J Cryptol (2018) 31:521–536
https://doi.org/10.1007/s00145-017-9262-z

Journal of
CRYPTOLOGY

Deterministic Encryption with the Thorp Shuffle

Ben Morris
Department of Mathematics, University of California, Davis, CA 95616, USA

Phillip Rogaway and Till Stegers[*]
Department of Computer Science, University of California, Davis, CA 95616, USA
rogaway@cs.ucdavis.edu

Communicated by Kenneth Paterson.

Received 3 April 2011 / Revised 26 July 2017
Online publication 25 August 2017

Abstract. We analyze the security of the Thorp shuffle, or, equivalently, a maximally unbalanced Feistel network. Roughly said, the Thorp shuffle on N cards mixes any $N^{1-1/r}$ of them in $O(r \lg N)$ steps. Correspondingly, making $O(r)$ passes of maximally unbalanced Feistel over an n-bit string ensures CCA security to $2^{n(1-1/r)}$ queries. Our results, which employ Markov chain techniques, particularly couplings, help to justify a practical, although still relatively inefficient, blockcipher-based scheme for deterministically enciphering credit card numbers and the like.

Keywords. Card shuffling, Coupling, Format-preserving encryption, Modes of operation, Symmetric encryption, Thorp shuffle, Unbalanced Feistel network.

1. Introduction

1.1. *Small-Space Encryption*

Suppose you want to encrypt a 9-decimal-digit plaintext, say a US social security number, into a ciphertext that is again a 9-decimal-digit number. A shared key K is used to control the encryption. Syntactically, you seek a cipher $E\colon \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ where $\mathcal{M} = \{0, 1, \ldots, N-1\}$, $N = 10^9$, and $E_K = E(K, \cdot)$ is a permutation for each key $K \in \mathcal{K}$. You aim to construct your scheme from a well-known primitive, say AES, and to prove your scheme is as secure as the primitive from which you start.

The problem is harder than it sounds. You cannot just encode each plaintext $M \in \mathcal{M}$ as a 128-bit string and then apply AES, say, as that will return a 128-bit string and projecting

[*] Research performed while at the University of California, Davis.

back onto $\mathcal{M}$ will destroy permutivity. Standard blockcipher modes of operation are of no use, and constructions such as balanced Feistel [21,35] or Benes [1,34] have security that falls off by, at best, the square root of the size of the domain, N. Here N is so small that such a result may provide no practically useful guarantee.

The above *small-space encryption* problem was first investigated by Black and Rogaway [7], but those authors could find no efficient and provably secure solution for small N values. While one does have the option to spend $\Omega(N)$ time and use a construction such as the "prefix cipher" [7], this rapidly becomes unattractive. This paper provides a provably secure solution for enciphering on these small domains.

1.2. *Thorp Shuffle*

Our approach is based on the *Thorp shuffle* [42], which works like this. Suppose you want to shuffle N cards, where N is even. Cut the deck into two equal piles. Drop the bottom card from either the left or right pile according to the outcome of a fair coin flip, and then, drop the card from the bottom of the other pile. Continue in this way, flipping $N/2$ independent coins and using each to decide whether you drop cards left-then-right or right-then-left. This is one *round* of the shuffle; repeat for as many rounds as you like. Expressed a bit more algebraically, for each round $r = 1, 2, \ldots, R$ the cards at positions x and $x + N/2$, where $x \in \{0, \ldots, N/2 - 1\}$, are moved either to positions $2x$ and $2x + 1$ or else to positions $2x + 1$ and $2x$, which one of these possibilities being determined by a uniform coin flip $c \in \{0, 1\}$. See the left-hand side of Fig. 1. Let Th$[N, R]$ denote the Thorp shuffle with message space $\mathcal{M} = \{0, \ldots, N-1\}$ and R rounds.

The potential utility of the Thorp shuffle to cryptography and complexity theory was first noticed by Naor nearly 30 years ago [31, p. 62], [37, p. 17]. He observed that the Thorp shuffle is *oblivious* in the following sense: one can trace the route of any given card in the deck without attending to the remaining cards in the deck. If the Thorp shuffle mixes cards quickly enough, this property would make it suitable for small-space encryption. Namely, the random bit c used for cards x and $x + N/2$ at round r could be determined by applying a pseudorandom function F, keyed by some underlying key K, to x and r. Conceptually, the string K compactly names all of the $(N/2) \cdot R$ random

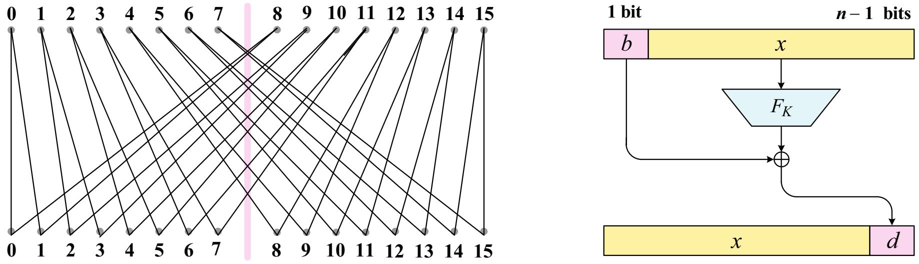

Fig. 1. Two views of the Thorp shuffle (one round). *Left*: each card is paired with the one $N/2$ positions away. For cards at positions x and $x + N/2$, a random bit c (not shown) determines whether the cards get mapped to $2x$ and $2x + 1$ or to $2x + 1$ and $2x$. *Right*: each card is regarded as an n-bit string X (assume $N = 2^n$). Now card $b \parallel x$ gets sent to $x \parallel b \oplus F_K(x)$ for a uniform (and round-dependent) random function F_K.

bits that would be needed to shuffle the entire deck. But because the Thorp shuffle is oblivious, only R of these bits, so only that many pseudorandom function (PRF) calls, would be needed to encipher an input x.

1.3. *Feistel View*

There are a variety of alternative views of what goes on in the Thorp shuffle. The one most resonant to cryptographers is this. Suppose that $N = 2^n$ is a power of two. In this case, the Thorp shuffle coincides with a maximally unbalanced Feistel network. In an unbalanced Feistel network [39] the left and right portions in the n-bit string that is being enciphered may have different lengths. Throughout this paper, "maximally unbalanced Feistel" means that the round function takes in $n - 1$ bits and outputs a single bit—a "source-heavy" scheme. See the right-hand side of Fig. 1. A moment's reflection will make clear that, if the round function F_K provides uniform random bits, independently selected for each round, then maximally unbalanced Feistel *is* the Thorp shuffle.

As it takes n rounds of maximally unbalanced Feistel until each bit gets its turn in being replaced, we term n rounds of maximally unbalanced Feistel (or $\lceil \lg N \rceil$ rounds of Thorp) a *pass*. One might hope that the Thorp shuffle mixes the deck well after a small number of passes.

