About the Security of Ciphers (Semantic Security and Pseudo-Random Permutations)

Duong Hieu Phan and David Pointcheval

CNRS/ENS – Dépt d'informatique – 45 rue d'Ulm, 75230 Paris Cedex 05, France
{duong.hieu.phan, david.pointcheval}@ens.fr

Abstract. Probabilistic symmetric encryption have already been widely studied, from a theoretical point of view. Nevertheless, many applications require length-preserving encryption, to be patched at a minimal cost to include privacy without modifying the format (e.g. encrypted filesystems). In this paper, we thus consider the security notions for length-preserving, deterministic and symmetric encryption schemes, also termed *ciphers*: semantic security under lunchtime and challenge-adaptive adversaries. We furthermore provide some relations for this notion between different models of adversaries, and the more classical security notions for ciphers: pseudo-random permutations (PRP) and super pseudo-random permutations (SPRP).

1 Introduction

The main goal for any encryption scheme is secrecy: ideally, such a notion means that a ciphertext should not reveal any information about the plaintext, however powerful is the adversary. This had been defined under "perfect secrecy" [11], but also showed to be impossible, unless one uses one-time pad, which is a symmetric encryption that uses a secret key as long as the messages to be encrypted. That is, if one wants to use a small symmetric key in order to protect many plaintexts or a long message, or asymmetric encryption, such perfect secrecy is impossible.

To overcome this theoretical impossibility, but which has no real practical impact since adversaries are computationally limited, several security notions have thereafter been defined, and namely the polynomial security [4], *a.k.a.* indistinguishability of ciphertexts or semantic security. This intuitively means that no *polynomially* bounded adversary can extract any information about the plaintext, given the ciphertext.

However, in practice, an adversary is not only given the challenge ciphertext about which plaintext it wants to learn some information. It may also have access to extra information, such as plaintext-ciphertext pairs. According to the way these pairs are obtained, several kinds of attacks may be mounted: known pairs, chosen-plaintext or chosen-ciphertext attacks, in an adaptive way or not. Furthermore, when considering semantic security, the choice of the plaintexts or the ciphertexts may be allowed before the adversary has been given the challenge ciphertext (lunchtime attacks [8]), or unlimited (*challenge*-adaptive attacks [10]).

H. Handschuh and A. Hasan (Eds.): SAC 2004, LNCS 3357, pp. 182–197, 2005.

1.1 Some Wordings

In order to make things clear, let us note that all the adversaries considered in this paper are implicitly *adaptive*, in the sense that their queries to any oracle may depend on previous answers, but not necessarily on the challenge ciphertext they want to break (when such a specific challenge exists, as in the semantic security game, or the indistinguishability one). To make the distinction between whether the challenge ciphertext may impact the queries or not, we will use the terms "adaptive attacks" and "lunchtime attacks" respectively: in lunchtime attacks the adversary has a full and adaptive access to oracles but before the challenge ciphertext is known only, while in adaptive attacks this access is unlimited in time.

1.2 Motivation

Relations between various security notions for symmetric encryption, under different kinds of attacks, have been deeply studied by Bellare *et al.* [1] and Katz and Yung [6]. But they were mainly restricted to the probabilistic case. Nevertheless, many applications of encryption require length-preserving schemes. For compatibility, one may indeed want the message format to be similar, whatever it is in clear (no privacy) or encrypted (enhanced with privacy). Another famous application of encryption is for encrypted filesystems [5], which need encryption schemes able to encipher the sectors of a disk in-place, while sectors have a fixed length. Length-preserving symmetric encryption thus means *deterministic* encryption schemes. In the following we thus focus on length-preserving, deterministic and symmetric encryption schemes, also termed *ciphers*. However, from our knowledge, no analysis of *ciphers* has ever been done so far. The main reason may be that, while the security goal is privacy, no semantic security definition fits the deterministic case: it is clear that the straightforward extension of the usual notion fails when considering deterministic encryption (probabilistic encryption is a basic requirement for semantic security, when an oracle —encryption and decryption— is available at least once). As a consequence, other notions are used: pseudo-random permutation or super pseudo-random permutation properties [3, 7].

The security notion one usually requires from a block cipher is indeed to look like perfectly random permutations for random keys (family of pseudo-random permutations if one just considers chosen-plaintext attacks, or family of super pseudo-random permutations if decryption queries are also possible). This is a very strong security notion useful when the block cipher is seen as a all-purpose primitive (for providing stream ciphers with encryption modes, message authentication codes, etc.). But for confidentiality, the useful notion is secrecy only: the view of the ciphertext does not leak any useful information about the plaintext to a (polynomial) adversary. While the former notion of super pseudo-random permutations is clearly stronger than the latter, the actual relations have never been studied.

1.3 Previous Work

Security notions for encryption have been defined a long time ago, namely with the definition of polynomial security [4] (*a.k.a.* semantic security or indistinguishability). Bellare *et al.* [1] studied several variants of the latter, for symmetric encryption, under the names of *find-then-guess*, *left-or-right* and *real-or-random*, and relations in the concrete setting. Katz and Yung [6] studied the actual difference between these various kinds of attacks, against probabilistic symmetric encryption. Indeed, whereas in the public-key setting chosen-plaintext attack is the basic scenario for an adversary, since it can encrypt any plaintext of its choice granted the public key, in the symmetric setting, simply some known plaintext-ciphertext pairs may give extra information. However, they showed that an adaptive chosen-plaintext attack (where queries are allowed even after the challenge ciphertext is known) does not help more than a lunchtime attack (where oracle accesses are limited up to the reception of the challenge ciphertext.)

As already noted, the security notion usually required from a block cipher is the (super) pseudo-randomness, which means to look like *perfectly random permutations*, for randomly chosen keys. Depending on whether a decryption oracle is available or not, one indeed considers either the super pseudo-randomness or the pseudo-randomness only, respectively. The latter notion (the weakest) has been recently studied by Desai and Miner [2]. They claimed the equivalence between this notion and the semantic security under lunchtime chosen-plaintext attacks. Halevy and Rogaway [5] showed the equivalence between the super pseudo-randomness and the *left-or-right* indistinguishability, with (almost) unlimited oracle accesses, for tweakable ciphers.

