resilient authenticated encryption with decryption leakages. IACR Trans. Symm. Cryptol. 2017(3), 271–293 (2017)
9. Borghoff, J., Canteaut, A., Güneysu, T., Kavun, E.B., Knežević, M., Knudsen, L.R., Leander, G., Nikov, V., Paar, C., Rechberger, C., Rombouts, P., Thomsen, S.S., Yalçin, T.: PRINCE - A low-latency block cipher for pervasive computing applications - extended abstract. In: Wang, X., Sako, K. (eds.) ASIACRYPT 2012. LNCS, vol. 7658, pp. 208–225. Springer, Heidelberg (Dec 2012)
10. Bozilov, D., Eichlseder, M., Knezevic, M., Lambin, B., Leander, G., Moos, T., Nikov, V., Rasoolzadeh, S., Todo, Y., Wiemer, F.: Princev2 - more security for (almost) no overhead. In: Dunkelman, O., Jr., M.J.J., O'Flynn, C. (eds.) Selected Areas in Cryptography - SAC 2020 - 27th International Conference, Halifax, NS, Canada (Virtual Event), October 21-23, 2020, Revised Selected Papers. Lecture Notes in Computer Science, vol. 12804, pp. 483–511. Springer (2020), https://doi.org/10.1007/978-3-030-81652-0_19
11. Chen, S., Steinberger, J.P.: Tight security bounds for key-alternating ciphers. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 327–350. Springer, Heidelberg (May 2014)
12. Cheng, H., Heys, H.M., Wang, C.: PUFFIN: A novel compact block cipher targeted to embedded digital systems. In: Fanucci, L. (ed.) 11th Euromicro Conference on Digital System Design: Architectures, Methods and Tools, DSD 2008, Parma, Italy, September 3-5, 2008. pp. 383–390. IEEE Computer Society (2008), https://doi.org/10.1109/DSD.2008.34
13. Chowla, S., Herstein, I., Moore, W.: On recursions connected with symmetric groups i. Canadian Journal of Mathematics 3, 328–334 (1951)
14. Cogliati, B., Dodis, Y., Katz, J., Lee, J., Steinberger, J.P., Thiruvengadam, A., Zhang, Z.: Provable security of (tweakable) block ciphers based on substitution-permutation networks. In: Shacham, H., Boldyreva, A. (eds.) CRYPTO 2018, Part I. LNCS, vol. 10991, pp. 722–753. Springer, Heidelberg (Aug 2018)
15. Daemen, J., Peeters, M., Van Assche, G., Rijmen, V.: Nessie proposal: NOEKEON. In: First open NESSIE workshop. pp. 213–230 (2000)
16. Dai, W., Hoang, V.T., Tessaro, S.: Information-theoretic indistinguishability via the chi-squared method. In: Katz, J., Shacham, H. (eds.) CRYPTO 2017, Part III. LNCS, vol. 10403, pp. 497–523. Springer, Heidelberg (Aug 2017)
17. Dutta, A., Nandi, M., Saha, A.: Proof of mirror theory for $\xi_{\max} = 2$. IEEE Trans. Inf. Theory 68(9), 6218–6232 (2022), https://doi.org/10.1109/TIT.2022.3171178
18. Gilboa, S., Gueron, S., Morris, B.: How many queries are needed to distinguish a truncated random permutation from a random function? Journal of Cryptology 31(1), 162–171 (Jan 2018)
19. Hall, C., Wagner, D., Kelsey, J., Schneier, B.: Building PRFs from PRPs. In: Krawczyk, H. (ed.) CRYPTO'98. LNCS, vol. 1462, pp. 370–389. Springer, Heidelberg (Aug 1998)
20. Lee, J.: Key alternating ciphers based on involutions. Des. Codes Cryptogr. 86(5), 955–988 (2018), https://doi.org/10.1007/s10623-017-0371-3
21. Lee, J., Koo, B.: Security of the misty structure using involutions as round functions. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. 93-A(9), 1612–1619 (2010), https://doi.org/10.1587/transfun.E93.A.1612
22. Li, S., Sun, S., Li, C., Wei, Z., Hu, L.: Constructing low-latency involutory MDS matrices with lightweight circuits. IACR Trans. Symm. Cryptol. 2019(1), 84–117 (2019)
23. Luby, M., Rackoff, C.: How to Construct Pseudorandom Permutations from Pseudorandom Functions. SIAM J. Comput. 17(2), 373–386 (1988)
24. Maurer, U.M., Oswald, Y.A., Pietrzak, K., Sjödin, J.: Luby-Rackoff ciphers from weak round functions? In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 391–408. Springer, Heidelberg (May / Jun 2006)
25. Nachef, V., Patarin, J., Volte, E.: Feistel Ciphers - Security Proofs and Cryptanalysis. Springer (2017)