1.4. *Our Results*

Assume $N = 2^n$ is a power of two, $r \geq 1$, and let $E = \mathrm{Th}[N, R]$ be the Thorp shuffle with $R = 2rn$ rounds (that is, with $2r$ passes over the input). We will show that an adversary mounting a nonadaptive chosen-plaintext attack and making q queries will have advantage that is at most $(q/(r + 1)) \cdot (4nq/N)^r$ at distinguishing E from a random permutation on n bits. We prove this bound by regarding the Thorp shuffle of a designated q out of N cards as a Markov chain and applying a coupling argument. This marks the first time that coupling has been used to prove security for a symmetric cryptographic primitive. Using a result of Maurer, Pietrzak, and Renner [23], we can then infer that $4rn$ rounds ($4r$ passes) are enough so that a q-query adversary making an adaptive chosen-ciphertext attack will have advantage at most $(2q/(r + 1)) \cdot (4nq/N)^r$ at distinguishing E from a random permutation and its inverse. Put in asymptotic terms, we are saying that a maximally unbalanced Feistel network is CCA secure to $2^{n(1-1/r)}$ queries using $O(r)$ passes (and a uniformly random function from $n - 1$ bits to 1 bit). This far exceeds what balanced Feistel delivers, providing a demonstrable separation between the information-theoretic security of balanced and unbalanced Feistel.

While our result is strong compared to what was known before, we emphasize that there remains a large gap between our security result and the best attack known. We cannot rule out the possibility that with a number of passes r that is a small constant—perhaps as small as two—one already achieves information-theoretic security up to $N = 2^n$ queries (or at least up to cN queries). Narrowing this gap is an intriguing open problem.

1.5. *Practicality*

While our work has been motivated by a practical problem, in the end, the number of rounds we require to get a good security bound is not competitive with what one can do, heuristically, using a balanced Feistel network. Because of this, when people actually want to encipher strings over a small domain, they opt for a maximally balanced Feistel network. This is what NIST elected in their 2016 standard, and what we had suggested to them [5,6,11]. It is also what the arbitrary-input-length blockcipher AEZ elected to do [14]. While the choice of balanced Feistel precludes information-theoretic security beyond $\sqrt{N}$ queries, no computationally practical attacks are known, even if one asks up to $N-2$ queries, as long as one is careful in selecting an adequate number of rounds for enciphering very short strings [3]. (The "minus 2" is needed because a permutation determined by a balanced Feistel network is always even [32, Theorem 6.1].)

1.6. *Related Work*

The conference version of our paper appeared in 2009 [30]. Here we describe some of what came before and after, and place our work in further context.

The problem of enciphering on a small or unusual domain is a special case of *format-preserving encryption*, a goal addressed (albeit incorrectly) in an old NBS cryptographic standard [29], described informally by Brightwell and Smith [8], named and popularized in industry by Spies [41], and formalized by Bellare et al. [4].

One could always solve the small domain encryption problem with a *de novo* construction, creating from scratch a confusion/diffusion primitive with an unusually rich domain. This was what Schroeppel did in his Hasty Pudding cipher [40], anticipating by several years a formulation of the general problem. It is also what Granboulan, Levieil, and Piret investigate [12] in their treatment of blockciphers with domains other than bit strings, and what is addressed in the subsequent work of Baignères et al. [2].

For balanced Feistel, the classical analysis by Luby and Rackoff [21] shows that three rounds provide CPA security (four rounds for CCA security) to nearly $2^{n/4}$ queries. This was improved by Maurer and Pietrzak [22], who showed that R rounds of balanced Feistel could withstand about $2^{n/2-1/R}$ queries (in the CCA setting). Patarin [33,35] went on to show that a constant number of rounds (six for CCA security) was already enough to withstand nearly $2^{n/2}$ queries. He suggested that enough rounds of maximally unbalanced Feistel ought to achieve security approaching 2^n queries [33, p. 527].

Naor and Reingold analyzed unbalanced Feistel constructions, showing that one pass over a maximally unbalanced Feistel network that operates on n bits remains secure to nearly $2^{n/2}$ queries [31]. In their treatment, pairwise independent permutations bracket the Feistel construction. Kaplan, Naor, and Reingold describe a method to reduce the number of bits needed to specify a permutation that will appear uniform against some number q of queries [17]. They do this by derandomizing a construction such as the Thorp shuffle.

Morris established that the mixing time for the Thorp shuffle—roughly, the number of steps until all $q = N$ cards are ordered nearly uniformly—is polylogarithmic: he showed it was $O(\lg^{44} N)$ [27]. This was later improved to $O(\lg^{19} N)$ [24] and then to $O(\lg^4 N)$ [25] (for arbitrary N) and $O(\lg^3 N)$ [26] (for N a power of 2). After our work demonstrated that the Thorp shuffle mixes $N^{1-\epsilon}$ of the cards in $O(\lg N)$ rounds, Czumaj and Vöcking show mixing of any cN cards ($c < 1$) in $O(\lg^2 N)$ rounds.

Granboulan and Pornin [13] describe a method to perfectly realize a random permutation using a shuffling procedure of Czumaj, Kanarek, Kutyłowski, and Loryś [9]. Their approach requires one to repeatedly sample from a hypergeometric distribution using parameters that are large and vary during the shuffle. In an implementation, they accomplish this with an arbitrary-precision floating-point package. In the end, about 10^9 machine cycles are used to encipher on a space of $N < 2^{32}$ points.

Building on the conference version of the current paper and its use of couplings [30], Hoang and Rogaway analyze a variety of Feistel variants, including alternating, unbalanced, and numerical Feistel networks [16]. The work includes a coupling-based analysis for the Thorp shuffle with the number of cards arbitrary, not just a power of two, extending what we establish here. Hoang, Morris, and Rogaway went on to describe the Swap-or-Not shuffle [15], an oblivious shuffle closely related to the Thorp shuffle and invented to provide better proven security—up to cN adversarial queries, for an arbitrary constant $c < 1$. Building on an idea of Ristenpart and Yilek [36], Morris and Rogaway then described the Sometimes-Recurse shuffle [28], which can permit all N points to be queried and has fast *expected* running time: it shuffles all N cards in $\log N$ expected rounds.

Following our work [16,30], coupling arguments have increasingly been used to analyze constructions in symmetric cryptography. A good and early example is Lampe, Patarin, and Seurin [18], who use a coupling argument to analyze the iterated Even–Mansour cipher.