1.4 Contributions

In this paper, we study the security notions of secrecy for ciphers, namely semantic security (indistinguishability of ciphertext) and (super) pseudo-randomness, with the existing relations between them.

We first show that the usual indistinguishability, modeled by the *find-then-guess* game, (with some natural restrictions) is still equivalent to the natural definition of semantic security (adapted for symmetric and deterministic encryption).

We then show that some results relative to the probabilistic case remain true for ciphers. Namely, adaptive chosen-plaintext attacks do not provide significant advantage against lunchtime attacks. More interestingly, we also consider the relation between adaptive and lunchtime chosen-ciphertext attacks, and prove that an adaptive access does not help either in the case where the cipher and its inverse are already both secure against lunchtime attacks.

Finally, for completeness, we provide relations between the above notions and the notion of (super) pseudo-random permutations. We namely prove that indistinguishability against lunchtime adversaries is equivalent to the notion of super pseudo-random permutations, when the cipher and its inverse have the

same security level against lunchtime attacks: challenge-adaptive security level is not necessary. All the proofs and some additional relations, under various assumptions, are provided in the full version [9].

We believe that these results have concrete applications for practical ciphers, since the encryption and the decryption algorithms are often very similar, and thus with a similar security level. For example, when considering DES possibly using some mode of operation, under the conjecture that a slight modification of the key schedule (replacement of the left rotation by a right rotation) does not affect the security against at least lunchtime adversaries, we can show that the above results hold without any additional assumption (see the full version [9] for the application.)

2 Security Notions for Encryption

2.1 Symmetric Encryption Schemes

Let us first review the formal definition of a symmetric encryption scheme $\pi = (k, \ell, \mathcal{E}, \mathcal{D})$. It is defined by two algorithms, parameterized by a key k that is assumed to be uniformly distributed in $\{0,1\}^k$. Note that the two main data in practice are k, the bit-length of the keys, and ℓ the bit-length of the block to be encrypted:

- the encryption algorithm $\mathcal{E}_\mathsf{k}$, which on a message m from the set $\{0,1\}^\ell$, and random coins r from $\{0,1\}^\mu$, outputs a ciphertext c in $\{0,1\}^\nu$;
- the decryption algorithm $\mathcal{D}_\mathsf{k}$, which on a ciphertext c outputs the corresponding plaintext m, or $\perp$ if there is no corresponding plaintext.

2.2 Ciphers: Length-Preserving, Deterministic and Symmetric Encryption Schemes

In the particular case of deterministic encryption, the encryption scheme does not use any random coin, since it is furthermore length-preserving, any ciphertext is valid: it is a permutation for each key (and thus $\mu = 0$ and $\nu = \ell$.) For a given cipher $\pi = (k, \ell, \mathcal{E}, \mathcal{D})$, we can denote the inverse cipher by:

$$\pi^{-1} = (k, \ell, \mathcal{E}^{-1} = \mathcal{D}, \mathcal{D}^{-1} = \mathcal{E}).$$

2.3 Semantic Security

The natural security notion for encryption is the computational variant of perfect secrecy: the view of the ciphertext does not help to learn any information about the plaintext. This has been formalized by the notion of *semantic security* [4], for which a SEM-adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ plays the following game, in two steps:

- a key k is first uniformly drawn from $\{0,1\}^k$;
- Stage 1: $\mathcal{A}_1$ outputs a samplable distribution D on the set $\{0,1\}^\ell$, together with a state information s to be forwarded to the second step of the attack;

- a message m is drawn from $\{0,1\}^\ell$ according to the distribution D (denoted $m \stackrel{D}{\leftarrow} \{0,1\}^\ell$), and a random tape r is uniformly drawn from $\{0,1\}^\mu$ (denoted $r \stackrel{R}{\leftarrow} \{0,1\}^\mu$) then one computes $c = \mathcal{E}_\mathsf{k}(m;r)$;
- Stage 2: $\mathcal{A}_2$ is given the state information s and the ciphertext c. It outputs a computable predicate f.

The adversary is said to be successful if $f(m)$ is true. It means that it has been able to learn at least one bit of information about m, from the ciphertext c. However it is easy for an adversary to win all the time, by outputting a constant predicate f. Then we say that $\mathcal{A}$ breaks the semantic security if the predicate f holds on m with probability significantly greater than for another random plaintext m' (following the same "a priori" distribution D).

Therefore, we define the advantage $\mathsf{Adv}^{\mathsf{sem}}_\pi(\mathcal{A})$ of an adversary $\mathcal{A}$, against the semantic security of an encryption scheme π, by $\Pr[f(m) = 1] - \Pr[f(m') = 1]$ on the distribution space $\mathcal{D} = \{\mathsf{k} \stackrel{R}{\leftarrow} \{0,1\}^k; (D,s) \leftarrow \mathcal{A}_1(); m, m' \stackrel{D}{\leftarrow} \{0,1\}^\ell; r \stackrel{R}{\leftarrow} \{0,1\}^\mu; c = \mathcal{E}_\mathsf{k}(m;r); f \leftarrow \mathcal{A}_2(s,c)\}$

Definition 1. *An encryption scheme π is said to be (ε, t)-semantically secure if for any adversary $\mathcal{A}$, that runs within time t, $\mathsf{Adv}^{\mathsf{sem}}_\pi(\mathcal{A}) \leq \varepsilon$.*

Adversaries. Adversary $\mathcal{A}$ may be given extra information than just the challenge ciphertext, such as plaintext-ciphertext pairs. According to the way these pairs are defined, several kinds of attacks may be mounted: known pairs, chosen-plaintext and/or chosen-ciphertext attacks. Furthermore, the choice of the plaintexts or the ciphertexts may be allowed before the adversary has been given the challenge ciphertext only, or unlimited.