26. Nandi, M.: An efficient SPRP-secure construction based on pseudo random involution. Cryptology ePrint Archive, Report 2008/092 (2008), `https://eprint.iacr.org/2008/092`
27. Nandi, M.: Improving upon HCTR and matching attacks for hash-counter-hash approach. Cryptology ePrint Archive, Report 2008/090 (2008), `https://eprint.iacr.org/2008/090`
28. Nandi, M.: The characterization of Luby-Rackoff and its optimum single-key variants. In: Gong, G., Gupta, K.C. (eds.) INDOCRYPT 2010. LNCS, vol. 6498, pp. 82–97. Springer, Heidelberg (Dec 2010)
29. Nandi, M.: On the optimality of non-linear computations of length-preserving encryption schemes. In: Iwata, T., Cheon, J.H. (eds.) ASIACRYPT 2015, Part II. LNCS, vol. 9453, pp. 113–133. Springer, Heidelberg (Nov / Dec 2015)
30. Naor, M., Reingold, O.: On the construction of pseudorandom permutations: Luby-Rackoff revisited. Journal of Cryptology 12(1), 29–66 (Jan 1999)
31. Naor, M., Reingold, O.: Constructing pseudo-random permutations with a prescribed structure. Journal of Cryptology 15(2), 97–102 (Mar 2002)
32. Patarin, J.: How to construct pseudorandom and super pseudorandom permutations from one single pseudorandom function. In: Rueppel, R.A. (ed.) EUROCRYPT'92. LNCS, vol. 658, pp. 256–266. Springer, Heidelberg (May 1993)
33. Patarin, J.: The "coefficients H" technique (invited talk). In: Avanzi, R.M., Keliher, L., Sica, F. (eds.) SAC 2008. LNCS, vol. 5381, pp. 328–345. Springer, Heidelberg (Aug 2009)
34. Pieprzyk, J.: How to construct pseudorandom permutations from single pseudorandom functions. In: Damgård, I. (ed.) EUROCRYPT'90. LNCS, vol. 473, pp. 140–150. Springer, Heidelberg (May 1991)
35. Piret, G., Quisquater, J.: Security of the MISTY structure in the luby-rackoff model: Improved results. In: Handschuh, H., Hasan, M.A. (eds.) Selected Areas in Cryptography, 11th International Workshop, SAC 2004, Waterloo, Canada, August 9-10, 2004, Revised Selected Papers. Lecture Notes in Computer Science, vol. 3357, pp. 100–113. Springer (2004), `https://doi.org/10.1007/978-3-540-30564-4\_7`
36. Rueppel, R.A.: On the security of Schnorr's pseudo random generator. In: Quisquater, J.J., Vandewalle, J. (eds.) EUROCRYPT'89. LNCS, vol. 434, pp. 423–428. Springer, Heidelberg (Apr 1990)
37. Sadeghiyan, B., Pieprzyk, J.: A construction for super pseudorandom permutations from a single pseudorandom function. In: Rueppel, R.A. (ed.) EUROCRYPT'92. LNCS, vol. 658, pp. 267–284. Springer, Heidelberg (May 1993)
38. Sadeghiyan, B., Pieprzyk, J.: On necessary and sufficient conditions for the construction of super pseudorandom permutations. In: Imai, H., Rivest, R.L., Matsumoto, T. (eds.) ASIACRYPT'91. LNCS, vol. 739, pp. 194–209. Springer, Heidelberg (Nov 1993)
39. Saha, S., Khairallah, M., Peyrin, T.: Exploring integrity of aeads with faults: Definitions and constructions. IACR Trans. Symmetric Cryptol. 2022(4), 291–324 (2022), `https://doi.org/10.46586/tosc.v2022.i4.291-324`
40. Schnorr, C.P.: On the construction of random number generators and random function generators. In: Günther, C.G. (ed.) EUROCRYPT'88. LNCS, vol. 330, pp. 225–232. Springer, Heidelberg (May 1988)
41. Smith, J.L.: Design of Lucifer, a Cryptographic Device for Data Communications. IBM Thomas J. Watson Research Center (1971)
42. Soni, P., Tessaro, S.: Naor-Reingold goes public: The complexity of known-key security. In: Nielsen, J.B., Rijmen, V. (eds.) EUROCRYPT 2018, Part III. LNCS, vol. 10822, pp. 653–684. Springer, Heidelberg (Apr / May 2018)
43. Standaert, F.X., Piret, G., Rouvroy, G., Quisquater, J.J., Legat, J.D.: ICEBERG: An involutional cipher efficient for block encryption in reconfigurable hardware. In: Roy, B.K., Meier, W. (eds.) FSE 2004. LNCS, vol. 3017, pp. 279–299. Springer, Heidelberg (Feb 2004)
44. Zheng, Y., Matsumoto, T., Imai, H.: Impossibility and optimality results on constructing pseudorandom permutations (extended abstract). In: Quisquater, J.J., Vandewalle, J. (eds.) EUROCRYPT'89. LNCS, vol. 434, pp. 412–422. Springer, Heidelberg (Apr 1990)