2. Preliminaries

By a *cipher*, we mean a map $E: \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ where $\mathcal{K}$ and $\mathcal{M}$ are finite nonempty sets (the *key space* and the *domain*) and $E_K(\cdot) = E(K, \cdot)$ is a permutation on $\mathcal{M}$ for every $K \in \mathcal{K}$. Let $\mathcal{A}$ be an adversary, meaning an algorithm with access to an oracle. For the game used to define E's indistinguishability from a random permutation, the oracle will depend on a permutation $f: \mathcal{M} \to \mathcal{M}$. It will respond to a query (enc, x) with $f(x)$ and it will respond to a query (dec, y) with $f^{-1}(y)$. Queries outside of $\{\mathsf{enc}, \mathsf{dec}\} \times \mathcal{M}$ are ignored. Define $\mathbf{Adv}_E^{\mathrm{cca}}(\mathcal{A}) = \mathbf{P}\,(K \twoheadleftarrow \mathcal{K}:\ \mathcal{A}^{\pm E_K} \Rightarrow 1) - \mathbf{P}\,(\pi \twoheadleftarrow \mathrm{Perm}(\mathcal{M}):\ \mathcal{A}^{\pm\pi} \Rightarrow 1)$ where $\mathcal{A}^{\pm f}$ denotes $\mathcal{A}$ interacting with the f-dependent oracle just described and $\mathcal{A}^f \Rightarrow 1$ is the event that it outputs a 1.

We say that adversary $\mathcal{A}$ is *nonadaptive* if its queries are the same on each and every run. It carries out a *chosen-plaintext* attack if each query is an encryption query, and a *chosen-ciphertext* attack if queries may be either encryption or decryption queries. Let $\mathbf{Adv}_E^{\mathrm{ncpa}}(q) = \max_{\mathcal{A}} \mathbf{Adv}_E^{\mathrm{cca}}(\mathcal{A})$ where the maximum is taken over all nonadaptive adversaries that ask at most q encryption queries and no decryption queries. By the standard averaging argument, the notion is unchanged if nonadaptive adversaries are assumed to be *deterministic*: they statically choose their queries $x_1, \ldots, x_q$. Let $\mathbf{Adv}_E^{\mathrm{cca}}(q) = \max_{\mathcal{A}} \mathbf{Adv}_E^{\mathrm{cca}}(\mathcal{A})$ where the maximum is taken over all adversaries that ask at most q queries.

3. Markov Chains and Coupling

We now give a brief description of the coupling technique for proving rapid mixing of a Markov chain; see any text on the subject, such as Levin, Peres, and Wilmer [19], for a more comprehensive account.

Let Ω be a finite nonempty set and let μ, ν be probability distributions on Ω. Let

$$\|\mu - \nu\| = \max_{S \subset \Omega} \mu(S) - \nu(S) = \tfrac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| \tag{1}$$

be the total variation distance between μ and ν.

A *coupling* of μ and ν is a pair of random variables (X, Y), defined on the same probability space, such that the marginal distributions of X and Y are μ and ν, respectively. The total variation distance has the following alternative description using couplings:

$$\|\mu - \nu\| = \min_{X \sim \mu,\ Y \sim \nu} \mathbf{P}(X \neq Y), \tag{2}$$

where $Z \sim \tau$ means that Z has distribution τ. The minimum in the right-hand side of (2) is taken over all couplings (X, Y) of μ and ν.

3.1. *The Coupling Technique*

Consider a Markov chain on state space Ω with transition matrix P and stationary distribution π. For a probability distribution μ on Ω, we write μ_t for the distribution of the chain at time t when the initial distribution is μ. Suppose we wish to bound the total variation distance $\|\mu_t - \nu_t\|$ for two different starting distributions μ and ν. (Typically, one wants to bound the distance $\|\mu_t - \pi\|$ for some starting distribution μ; this corresponds to letting $\nu = \pi$, so that $\nu_t = \pi$ for all t.) To use the coupling technique, one constructs a pair process $\{(X_t, Y_t) : t \geq 0\}$, the *coupling*, that satisfies the following conditions:

1. We have $X_0 \sim \mu$ and $Y_0 \sim \nu$.
2. Individually, $\{X_t\}$ and $\{Y_t\}$ are Markov chains with transition matrix P.
3. For every $t \geq 0$, if $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$.

The random variable $T = \min\{t : X_t = Y_t\}$ is called the *coupling time.* Then, from (2), we have

$$\begin{aligned} \|\mu_t - \nu_t\| &\leq \mathbf{P}(X_t \neq Y_t) \\ &= \mathbf{P}(T > t). \end{aligned}$$

The idea is to define the coupling in such a way that T is unlikely to be large.

4. The Projected Thorp Shuffle

Fix $N = 2^n$. Let $\{\mathrm{Th}_t : t \geq 0\}$ be the Markov chain representing the Thorp shuffle with N cards. More formally, let $\mathcal{C}$ be a set of cardinality N, whose elements we call *cards*.

For concreteness, we take $\mathcal{C} = \{0, 1\}^n$. The state space of $\{\mathrm{Th}_t\}$ is the set of bijections from $\mathcal{C}$ to $\{0, 1\}^n$. For a card $z \in \mathcal{C}$, we interpret $\mathrm{Th}_t(z)$ as the position of card z at time t. For cards $z_1, \dots, z_j$, we define $\mathrm{Th}_t(z_1, \dots, z_j) = (\mathrm{Th}_t(z_1), \dots, \mathrm{Th}_t(z_j))$.

Let $\mathcal{A}$ be a deterministic adversary that makes exactly q queries. Our proof is based on an analysis of the mixing rate of the Thorp shuffle. However, since $\mathcal{A}$ makes only q queries, we need only bound the rate at which a subset of the cards mixes. Fix m with $1 \leq m \leq N$, let $z_1, \dots, z_m$ be distinct cards in $\mathcal{C}$, and let X_t be the vector of positions of cards $z_1, \dots, z_m$ at time t. For j in $\{1, \dots, m\}$, we write *particle* j for card z_j, we write $X_t(j)$ for the position of particle j at time t, and define $X_t(1, \dots, j) = (X_t(1), \dots, X_t(j))$.

We shall call X_t the *projected Thorp shuffle with m particles*. Note that since the Thorp shuffle is a random walk on a group (see, e.g., [38]), it has uniform stationary distribution. Hence the stationary distribution of X_t, which we denote by π, is uniform over the set of distinct m tuples of elements from $\{0, 1\}^n$. Equivalently, π is the distribution of m samples without replacement from $\{0, 1\}^n$. Let τ_t denote the distribution of X_t.

Let X_t be the projected Thorp shuffle. It will be convenient to use a rule for generating the evolution of X_t that uses m fair coins, $c^1, \dots, c^m$, each of which is flipped at each step. Formally, each c^j is a sequence $\{c^j_t : t \geq 0\}$ of Bernoulli$(1/2)$ random variables, where we interpret c^j_t as the outcome of coin c^j at time t. We assume that all of the c^j_t are independent.