Such additional information is modeled by (un)limited access to oracles that compute encryptions or decryptions. A (t, e_1, d_1, e_2, d_2)-adversary $\mathcal{A} = (\mathcal{A}_1^{\mathcal{E}_\mathsf{k}, \mathcal{D}_\mathsf{k}}, \mathcal{A}_2^{\mathcal{E}_\mathsf{k}, \mathcal{D}_\mathsf{k}})$ is a 2-stage adversary $\mathcal{A}$ where $\mathcal{A}_1$ (resp. $\mathcal{A}_2$) can ask up to e_1 and d_1 (resp. e_2 and d_2) queries to the encryption and decryption oracles $\mathcal{E}_\mathsf{k}$ and $\mathcal{D}_\mathsf{k}$, with a running time bounded by t. We cover this way the passive adversary, where $e_1 = e_2 = d_1 = d_2 = 0$ that is denoted P0-C0, or any active adversary that is denoted PX-CY, according to the oracles access:

$X =$ '1' – $e_1 > 0$ but $e_2 = 0$, lunchtime chosen-plaintext (P1-CY);
$Y =$ '1' – $d_1 > 0$ but $d_2 = 0$, lunchtime chosen-ciphertext (PX-C1);
$X =$ '2' – $e_2 > 0$ whatever e_1 is, adaptive chosen-plaintext (P2-CY);
$Y =$ '2' – $d_2 > 0$ whatever d_1 is, adaptive chosen-ciphertext (PX-C2).

We remind that all the adversaries are adaptive w.r.t. the previous oracle answers, and thus by "adaptive" we mean "challenge-adaptive", while "lunchtime" stands for "challenge-non-adaptive".

Such a PX-CY adversary can play the attack game against semantic security, but there are natural restrictions in case of oracle access. Let us denote by $\Lambda_\mathcal{E}$ ($\Lambda_\mathcal{D}$ resp.) the lists of plaintext-ciphertext (m,c) pairs obtained from the encryption oracle (and the decryption oracle resp.). The superscript m (resp.

c) will be used to restrict these lists to the first coordinates (resp. the second coordinates), which thus leads to two lists of plaintexts $\Lambda^m_{\mathcal{E}}$ and $\Lambda^m_{\mathcal{D}}$, and two lists of ciphertexts $\Lambda^c_{\mathcal{E}}$ and $\Lambda^c_{\mathcal{D}}$. The restrictions are thus:

- if the adversary has access to the decryption oracle (that is C1 or C2), it is restricted not to ask the challenge ciphertext c in the second stage;
- in the deterministic case, if the adversary has access to the encryption oracle (that is P1 or P2), the support S_D of D (the set of plaintexts that have a non-zero probability in D) must be disjoint with the list of the plaintexts asked to the encryption oracle at any time, or obtained from the decryption oracle during the first stage.

The former restriction is the classical one, and the latter one is quite natural for deterministic encryption. We show later (by proving equivalence with the *find-then-guess* notion) that it is a minimal restriction.

Definition 2. *An encryption scheme π is said to be $(\varepsilon, t, e_1, d_1, e_2, d_2)$-semantically secure if for any (t, e_1, d_1, e_2, d_2)-SEM adversary $\mathcal{A}$, that asks at most e_1 and d_1 (resp. e_2 and d_2) encryption and decryption queries in the first stage (resp. in the second stage) within time t, $\mathsf{Adv}^{\mathsf{sem}}_{\pi}(\mathcal{A}) \leq \varepsilon$.*

2.4 Indistinguishability: Find-Then-Guess

The *indistinguishability* security notion (also known as *find-then-guess* [1]) involves a (t, e_1, d_1, e_2, d_2)-IND adversary $\mathcal{A} = (\mathcal{A}_1^{\mathcal{E}_k,\mathcal{D}_k}, \mathcal{A}_2^{\mathcal{E}_k,\mathcal{D}_k})$ that plays the following game:

- a key k is first uniformly drawn from $\{0,1\}^k$;
- Stage 1 (find): $\mathcal{A}_1^{\mathcal{E}_k,\mathcal{D}_k}$ outputs two plaintexts (m_0, m_1) together with a state information s;
- a bit b is randomly drawn, and a random tape r is uniformly drawn from $\{0,1\}^\mu$ then one computes $c = \mathcal{E}_{\mathsf{k}}(m_b; r)$;
- Stage 2 (guess): $\mathcal{A}_2^{\mathcal{E}_k,\mathcal{D}_k}$ is given the state information s and the ciphertext c. It outputs its guess b' for b.

The adversary is said to be successful if $b' = b$. It means that it has been able to distinguish the encryption of m_0 from the encryption of m_1. However it is easy for an adversary to win half the time, by simply flipping a random coin. Then we say that $\mathcal{A}$ breaks the *find-then-guess* security if $b' = b$ with probability significantly greater than $1/2$. Therefore, we define the advantage of an adversary $\mathcal{A}$, against the *find-then-guess* security, or *indistinguishability*, of an encryption scheme π, by the following formula:

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}) = 2 \times \Pr\left[\begin{array}{l} \mathsf{k} \stackrel{R}{\leftarrow} \{0,1\}^k; (m_0, m_1, s) \leftarrow \mathcal{A}_1^{\mathcal{E}_k,\mathcal{D}_k}(); b \stackrel{R}{\leftarrow} \{0,1\}; \\ r \stackrel{R}{\leftarrow} \{0,1\}^\mu; c = \mathcal{E}_{\mathsf{k}}(m_b; r); b' \leftarrow \mathcal{A}_2^{\mathcal{E}_k,\mathcal{D}_k}(s, c) : b' = b \end{array}\right] - 1.$$

As above, there are also natural restrictions in case of oracle access:

– if the adversary has access to the decryption oracle (that is C1 or C2), it is restricted not to ask the challenge ciphertext c in the second stage;
– in the deterministic case, if the adversary has access to the encryption oracle (that is P1 or P2) it is restricted not to ask m_0 or m_1 to the encryption oracle at any time, or to have obtained m_0 or m_1 from the decryption oracle during the first stage.