We now give the rule for updating the positions of each of the m particles in the projected Thorp shuffle. Say that two cards are *adjacent* at time t if their positions (viewed as elements of $\{0, 1\}^n$) are the same except for the first bit (or, viewed as elements of $\{0, \dots, N-1\}$, they differ by $N/2$). Note that if two particles are adjacent, then it is sufficient to describe how the position of one of them is updated, since this implies how the position of the other is updated. We shall use the following rule, which we call Rule R:

> **Update rule R** :
> If particle i is adjacent to particle j with $j < i$, then we determine the position of particle i at time $t+1$ using coin c^j and coin flip c^j_t as follows:
>
> 1. the first (leftmost) bit of the position of particle i is set to $1 - c^j_t$, and then
> 2. the position of particle i undergoes a cyclic left bit shift.
>
> If, on the other hand, particle i is not adjacent to a particle of smaller label, then we determine the position of particle i at time $t+1$ using coin c^i and coin flip c^i_t as follows:
>
> 1. the first (leftmost) bit of the position of particle i is set to c^i_t, and then
> 2. the position of particle i undergoes a cyclic left bit shift.

Note that when two particles are adjacent, the way that they are updated is entirely determined by the coin of the particle with the smaller label.

A key property of the projected Thorp shuffle is that when particles are not adjacent they move independently. The following lemma shows that particles are very unlikely to be adjacent after sufficiently many steps of the chain.

Lemma 1. *If $t \geq n-1$ then for any pair of particles i and j we have*

$$\mathbf{P}\,(i \text{ and } j \text{ are adjacent at time } t) \leq 2^{1-n}. \tag{3}$$

Proof. Assume that update rule R is used to generate X_t. Note that (by reordering if necessary) we may assume that $i = 1$ and $j = 2$. Let E be the event that particles 1 and 2 are adjacent at time t. If E occurs, then at each step during times $t-1, \ldots, t-n+1$, when their bits are changed (in step 1 of the update rule), the same change must occur for both particles. This occurs only if coins c^1 and c^2 have the same outcomes at times $t-1, \ldots, t-n+1$, which occurs with probability 2^{1-n}. It follows that $\mathbf{P}\,(E) \leq 2^{1-n}$. □

Theorem 1. (Rapid mixing) *Let $N = 2^n$, suppose $q \in \{1, \ldots, N\}$, and let $\{X_t : t \geq 0\}$ be the projected Thorp shuffle with q particles, π its stationary distribution, and τ_t the distribution of X_t. Then, for any $r \geq 1$,*

$$\|\tau_{r(2n-1)} - \pi\| \leq \frac{q}{r+1}\left(\frac{4nq}{N}\right)^r.$$

Proof. We use a hybrid argument. Fix cards $z_1, \ldots, z_q$. For each $l \leq q$, consider the vector $(Z_1, \ldots, Z_q)$ such that $Z_i = z_i$ if $i \leq l$ and Z_i is chosen uniformly from $\{0,1\}^n \backslash \{Z_1, \ldots, Z_{i-1}\}$ otherwise, and let μ^l be the distribution of the location of cards $Z_1, \ldots, Z_q$ after $r(2n-1)$ Thorp shuffles. Note that μ^q is $\tau_{r(2n-1)}$ and μ^0 is π.

The distributions μ^l and μ^{l+1} have the following property: The conditional distribution of the last $q-l-1$ particles, given the positions of the first $l+1$ particles, is uniform over all the possibilities. Thus, if we write μ^l_{l+1} (respectively, μ^{l+1}_{l+1}) for the marginal distribution, with respect to μ^l (respectively, μ^{l+1}), of the first $l+1$ particles, then

$$\mu^l(x_1, \ldots, x_q) = \mu^l_{l+1}(x_1, \ldots, x_{l+1})/A; \quad \mu^{l+1}(x_1, \ldots, x_q) = \mu^{l+1}_{l+1}(x_1, \ldots, x_{l+1})/A,$$

for some normalizing constant A. Hence

$$\begin{aligned}
\|\mu^l - \mu^{l+1}\| &= \tfrac{1}{2} \sum_{x_1, \ldots x_q} |\mu^l(x_1, \ldots, x_q) - \mu^{l+1}(x_1, \ldots, x_q)| \\
&= \tfrac{1}{2} \sum_{x_1, \ldots x_q} A^{-1} |\mu^l_{l+1}(x_1, \ldots, x_{l+1}) - \mu^{l+1}_{l+1}(x_1, \ldots, x_{l+1})| \\
&= \tfrac{1}{2} \sum_{x_1, \ldots x_{l+1}} |\mu^l_{l+1}(x_1, \ldots, x_{l+1}) - \mu^{l+1}_{l+1}(x_1, \ldots, x_{l+1})|. \qquad (4)
\end{aligned}$$

Thus, the total variation distance between μ^l and μ^{l+1} depends only on the marginal distributions of the first $l+1$ particles. If we write random variables as shorthand for their distributions, then (4) implies that

$$\|\mu^l - \mu^{l+1}\| = \|\text{Th}_{r(2n-1)}(z_1, \ldots, z_{l+1}) - \text{Th}_{r(2n-1)}(z_1, \ldots, Z_{l+1})\|. \tag{5}$$

To bound this total variation distance, we will use coupling. We will construct a pair process $\{(W_t, \tilde{W}_t) : t \geq 0\}$ such that W_t (resp. $\tilde{W}_t$) has the distribution of cards $z_1, \ldots, z_{l+1}$ (resp. cards $z_1, \ldots, z_l, Z_{l+1}$) at time t, in the Thorp shuffle. Let

$$W_0 = (\mathrm{Th}_0(z_1), \ldots, \mathrm{Th}_0(z_{l+1})) \qquad \tilde{W}_0 = (\mathrm{Th}_0(z_1), \ldots, \mathrm{Th}_0(z_l), \mathrm{Th}_0(Z_{l+1})).$$

It remains to describe the rule for generating $(W_{t+1}, \tilde{W}_{t+1})$ from $(W_t, \tilde{W}_t)$. Note that both W_t and $\tilde{W}_t$ are projected Thorp shuffles with $l+1$ particles. We shall generate the evolution of $\{(W_t, \tilde{W}_t) : t \geq 0\}$ using rule R with the same coins $c^1, \ldots, c^{l+1}$ updating both W_t and $\tilde{W}_t$. Note that under rule R, the updates for the first l particles are determined by coins $c^1, \ldots, c^l$ alone (since coin c^{l+1} is only used to update particle $l+1$ when it is not adjacent to a particle in $\{1, \ldots, l\}$.) Since the positions of the first l particles initially agree in both W_0 and $\tilde{W}_0$ (that is, $W_0(j) = \tilde{W}_0(j)$ for all $j \leq l$), and we are using the same coin flips $c_t^1, \ldots, c_t^l$ to update them each step, the positions of these particles remain matched for all times t. Furthermore, if at any point the position of particle $l+1$ becomes matched, then at that point all of particles $1, \ldots, l+1$ are matched and hence will stay matched from that point on. Thus, we have

$$\begin{aligned} \|W_t - \tilde{W}_t\| &\leq \mathbf{P}\left(W_t(l+1) \neq \tilde{W}_t(l+1)\right) \\ &= \mathbf{P}\left(T > t\right), \end{aligned} \tag{6}$$

where $T = \min\{t : W_t(l+1) = \tilde{W}_t(l+1)\}$ is the coupling time.

Let A be the event that at some time in $\{n-1, n, \ldots, 2n-2\}$, particle $l+1$ is adjacent to particle j for some $j \leq l$, in either the W process, or the $\tilde{W}$ process. Unless A occurs, coupling occurs by time $2n-1$. Summing equation (3) over 2 processes, n time steps, and l smaller indices shows that

$$\mathbf{P}\left(A\right) \leq 2nl \cdot 2^{1-n}. \tag{7}$$

It follows that

$$\mathbf{P}\left(T > 2n-1\right) \leq p\,, \tag{8}$$

where $p = nl2^{2-n}$. Note that (8) holds regardless of the initial state $(W_0, \tilde{W}_0)$ and that the process $\{(W_t, \tilde{W}_t) : t \geq 0\}$ is itself a Markov chain. Now imagine that we have a sequence of trials where in each trial we run the coupling for an additional $2n-1$ steps. The probability that particle $l+1$ remains unmatched after the first trial is at most p. Furthermore, by the memoryless property of Markov chains, given that it remained unmatched after the first $r-1$ trials, the conditional probability that it remains unmatched after the r-th trial is again at most p. Hence, by induction, $\mathbf{P}$ (particle $l+1$ remains unmatched after r trials) $\leq p^r = (nl2^{2-n})^r$, that is,

$$\mathbf{P}\left(T > r(2n-1)\right) \leq \left(nl2^{2-n}\right)^r. \tag{9}$$

Combining this with (6) and (5) shows that $\|\mu^l - \mu^{l+1}\|$ is at most $(nl2^{2-n})^r$. This holds for every l with $l \leq q$. Since μ^q is $\tau_{r(2n-1)}$ and μ^0 is π, we have, by the triangle inequality,

$$\begin{aligned}\|\tau_{r(2n-1)} - \pi\| \quad &\leq \quad \sum_{l=0}^{q-1} (nl2^{2-n})^r \leq \int_0^q (n2^{2-n})^r x^r \, \mathrm{d}x \\ &\leq \frac{q^{r+1}}{r+1} \cdot n^r 2^{2r-nr} = \frac{q}{r+1} \left(\frac{4nq}{N}\right)^r .\end{aligned}$$

□

5. Pseudorandomness of the Thorp Shuffle

5.1. *CPA Security*

The total variation distance is identical to the advantage with respect to a (deterministic) nonadaptive chosen-plaintext attack. So, reformulating Theorem 1 in cryptographic terms, what we have shown is the following.

Theorem 2. (nCPA security, concrete) *Let* $N = 2^n$ *and* $1 \leq q \leq N$. *Then, for any* $r \geq 1$,

$$\mathbf{Adv}^{\mathrm{ncpa}}_{\mathrm{Th}[N, r(2n-1)]}(q) \leq \frac{q}{r+1} \left(\frac{4nq}{N}\right)^r .$$

5.2. *Time-Reverse Thorp Shuffle*

Let $\mathrm{Th}^{-1}[N, R] = (\mathrm{Th}[N, R])^{-1}$ denote the time-reverse Thorp shuffle on N cards with R rounds: in round $r \in \{1, \ldots, R\}$ it sends cards $2x$ and $2x + 1$, where $0 \leq x < N/2$, either to x and $x + N/2$, or to $x + N/2$ and x, depending on a random bit $c(x, r)$. The inverse of the time-reverse Thorp shuffle is the Thorp shuffle itself. For N a power of two the forward and reverse Thorp shuffle are "isomorphic" Markov chains in the sense that there is a relabeling of states from $\mathrm{Th}[N, R]$ to $\mathrm{Th}^{-1}[N, R]$ that preserves the transition rule. As a consequence, the bound of Theorem 1 applies to the time-reverse Thorp shuffle as well as to the original. The observation is needed for concluding the theorem below.

5.3. *CCA Security*

A lovely result of Maurer, Pietrzak, and Renner [23] allows us to extend Theorem 2 to a larger class of adversaries—namely, we can trade our nCPA adversaries for CCA ones. The cost of doing so is just a doubling in the number of rounds, as well as the advantage bound.

Theorem 3. (CCAsecurity, concrete) *Let* $N = 2^n$ *and* $1 \leq q \leq N$. *Then, for any* $r \geq 1$,

$$\mathbf{Adv}^{\mathrm{cca}}_{\mathrm{Th}[N,r(4n-2)]}(q) \le \frac{2q}{r+1}\left(\frac{4nq}{N}\right)^{r}.$$

Proof. We use the second half of Corollary 5 from Maurer et al. [23, p. 148]. Although the notation of that paper is very different from what we use here, the authors' result implies that for any n-bit blockciphers F and G, we have that $\mathbf{Adv}^{\mathrm{cca}}_{H}(q) \le \mathbf{Adv}^{\mathrm{ncpa}}_{F}(q) + \mathbf{Adv}^{\mathrm{ncpa}}_{G}(q)$, where H is the cipher defined by $H_{K,K'}(X) = G^{-1}_{K'}(F_K(x)))$. Note that F and G are independently keyed. We use this fact with F being the $(R/2)$-round Thorp shuffle, $F = \mathrm{Th}[N, R/2]$, and G being the $(R/2)$-round time-reverse Thorp shuffle, $G = \mathrm{Th}^{-1}[N, R/2]$. Applying Theorem 2 and the observation that the time-reverse Thorp shuffle coincides with the usual Thorp shuffle gives the result. □

5.4. *Graphical Illustration*

The bounds of Theorems 2 and 3 are illustrated in Fig. 2. For example, for 16 passes and $N = 2^{40}$ points (third curve on the bottom right), an adversary must ask at least $2^{26.2}$ queries to have CCA advantage 0.5. For comparison, when applied to a maximally unbalanced Feistel network, the earlier analysis of Naor and Reingold [31, Theorem 6.2] would have topped out—with one pass—at $2^{16.8}$ queries. Had we enciphered strings using a balanced Feistel network instead, then the result of Maurer and Pietrzak [22, Theorem 1] would give a family of curves (depending, like ours, on how many rounds were performed) that would top out by $2^{18.5}$ queries. Patarin's result for six-round Feistel [35] would apparently be similar, but the concrete security is not explicitly given in that work, and the quantitative bounds are difficult to extract.