Since we focus this paper on the deterministic case, one can note that the above restrictions sum up to

$$m_0, m_1 \notin \Lambda_{\mathcal{E}}^{m} \qquad c \notin \Lambda_{\mathcal{D}}^{c}.$$

Definition 3. *An encryption scheme π is said to be $(\varepsilon, t, e_1, d_1, e_2, d_2)$-indistinguishable if for any (t, e_1, d_1, e_2, d_2)-*IND *adversary $\mathcal{A}$, that asks at most e_1 and d_1 (resp. e_2 and d_2) encryption and decryption queries in the first stage,* a.k.a. *the find stage (resp. in the second stage,* a.k.a. *the guess stage) within time t,* $\mathsf{Adv}_{\pi}^{\mathsf{ind}}(\mathcal{A}) \leq \varepsilon$.

2.5 Pseudo-Random and Super Pseudo-Random Permutations

Pseudo-Random Permutation. The usual security notion one requires from a block cipher is to look like perfectly random permutations, for the keys uniformly drawn. This notion can be formalized as follows: any adversary accessing an oracle $\mathcal{O}_b$ ($\mathcal{O}_0$ corresponds to the perfectly random permutation $\mathcal{P}$ —a permutation randomly chosen in the set $\mathcal{SP}_\ell$ of the permutations onto $\{0,1\}^\ell$— and $\mathcal{O}_1$ corresponds to an encryption permutation $\mathcal{E}_\mathsf{k}$, for a random key k) cannot guess b (i.e, it cannot distinguish if it accesses the perfectly random permutation $\mathcal{P}$ or the actual encryption algorithm $\mathcal{E}_\mathsf{k}$, with a random key):

$$\mathsf{Adv}_{\pi}^{\mathsf{prp}}(\mathcal{A}) = 2 \times \Pr\left[\begin{array}{l}\mathsf{k} \stackrel{R}{\leftarrow} \{0,1\}^k; \mathcal{P} \stackrel{R}{\leftarrow} \mathcal{SP}_\ell; \mathcal{O}_0 = \mathcal{P}; \mathcal{O}_1 = \mathcal{E}_\mathsf{k}; \\ b \stackrel{R}{\leftarrow} \{0,1\}; b' \leftarrow \mathcal{A}^{\mathcal{O}_b}() : b' = b\end{array}\right] - 1.$$

Definition 4. *An encryption scheme π is said to be a (ε, t, n)-pseudo-random permutation, denoted (ε, t, n)-*PRP *if for any (t, n)-*PRP *adversary $\mathcal{A}$, that asks at most n encryption queries within time t,* $\mathsf{Adv}_{\pi}^{\mathsf{prp}}(\mathcal{A}) \leq \varepsilon$.

Super Pseudo-Random Permutation. The above notion does not take into account the decryption oracle access. Hence the stronger notion: as above, one requires that no adversary can distinguish if it accesses the perfectly random permutation $\mathcal{P}$ or the actual cipher. But in this case, the adversary not only accesses the permutation $\mathcal{O}_b$ itself, which is either $\mathcal{P}$ or $\mathcal{E}_\mathsf{k}$, but also its inverse $\mathcal{O}_b^{-1}$, which is thus either $\mathcal{P}^{-1}$ or $\mathcal{D}_\mathsf{k}$:

$$\mathsf{Adv}_{\pi}^{\mathsf{sprp}}(\mathcal{A}) = 2 \times \Pr\left[\begin{array}{l}\mathsf{k} \stackrel{R}{\leftarrow} \{0,1\}^k; \mathcal{P} \stackrel{R}{\leftarrow} \mathcal{SP}_\ell; \\ (\mathcal{O}_0, \mathcal{O}_0^{-1}) = (\mathcal{P}, \mathcal{P}^{-1}); (\mathcal{O}_1, \mathcal{O}_1^{-1}) = (\mathcal{E}_\mathsf{k}, \mathcal{D}_\mathsf{k}); \\ b \stackrel{R}{\leftarrow} \{0,1\}; b' \leftarrow \mathcal{A}^{\mathcal{O}_b, \mathcal{O}_b^{-1}}() : b' = b\end{array}\right] - 1.$$

Definition 5. *An encryption scheme π is said to be a (ε, t, n, m)-super pseudo-random permutation, denoted (ε, t, n, m)-*SPRP *if for any (t, n, m)-*SPRP *adversary $\mathcal{A}$, that asks at most n encryption queries and m decryption queries within time t, $\mathsf{Adv}_{\pi}^{\mathsf{sprp}}(\mathcal{A}) \leq \varepsilon$.*

2.6 Equivalences

For completeness, let us briefly recall a well-known result: indistinguishability and semantic security are equivalent security notions, if D is required to be efficiently samplable, and the predicate f to be efficiently computable. From a more concrete point of view, we can state the following theorem.

Theorem 6. *For any encryption scheme $\pi = (k, \ell, \mathcal{E}, \mathcal{D})$:*

$$\frac{1}{2} \times \mathsf{Adv}_{\pi}^{\mathsf{ind}}(t, e_1, d_1, e_2, d_2) \leq \mathsf{Adv}_{\pi}^{\mathsf{sem}}(t, e_1, d_1, e_2, d_2) \leq \mathsf{Adv}_{\pi}^{\mathsf{ind}}(t', e_1, d_1, e_2, d_2),$$

where $t' \leq t + 2T_D + T_f$, if the sampling time for D is bounded by T_D and the time to evaluate predicate f is bounded by T_f.

3 About the Indistinguishability of Ciphers

First, as already remarked, contrary to the probabilistic case, restrictions do not exist for the challenge only, which should not have been asked to the decryption oracle, but also for m_0 and m_1: they should not have been asked to the encryption oracle either, hence $m_0, m_1 \notin \Lambda_{\mathcal{E}}^{m}$ and $c \notin \Lambda_{\mathcal{D}}^{c}$.

3.1 Normal Adversary

Moreover, in the following, we restrict any adversary to behave like a *normal* adversary, which means that

- each query is asked at most once;
- if m has been asked as an encryption query (or to $\mathcal{O}_b$), with answer c, the query c will never be asked to the decryption oracle (or to $\mathcal{O}_b^{-1}$) later;
- if c has been asked as a decryption query (or to $\mathcal{O}_b^{-1}$), with answer m, the query m will never be asked to the encryption oracle (or to $\mathcal{O}_b$) later;
- for a (t, n)-PRP adversary (or (t, n, m)-SPRP adversary, respectively), the adversary makes exactly n queries to $\mathcal{O}_b$ (n queries to $\mathcal{O}_b$ and m queries to $\mathcal{O}_b^{-1}$, respectively) .

Proposition 7. *Any adversary can be made* normal *(with just additional look up in tables.)*

3.2 Adaptive Adversaries

Since we consider general adversaries, with possible oracle access, according to the values e_1, d_1, e_2 and d_2, for simpler notations we omit the oracle notation $\mathcal{A} = (\mathcal{A}_1^{\mathcal{E}_k, \mathcal{D}_k}, \mathcal{A}_2^{\mathcal{E}_k, \mathcal{D}_k})$ but simply use $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$. Oracle access is now implicit.

Adaptive Chosen-Plaintext Attacks. First, we review the property showed by Katz and Yung [6] about probabilistic symmetric encryption schemes. By the Corollary 10 below, we prove that it still holds for ciphers: an adaptive access to the encryption oracle after the challenge ciphertext is known does not significantly increase the power of an adversary which already had adaptive access to this oracle in the first stage.

Theorem 8. *For any cipher π:*

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1, d_1, e_2, d_2) \leq (2e_2 + 1) \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, d_2).$$

Proof. Let $\mathcal{A}$ be a (t, e_1, e_2, d_1, d_2)-normal adversary against indistinguishability. We denote by $\mathcal{A}[\epsilon_2]$ the new adversary $\mathcal{B}$ we build using $\mathcal{A}$, by restricting the interactions $\mathcal{A}$ actually has with the world. We indeed filter the queries it asks: all the queries asked by $\mathcal{A}_1$ are forwarded (as well as the answers); however, only the first ϵ_2 encryption queries are forwarded in the second stage, extra encryption queries are answered at random, but different from any previously involved ciphertext (the decryption queries, the ciphertext answers to encryption queries, and the challenge ciphertext.) We easily see that $\mathcal{A}[\epsilon_2]$ is normal. Note that $\mathcal{A}[e_2] = \mathcal{A}$ since in this case all the queries are forwarded, as well as the answers, whereas $\mathcal{A}[0]$ is in fact an adversary who makes no encryption query in the second stage, since the queries asked by $\mathcal{A}_2$ are answered at random, without querying $\mathcal{E}_{\mathsf{k}}$.

Lemma 9. *For any $1 \leq \epsilon_2 \leq e_2$:*

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[\epsilon_2]) - \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[\epsilon_2 - 1]) \leq 2 \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, d_2).$$

The proof of this lemma is quite similar but simpler than the proof of the Lemma 12 below. The full proof of the Lemma 12 is included below. By applying e_2 times this lemma, using a classical hybrid argument, one gets

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1, d_1, 0, d_2) + 2e_2 \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, d_2),$$

which implies the claimed result. □

In the particular case where $d_2 = 0$, one gets the following corollary which means that adaptive chosen-plaintext attacks do not give any additional power to an adversary.

Corollary 10. *For any cipher π:*

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1, d_1, e_2, 0) \leq (2e_2 + 1) \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1, 0, 0).$$

Adaptive Chosen-Plaintext and Chosen-Ciphertext Attacks. This result was already known. But the particular case of deterministic encryption admits improvements: under specific assumptions, an adaptive access to both the encryption oracle and the decryption oracle after the challenge ciphertext is known does not significantly increase the power of an adversary which already had access to these oracles in the first stage. Interestingly, the cost of the reduction is only linear in the (total) number of queries.

Theorem 11. *For any cipher* π: $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1, d_1, e_2, d_2)$ *is upper-bounded by*

$$\left(2(e_2+d_2)+1\right)\left(\begin{array}{c}\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1+e_2, d_1+d_2, 0, 0)\\ +\mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(t, d_1+d_2-1, e_1+e_2+2, 0, 0)\end{array}\right).$$

Proof. Let $\mathcal{A}$ be a (t, e_1, e_2, d_1, d_2)-normal adversary against indistinguishability. As above, we denote by $\mathcal{A}[n]$ the new adversary $\mathcal{B}$ we build using $\mathcal{A}$, by restricting the interactions $\mathcal{A}$ actually has with the world: all the queries in the first stage are forwarded, and the answers too, but only the first n queries are answered correctly in the second stage, while extra queries are answered at random but different from any message which previously appeared in the same category: if it is an encryption query, the answer must be different from any previously involved ciphertext (the decryption queries, the ciphertext answers to encryption queries, and the challenge); if it is a decryption query, the answer must be different from any previously involved plaintext (the encryption queries, the plaintext answers to decryption queries, and the two plaintexts output of $\mathcal{A}_1$). We easily see that $\mathcal{A}[n]$ is normal. Note that $\mathcal{A}[e_2+d_2] = \mathcal{A}$, since there are at most e_2+d_2 oracle queries in the second stage. However, $\mathcal{A}[0]$ is a lunchtime adversary, since all the queries in the second stage are answered at random, without querying any oracle.

Lemma 12. *For any* $n \leq e_2+d_2$*: the difference* $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n]) - \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n-1])$ *is upper-bounded by*

$$2 \times \left(\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1+e_2, d_1+d_2, 0, 0) + \mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(t, d_1+d_2-1, e_1+e_2+2, 0, 0)\right),$$

where t *is the running time of* $\mathcal{A}$*.*

Proof. We construct two adversaries $\mathcal{B}$ and $\mathcal{C}$, such that for each successful execution of $\mathcal{A}$, one of $\mathcal{B}$ or $\mathcal{C}$ is successful. The former is a $(t, e_1+e_2, d_1+d_2, 0, 0)$-$\mathsf{IND}$ adversary against π, while the latter is a $(t, d_1+d_2-1, e_1+e_2+2, 0, 0)$-$\mathsf{IND}$ adversary against π^{-1}.