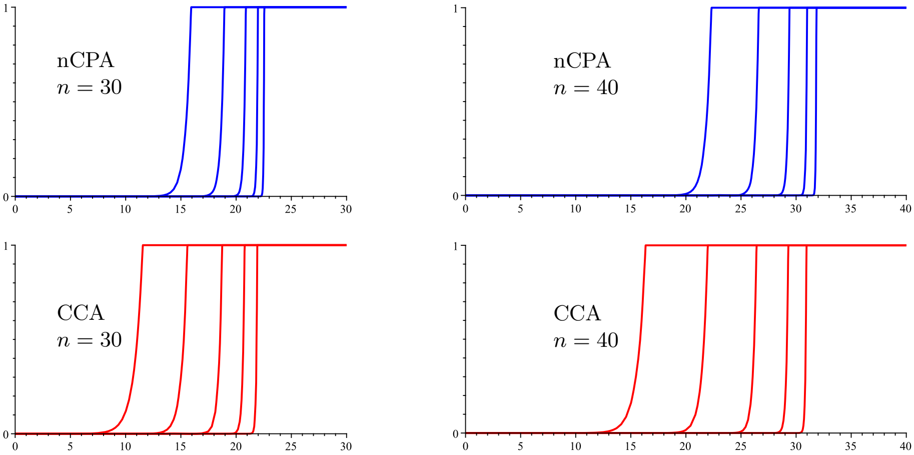

Fig. 2. Proven security of the Thorp shuffle. Each x-axis gives the log (base 2) of the number of queries. Each y-axis gives the upper bound on an adversary's nCPA advantage by Theorem 2 (*top*) or its CCA advantage by Theorem 3 (*bottom*). The curves in each plot are for $N = 2^{30}$ points (*left*) or $N = 2^{40}$ points (*right*). The five curves in each plot are for 4, 8, 16, 32, and 64 passes.

5.5. *Asymptotic Interpretation*

For an asymptotic interpretation of Theorem 2, fix $r > 0$ and suppose that $q = N^{1-1/r}$ and where, as before, $N = 2^n$. Then

$$\mathbf{Adv}^{\mathrm{ncpa}}_{\mathrm{Th}[N,2rn]}(q) \le \frac{q}{r+1}\left(\frac{4nq}{N}\right)^r = \frac{4^r n^r}{r+1} \cdot \frac{1}{N^{1/r}}.$$

In other words, we have upper-bounded the advantage by an expression of the form $(a\ \log^b N)/N^{1/r}$ for r-dependent constants a and b. Since this goes to 0 as $n \to \infty$, we conclude the following.

Corollary 4. (nCPA-security, asymptotic) *Let $r \ge 1$ be an integer. Then*

$$\lim_{n\to\infty} \mathbf{Adv}^{\mathrm{ncpa}}_{\mathrm{Th}[2^n,2rn]}\left(2^{n(1-1/r)}\right) = 0\,.$$

In English, a maximally unbalanced Feistel network on n bits employing $2r$ passes maintains security to nearly 2^n queries: specifically, to $2^{n(1-1/r)}$ queries for large enough n. Said differently, you can achieve security up to $N^{1-\varepsilon}$ nonadaptive queries, for any $\varepsilon > 0$, provided you make at least $2 \cdot \lceil 1/\varepsilon \rceil$ passes. This is far better than what a balanced Feistel network can achieve. The asymptotic version of Theorem 3 is similar.

Corollary 5. (CCA security, asymptotic) *Let $r \ge 1$ be an integer. Then*

$$\lim_{n\to\infty} \mathbf{Adv}^{\mathrm{cca}}_{\mathrm{Th}[2^n,4rn]}\left(2^{n(1-1/r)}\right) = 0\,.$$

5.6. *Designated-Point Security*

The PRP notion of security formalizes an adversary's inability to detect nonuniform behavior when it sees a *population* of plaintext/ciphertext pairs. Many security notions instead demand that the adversary figure something out about a designated point that it selects: the customary formulations for find-then-guess security, semantic security, unforgeability, and nonmalleability are all this way. Weakening the security notion along these lines facilitates a stronger bound for the Thorp shuffle.

Let $E\colon \mathcal{K} \times \mathcal{M} \to \mathcal{M}$ be a cipher and let $\mathcal{A}$ be an adversary. We measure the effectiveness of $\mathcal{A}$ in carrying out what we call a (nonadaptive) *designated-point attack* on E by $\mathbf{Adv}^{\mathrm{dpa}}_E(\mathcal{A}) = \mathbf{P}\left(\mathcal{A}^G \Rightarrow 1\right) - \mathbf{P}\left(\mathcal{A}^H \Rightarrow 1\right)$ where oracles G and H behave like this. Both oracles begin by sampling $K \twoheadleftarrow \mathcal{K}$ and then answering queries (enc, x) by $E_K(x)$. Oracle G answers the same way for a query (test, x), but H answers such a query by a uniformly chosen value that has not yet been returned to $\mathcal{A}$. No other types of queries are allowed. The adversary may ask a single test query, its last: once a test query is asked, any subsequent query returns $\perp$. Let $\mathbf{Adv}^{\mathrm{dpa}}_E(q) = \max_{\mathcal{A}} \mathbf{Adv}^{\mathrm{dpa}}_E(\mathcal{A})$ where the maximum is taken over all deterministic, nonadaptive adversaries that ask exactly q enc queries. The DPA notion is similar to but weaker than the IUP notion investigated by Desai and Miner [10], the main difference being that, with IUP security, the value x may depend on prior oracle responses.

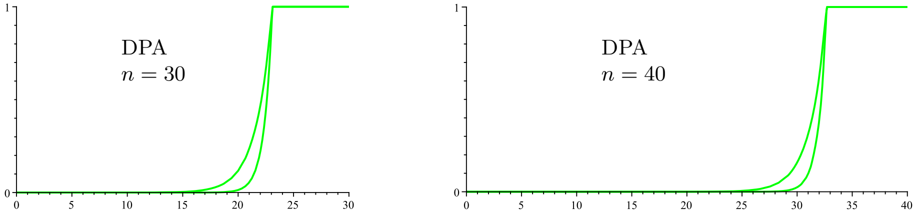

Fig. 3. Proven security of the Thorp shuffle, continued. The x-axis gives the log (base 2) of the number of queries. The y-axis gives an upper bound on an adversary's (nonadaptive) DPA advantage by Theorem 6, both for $N = 2^{30}$ points (*left*) and $N = 2^{40}$ points (*right*). The two curves in each figure are for two passes and then four.