Description of $\mathcal{B}$ *and* $\mathcal{C}$*.* Our adversaries $\mathcal{B}$ and $\mathcal{C}$ actually restrict the interactions $\mathcal{A}$ has, the same way as $\mathcal{A}[n-1]$ or $\mathcal{A}[n]$ would do: $\mathcal{B}_1$ and $\mathcal{C}_1$ run $\mathcal{A}_1$, forwarding any query/answer to their corresponding encryption/decryption oracles[1]. When $\mathcal{A}_1$ outputs (m_0, m_1), $\mathcal{B}_1$ and $\mathcal{C}_1$ choose a random bit b and get $c = \mathcal{E}_{\mathsf{k}}(m_b)$. This value requires one more encryption query to π for $\mathcal{B}_1$, while it requires one more decryption query to π^{-1} for $\mathcal{C}_1$. Then $\mathcal{B}_1$ and $\mathcal{C}_1$ run $\mathcal{A}_2(c)$ up to the n^{th} query q, still forwarding any query/answer to their corresponding oracles, except that last q one (the n^{th} query of $\mathcal{A}_2$). In the case that $\mathcal{A}_2$ makes less than n queries, $\mathcal{B}$ and $\mathcal{C}$ complete randomly their games by choosing immediately two random plaintexts different from any previous plaintext and outputting randomly the guesses. The advantages are thus exactly zero in this case. We thus now turn to the case where such a query q exists:

[1] Note that a query to $\mathcal{E}_{\mathsf{k}}$ corresponds to an encryption query to π (for $\mathcal{B}_1$), while it corresponds to a decryption query to π^{-1} (for $\mathcal{C}_1$), and similarly for a query to $\mathcal{D}_{\mathsf{k}}$.

- If q is an encryption query, $\mathcal{C}$ completes randomly its game in the above sense with a random answer since we do not care about it but only about $\mathcal{B}$, which attacks π as follows. $\mathcal{B}_1$ chooses a random plaintext q_0 for π, different from any previous plaintext (encryption queries and decryption answers), and then outputs $(q_0, q_1 = q)$. Thereafter, the challenge ciphertext $a = \mathcal{E}_{\mathsf{k}}(q_d)$ is produced, for a random bit d. On input a, $\mathcal{B}_2$ resumes $\mathcal{A}_2$ using a for answering the query q (note that $\mathcal{B}_2$ does not query on q). When $\mathcal{A}_2$ outputs its guess b' for the bit b, $\mathcal{B}_2$ outputs its guess d', for the bit d, that is defined by the boolean value of the test $b' = b$ (in other words, if $b' = b$, then $d' = 1$, else $d' = 0$).
- If q is a decryption query, $\mathcal{B}$ completes randomly its game in the above sense with a random answer since we do not care about it but only about $\mathcal{C}$, which attacks π^{-1} as follows. $\mathcal{C}_1$ chooses a random plaintext q_0 for π^{-1} (and thus a ciphertext for π), different from any previous plaintext for π^{-1} ($\mathcal{D}_{\mathsf{k}}$ queries and $\mathcal{E}_{\mathsf{k}}$ answers) but also from $\mathcal{E}_{\mathsf{k}}(m_{\overline{b}})$ ($\mathcal{C}_1$ must ask this further query —a decryption query for π^{-1}— to learn this value and avoid the collision), and then outputs $(q_0, q_1 = q)$. Thereafter, the challenge $a = \mathcal{D}_{\mathsf{k}}(q_d)$, a ciphertext for π^{-1}, is produced for a random bit d. On input a, $\mathcal{C}_2$ resumes $\mathcal{A}_2$ using a for answering the query q. When $\mathcal{A}_2$ outputs its guess b' for the bit b, $\mathcal{C}_2$ outputs its guess d', for the bit d, that is defined by the boolean value of the test $b' = b$ (in other words, if $b' = b$, then $d' = 1$, else $d' = 0$).

Advantages of $\mathcal{B}$ and $\mathcal{C}$. We first check that $\mathcal{B}$ and $\mathcal{C}$ satisfy the access restriction to the oracles, which is easy. Indeed, in the case $\mathcal{B}_1$ and $\mathcal{C}_1$ choose a random plaintext q_0 (when $\mathcal{A}$ makes the n^{th} query), they choose it different from any previous plaintext. Then, we know that $\mathcal{B}_2$ and $\mathcal{C}_2$ do not ask any other query, the access restriction to the decryption oracle is then satisfied. Let us now evaluate the number of queries:

- Algorithm $\mathcal{B}_1$ makes at most $e_1 + e_2$ encryption queries (all the encryption queries that $\mathcal{A}$ makes up to the n^{th} query q excepted q itself and it must make one more encryption query to get $c = \mathcal{E}_{\mathsf{k}}(m_b)$), and $d_1 + d_2$ decryption queries (all the decryption queries that $\mathcal{A}$ makes up to the n^{th} query);
- Algorithm $\mathcal{C}_1$ makes at most $d_1 + d_2 - 1$ encryption queries (all the decryption queries that $\mathcal{A}$ makes up to the n^{th} query q excepted the query q itself) and $e_1 + e_2 + 2$ decryption queries (all the encryption queries that $\mathcal{A}$ makes up to the n^{th} query, one more query to get $c = \mathcal{E}_{\mathsf{k}}(m_b)$, and one more query to learn the value $\mathcal{E}_{\mathsf{k}}(m_{\overline{b}})$).