Theorem 6. (Designated-point security) *Let $N = 2^n$ and $1 \le q \le N$. Then, for any $r \ge 1$,*

$$\mathbf{Adv}^{\mathrm{dpa}}_{\mathrm{Th}[N,r(2n-1)]}(q) \le \left(\frac{4nq}{N}\right)^r.$$

The proof follows immediately from Eqs. (6) and (9). The bounds are illustrated in Fig. 3. An asymptotic counterpart for the result is as follows.

Corollary 7. (Designated-point security, asymptotic) *For any $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbf{Adv}^{\mathrm{dpa}}_{\mathrm{Th}[2^n,2n-2)]}\left(2^{n(1-\varepsilon)}\right) = 0.$$

5.7. *More General Message Spaces*

We emphasize that our results on the Thorp shuffle have assumed that the size of the message is a power of two. By using the cycle-walking construction [7], this suffices to encipher messages on any message space $\{0, \ldots, N-1\}$. But the cost of applying this domain transformation can be nearly as bad as an expected doubling in the encryption and decryption time. It would be more desirable for the results to directly apply to Thorp-enciphering for any even N. Precisely such a result has been achieved in the follow-on work by Hoang and Rogaway [16]. The bounds are little unchanged.

6. Practical Considerations

In this section, we sketch some practical consideration in implementing Thorp shuffle encryption. We keep this section short in view of the fact that the most practical choice of all is not to implement the Thorp shuffle in the first place, but to use a maximally balanced Feistel network of many fewer rounds. One does not get the sort of strong provable security guarantee that is the focus of this paper, but one saves about an order of magnitude in rounds. For this reason, maximally balanced Feistel was chosen for

NIST Recommendation 800-38G [11] and for the AEZ-tiny portion of the arbitrary-input-length blockcipher AEZ [14].

6.1. *Round Function*

A practical realization of $\mathrm{Th}[N, R]$ must associate with each pair $x, x + N/2$ and each round r a coin flip $c(x, r)$. The function $c(x, r)$ would normally be built directly from AES. The number of rounds would then corresponds precisely to the number of AES calls.

The approach is wasteful in the sense that one is using only one of the 128 output bits from AES. With care, the 128 bits of AES output can be used to compute up to five rounds of the Thorp shuffle at once. The trick can be easily rediscovered and was described in the proceedings version of our paper [30]. We note that many modern processors now compute AES extremely quickly, making it unclear whether it is worth the conceptual and algorithmic complexity manipulate the 128-bit AES output to save on AES calculations.

Figure 4 illustrates our security bounds for Thorp shuffle enciphering messages of 20, 30, and 40 bits. A practical but fairly large number of rounds, typically hundreds, are needed to get good guarantees. Even then, guarantees do not reach to $q = N/(4 \lg N)$ queries. For comparison, FF1 [11] uses 10 rounds to encipher a 40-bit string, while AEZ [14] uses 16 rounds, both with anticipated computational security to $q = N - 2$ queries.

6.2. *Tweaking*

A practical realization for small-space encryption should be *tweakable*, a notion formalized by Liskov, Rivest, and Wagner [20]. The syntax of the cipher is extended to take an additional argument, the tweak, and each tweak effectively names a random independent cipher. A PRF-based scheme for small-space encryption can be modified to accommodate a tweak by including it in the scope of the PRF. As a simple example of the utility of doing so, an application might encipher the middle digits of a US social security number using a tweak that is the remaining digits.

	$n = 20$				$n = 30$				$n = 40$			
#passes	#rounds	dpa	ncpa	cca	#rounds	dpa	ncpa	cca	#rounds	dpa	ncpa	cca
2	40	12.7	6.9	—	60	22.1	11.6	—	80	31.7	16.4	—
4	80	13.2	9.4	6.4	120	22.6	15.7	11.2	160	32.2	22.1	15.9
8	160	13.4	11.3	9.1	240	22.8	18.8	15.3	320	32.4	26.5	21.7
16	320	13.6	12.4	11.1	480	23.0	20.8	18.6	640	32.6	29.3	26.3
32	640	13.6	13.4	13.0	960	23.1	22.5	20.0	1280	32.6	31.8	30.9
64	1280	13.6	13.4	13.0	1920	23.1	22.6	21.9	2560	32.6	31.8	30.9

Fig. 4. Concrete security of the Thorp shuffle. The columns indicate the domain size $N = 2^n$; the number of passes r; the number of rounds R (or the number of AES calls, using a naive implementation); and the log (base 2) of the number of queries q that can be tolerated until our bound on $\mathbf{Adv}^{\mathrm{xxx}}_{\mathrm{Th}[2^n, rn]}(q)$ is about 0.5, for xxx $\in$ {dpa, ncpa, cca}. With substantially fewer queries than this, the advantage is very small.

6.3. *Variable Input Length*

It may be desirable that a realization of the Thorp shuffle achieve security in the sense of a variable-input-length (VIL) cipher; the size of the message space may vary from query to query in an adversarially selected way. This is easily achieved by including the domain size N within the scope of the PRF that is used to realize the shuffle. Use of a VIL enciphering scheme facilitates, for example, enciphering database fields that have various lengths.

Acknowledgements

Many thanks to Terence Spies and Voltage Security, for renewing the second author's interest in the small-space encryption problem. Many thanks to the *JoC* referees, who provided excellent comments. This research was supported in part by NSF grants CNS 0904380, CNS 1228828, CNS 1314885, and CNS 1717542. Many thanks to the NSF for their support. Till Stegers is now employed at Google Inc., in Los Angeles.

References

[1] W. Aiello, R. Venkatesan, Foiling birthday attacks in length-doubling transformations: Benes: a non-reversible alternative to Feistel, in *EUROCRYPT 1996*, LNCS, vol. 1070 (Springer, 1996), pp. 307–320

[2] T. Baignères, J. Stern, S. Vaudenay, Linear cryptanalysis of non binary ciphers (with an application to SAFER), in *Selected Areas in Cryptography 2007*, LNCS vol. 4876 (Springer, 2007), pp. 184–211

[3] M. Bellare, V.T. Hoang, S. Tessaro, Message-recovery attacks on Feistel-based format preserving encryption, in *ACM Conference on Computer and Communications Security (CCS 2016)* (ACM Press, 2016), pp. 444–455

[4] M. Bellare, T. Ristenpart, P. Rogaway, T. Stegers, Format-preserving encryption, in *Selected Areas in Cryptography (SAC 2009)*, LNCS vol. 5867 (Springer, 2009), pp. 295–312

[5] M. Bellare, P. Rogaway, T. Spies, The FFX mode of operation for format-preserving encryption. NIST submission (2010). http://ccsrc.nist.gov/groups/ST/toolkit/BCM/modes_development.html

[6] M. Bellare, P. Rogaway, T. Spies, Addendum to the "The FFX mode of operation for format-preserving encryption", NIST submission (2010). http://csrc.nist.gov/groups/ST/toolkit/BCM/modes_development.html

[7] J. Black, P. Rogaway, Ciphers with arbitrary finite domains, in *Topics in Cryptology – CT-RSA 2002* LNCS, vol. 2271 (Springer 2002), pp. 114–130

[8] M. Brightwell, H. Smith, Using datatype-preserving encryption to enhance data warehouse security, in *20th National Information Systems Security Conference Proceedings (NISSC)* (1997) pp. 141–149

[9] A. Czumaj, P. Kanarek, M. Kutyłowski, K. Loryś, Fast generation of random permutations via networks simulation. *Algorithmica*: 21(1), (Springer, 1998)

[10] A. Desai, S. Miner, Concrete security characterizations of PRFs and PRPs: reductions and applications, in *ASIACRYPT 2000*. LNCS, vol. 1976 (Springer, 2000), pp. 503–516

[11] M. Dworkin, Recommendation for block cipher modes of operation: methods for format-preserving encryption. NIST Special Publication 800-38G (2016)

[12] L. Granboulan, E. Levieil, G. Piret, Pseudorandom permutation families over abelian groups, in *FSE 2006*. LNCS vol. 4047 (Springer, 2006), pp. 57–77

[13] L. Granboulan, T. Pornin, Perfect block ciphers with small blocks, in *Fast Software Encryption (FSE 2007)*. LNCS vol. 4593 (Springer, 2007), pp. 452–465

[14] V. T. Hoang, T. Krovetz, P. Rogaway, Robust authenticated-encryption: AEZ and the problem that it solves, in *EUROCRYPT 2016 (1)*. LNCS vol. 9056 (Springer, 2015), pp. 15–44

[15] V. T. Hoang, B. Morris, P. Rogaway, An enciphering scheme based on a card shuffle, in *CRYPTO 2012*. LNCS vol. 7417 (Springer, 2012), pp. 1–13

[16] V. T. Hoang, P. Rogaway, On generalized Feistel networks, in *CRYPTO 2010*. LNCS vol. 6223 (2010), pp. 613–630

[17] E. Kaplan, M. Naor, O. Reingold, Derandomized constructions of k-wise (almost) independent permutations, in *Randomization and Computation (RANDOM 2005)*. LNCS vol. 3624 (Springer, 2005), pp. 354–365

[18] R. Lampel, J. Patarin, Y. Seurin. An asymptotically tight security analysis of the iterated Even-Mansour cipher, in *ASIACRYPT 2012*. LNCS 7658 (Springer, 2012), pp. 278–295

[19] D. Levin, Y. Peres, E. Wilmer, *Markov chains and mixing times* (American Mathematical Society 2008)

[20] M. Liskov, R. Rivest, D. Wagner, Tweakable block ciphers, in *CRYPTO 2002*, LNCS vol. 2442 (Springer, 2002), pp. 31–46

[21] M. Luby, C. Rackoff, How to construct pseudorandom permutations from pseudorandom functions. *SIAM J. Comput.* **17**(2), 373–386 (1988)

[22] U. Maurer, K. Pietrzak, The security of many-round Luby-Rackoff pseudo-random permutations, in *EUROCRYPT 2003*. LNCS vol. 2656 (Springer, 2003), pp. 544–561

[23] U. Maurer, K. Pietrzak, R. Renner, Indistinguishability amplification, in *CRYPTO 2007*. LNCS vol. 4622 (Springer, 2007), pp. 130–149

[24] R. Montenegro, P. Tetali, Mathematical aspects of mixing times in Markov chains, in *Foundations and Trends in Theoretical Computer Science* 1(3) (Now Publishers, 2006)

[25] B. Morris, Improved mixing time bounds for the Thorp shuffle and L-reversal chain. *Ann. Probab*. **37**(2), 453–477 (2009)

[26] B. Morris, Improved mixing time bounds for the Thorp shuffle. *Comb. Probab. Comput*. **22**(1) (2013)

[27] B. Morris, The mixing time of the Thorp shuffle. SIAM *J. Comput.* **38**(2), 484–504 (2008) *Earlier version in STOC 2005*

[28] B. Morris, P. Rogaway, Sometimes-Recurse shuffle: almost-random permutations in logarithmic expected time, in *EUROCRYPT 2014*. LNCS vol. 8441 (Springer, 2014), pp. 311–326

[29] National Bureau of Standards. FIPS PUB 74: Guidelines for Implementing and Using the NBS Data Encryption Standard (1981)

[30] B. Morris, P. Rogaway, T. Stegers, How to encipher messages on a small domain: deterministic encryption and the Thorp shuffle, in *CRYPTO 2009*. LNCS vol. 2009 (Springer, 2009), pp. 286–302

[31] M. Naor, O. Reingold, On the construction of pseudo-random permutations: Luby-Rackoff revisited. *J. Cryptol.* **12**(1), 29–66 (1999)

[32] J. Patarin, Generic attacks on Feistel schemes. Cryptology ePrint report 2008/036

[33] J. Patarin, Luby-Rackoff: 7 rounds are enough for $2^{n(1-\varepsilon)}$ security, in *CRYPTO 2003*. LNCS vol. 2729 (Springer, 2003), pp. 513–529

[34] J. Patarin, A proof of security in $O(2^n)$ for the Benes scheme, in *Progress in Cryptology – AFRICACRYPT 2008*. LNCS vol. 5023 (Springer, 2008), pp. 209–220

[35] J. Patarin, Security of random Feistel schemes with 5 or more rounds, in *CRYPTO 2004*. LNCS vol. 3152 (Springer, 2004), pp. 106–122

[36] T. Ristenpart, S. Yilek, The Mix-and-Cut shuffle: small-domain encryption secure against N queries, in *CRYPTO 2013, Part I*. LNCS vol. 8042 (Springer, 2013), pp. 392–409

[37] S. Rudich, Limits on the provable consequences of one-way functions. Ph.D. Thesis, UC Berkeley (1989)

[38] L. Saloff-Coste, Random walks on finite groups, in *Probability on Discrete Structures, Encyclopedia of Mathematical Sciences*, vol. 110, H. Kesten, editor (Springer 2004), pp. 263–346

[39] B. Schneier, J. Kelsey, Unbalanced Feistel networks and block-cipher design, in *Fast Software Encryption (FSE 1996)*. LNCS vol. 1039 (Springer, 1996) pp. 121–144

[40] R. Schroeppel, Hasty Pudding Cipher specification (1998). http://richard.schroeppel.name:8015/hpc/hpc-spec (revised 5/99)

[41] T. Spies, Personal communications (2009)

[42] E. Thorp, Nonrandom shuffling with applications to the game of Faro. *J. Am. Stat. Assoc.* **68**, 842–847 (1973)