About the running time, no extra computation has to be perform by either $\mathcal{B}$ or $\mathcal{C}$. We thus get the following upper-bounds, where t is the running time of $\mathcal{A}$:

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, 0),$$
$$\mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(\mathcal{C}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(t, d_1 + d_2 - 1, e_1 + e_2 + 2, 0, 0).$$

Let us now analyze the relation between the advantages of $\mathcal{B}$ and $\mathcal{C}$, and those of $\mathcal{A}[n]$ and $\mathcal{A}[n-1]$. We denote by Enc^q the event in which q is an encryption

query and we also denote by $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A} \,|\, \mathsf{Enc}^q)$ the conditional advantage of $\mathcal{A}$ providing the event Enc^q holds, that is

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A} \,|\, \mathsf{Enc}^q) = \Pr[\mathcal{A}() = 1 \,|\, b = 1 \wedge \mathsf{Enc}^q] - \Pr[\mathcal{A}() = 1 \,|\, b = 0 \wedge \mathsf{Enc}^q].$$

– if q is an encryption query, we have a non trivial adversary $\mathcal{B}$:

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B}) = 2\Pr[d' = d] - 1 = \Pr[d' = 1 \,|\, d = 1] - \Pr[d' = 1 \,|\, d = 0].$$

When $d = 1$, the distribution of b and b' used by $\mathcal{B}$ is exactly the same as the usual attack game for $\mathcal{A}[n]$, since a is the correct answer of $q_1 = q$. When $d = 0$, the answer of the encryption query q (w.r.t. π) is a, the encryption of q_0 (a random distinct message), and thus a random ciphertext different from any previously involved ciphertext because of the permutation propriety of the cipher. The last remark shows that $\mathcal{B}$ is identical to $\mathcal{A}[n-1]$. Since $d' = 1$ means $b' = b$, we have[2]:

$$\begin{aligned}\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B} \,|\, \mathsf{Enc}^q) &= \Pr[d' = 1 \,|\, d = 1 \wedge \mathsf{Enc}^q] - \Pr[d' = 1 \,|\, d = 0 \wedge \mathsf{Enc}^q] \\ &= \frac{1}{2} \cdot \left(\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n] \,|\, \mathsf{Enc}^q) - \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n-1] \,|\, \mathsf{Enc}^q)\right).\end{aligned}$$

– if q is a decryption query, a similar argument can be provided for the adversary $\mathcal{C}$: when $d = 1$, $\mathcal{C}$ is identical to $\mathcal{A}[n]$ and when $d = 0$, $\mathcal{C}$ is identical to $\mathcal{A}[n-1]$ because the encryption of q_0 (a random distinct message) for $\mathcal{C}$ is a random plaintext different from any previous plaintext (included m_0 and m_1.) Therefore, we have[2]:

$$\begin{aligned}\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{C} \,|\, \overline{\mathsf{Enc}^q}) &= \Pr[d' = 1 \,|\, d = 1 \wedge \overline{\mathsf{Enc}^q}] - \Pr[d' = 1 \,|\, d = 0 \wedge \overline{\mathsf{Enc}^q}] \\ &= \frac{1}{2} \cdot \left(\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n] \,|\, \overline{\mathsf{Enc}^q}) - \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[n-1] \,|\, \overline{\mathsf{Enc}^q})\right).\end{aligned}$$

In the above formula, $\overline{\mathsf{Enc}^q}$ denotes the negation of event Enc^q. With the remark that $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B} \,|\, \overline{\mathsf{Enc}^q}) = 0$ and $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{C} \,|\, \mathsf{Enc}^q) = 0$, we have:

$$\begin{aligned}\Pr[\mathsf{Enc}^q] \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B} \,|\, \mathsf{Enc}^q) &= \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{B}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi}(e_1 + e_2, d_1 + d_2, 0, 0), \\ \Pr[\overline{\mathsf{Enc}^q}] \times \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{C} \,|\, \overline{\mathsf{Enc}^q}) &= \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{C}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(d_1 + d_2 - 1, e_1 + e_2 + 2, 0, 0).\end{aligned}$$

Combined with the two above equations, this leads to the expected result. □

Starting from $\mathcal{A} = \mathcal{A}[e_2 + d_2]$, and applying $e_2 + d_2$ times the above relation, one gets:

$$\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}[0]) + 2(e_2+d_2)\begin{pmatrix}\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, 0) \\ +\mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(t, d_1 + d_2 - 1, e_1 + e_2 + 2, 0, 0)\end{pmatrix}.$$

Since $\mathcal{A}[0]$ is a $(t, e_1, d_1, 0, 0)$-IND adversary, and thus its advantage is bounded by $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, e_1 + e_2, d_1 + d_2, 0, 0)$, one gets the result. □

[2] We remind that $\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A} \,|\, \mathsf{E})$ denotes the conditional advantage of any adversary $\mathcal{A}$ providing the event E holds.

In many ciphers, the encryption algorithm and the decryption algorithm are similar. Therefore, if the cipher is secure against any lunchtime adversary (IND-P1-C1), its inverse achieves a similar security level. The above theorem implies that the cipher is actually secure against any adaptive adversary (IND-P2-C2): thus, adaptive attacks do not help against symmetric and deterministic encryption schemes.

4 Indistinguishability and Pseudo-Randomness

In this section, we give a relation between the notion of indistinguishability defined above and the classical security notions for ciphers, namely to provide a pseudo-random permutation family or a super pseudo-random permutation family.

4.1 IND-P1-C0 is Equivalent to Pseudo-Randomness

In [2], Desai and Miner claimed that:

Proposition 13. *For any cipher π:*

$$\frac{1}{2} \times \mathsf{Adv}_{\pi}^{\mathsf{ind}}(t, e_1, 0, 0, 0) \leq \mathsf{Adv}_{\pi}^{\mathsf{prp}}(t, e_1 + 1) \leq (e_1 + 1) \times \mathsf{Adv}_{\pi}^{\mathsf{ind}}(t, e_1 + 1, 0, 0, 0).$$

We prove this proposition (which has not been published anywhere) in the following two theorems whose results are more general. In fact, the left relation is a particular case of Theorem 14 where $d_1 = e_2 = d_2 = 0$, while the right relation is a particular case of the proof of Theorem 15 where $n = e_1 + 1$ and $m = 0$. Since we know that the last query is always an encryption query, the second term disappears. We just have to build the adversary $\mathcal{B}$.

4.2 IND-P2-C2 is "Almost" Equivalent to Super Pseudo-Randomness

The first theorem is the intuitive and easy direction:

Theorem 14. *For any cipher π:*

$$\mathsf{Adv}_{\pi}^{\mathsf{ind}}(t, e_1, d_1, e_2, d_2) \leq 2 \times \mathsf{Adv}_{\pi}^{\mathsf{sprp}}(t, e_1 + e_2 + 1, d_1 + d_2).$$

Proof. We are assuming that π is SPRP-secure. We then show that π is also secure in the sense of IND-P2-C2. Let $\mathcal{A}$ to be a (t, e_1, d_1, e_2, d_2)-IND adversary attacking π. We want to show that $\mathsf{Adv}_{\pi}^{\mathsf{ind}}(\mathcal{A})$ is negligible. To this end, we describe a SPRP adversary $\mathcal{B}$ which attacks π by using $\mathcal{A}$ as a sub-program.

Description of $\mathcal{B}^{\mathcal{O}_b, \mathcal{O}_b^{-1}}$. Our adversary $\mathcal{B}$ runs $\mathcal{A}_1$ by answering its encryption/decryption queries, which are simply forwarded to the oracles $\mathcal{O}_b$ and $\mathcal{O}_b^{-1}$, respectively. When $\mathcal{A}_1$ outputs (m_0, m_1), $\mathcal{B}$ chooses a random bit d and gets

$y_d = \mathcal{O}_b(m_d)$. $\mathcal{B}$ then runs $\mathcal{A}_2(y_d)$, still forwarding all the encryption/decryption queries of $\mathcal{A}$ to the oracles $\mathcal{O}_b$ and $\mathcal{O}_b^{-1}$, respectively. When $\mathcal{A}_2$ outputs its guess d' for the bit d, $\mathcal{B}$ outputs its guess b', for the bit b, that is defined by the boolean value of the test $d' = d$ (i.e, if $d' = d$, then $b' = 1$, else $b' = 0$).

Advantage of $\mathcal{B}$. We now consider the relation between the advantage of $\mathcal{B}$ and the advantage of $\mathcal{A}$.

- in the case $b = 1$, this game is exactly the game in which $\mathcal{A}$ plays against π. The probability that $\mathcal{B}$ outputs $b' = 1$ is therefore the probability that $d' = d$: $(\mathsf{Adv}^{\mathsf{ind}}_{\pi}(\mathcal{A}) + 1)/2$.
- in the case $b = 0$, because $\mathcal{A}$ queries a random permutation, and $y_d = \mathcal{P}(m_d)$ is perfectly independent with m_0 and m_1, $\mathcal{A}_2$ therefore gives an answer $d' = d$ with probability $1/2$. Consequently, $\mathcal{B}$ gives $b' = 1$ with probability $1/2$.

Combining these two cases, in which $\mathcal{A}$ is a (t, e_1, d_1, e_2, d_2)-IND adversary and $\mathcal{B}$ is a $(t, e_1 + e_2 + 1, d_1 + d_2)$-$\mathsf{SPRP}$ adversary, we get the expected result. $\mathcal{B}$ indeed asks $e_1 + e_2 + 1$ queries to $\mathcal{O}_b$, because of the extra query to get y_d. □

The other direction is less natural, and much more surprising:

Theorem 15. *For any cipher π:*

$$\mathsf{Adv}^{\mathsf{sprp}}_{\pi}(t, n, m) \leq (n + m) \times \left(\mathsf{Adv}^{\mathsf{ind}}_{\pi}(t, n, m, 0, 0) + \mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(t, m, n, 0, 0) \right).$$

Proof. Let $\mathcal{A}$ be a (t, n, m)-SPRP normal adversary against π. We denote by $\mathcal{A}[\eta]$ the hybrid adversary $\mathcal{B}$, built using $\mathcal{A}$ by restricting its interactions: the first η queries to the oracles are answered by $\mathcal{E}_{\mathsf{k}}$ (for an encryption query – oracle $\mathcal{O}$) and by $\mathcal{D}_{\mathsf{k}}$ (for a decryption query – oracle $\mathcal{O}^{-1}$), the following queries are answered by $\mathcal{P}$ and $\mathcal{P}^{-1}$ respectively. The goal of the adversary is always to output a bit b'. We define $\mathsf{PI}(\mathcal{B})$ to be the probability that any adversary $\mathcal{B}$ gives the answer $b' = 1$. We thus have:

$$\begin{aligned}\mathsf{Adv}^{\mathsf{sprp}}_{\pi}(\mathcal{A}) &= \Pr[\mathcal{A}() = 1 \mid b = 1] - \Pr[\mathcal{A}() = 1 \mid b = 0] \\ &= \mathsf{PI}[\mathcal{A}^{\mathcal{E}_{\mathsf{k}}, \mathcal{D}_{\mathsf{k}}}() = 1] - \Pr[\mathcal{A}^{\mathcal{P}, \mathcal{P}^{-1}}() = 1] = \mathsf{PI}(\mathcal{A}[n + m]) - \mathsf{PI}(\mathcal{A}[0]).\end{aligned}$$

Lemma 16. *For any $\eta \leq n + m$:*

$$\mathsf{PI}(\mathcal{A}[\eta]) - \mathsf{PI}(\mathcal{A}[\eta - 1]) \leq \mathsf{Adv}^{\mathsf{ind}}_{\pi}(n, m, 0, 0) + \mathsf{Adv}^{\mathsf{ind}}_{\pi^{-1}}(m, n, 0, 0).$$

This proof is similar to the one of the Lemma 12. The idea is that we construct two adversaries, a $(t, n, m, 0, 0)$-adversary $\mathcal{B}$ against π and a $(t, m, n, 0, 0)$-adversary $\mathcal{C}$ against π^{-1} such that one of their advantages is exactly equal to the left-hand side. These two adversaries run $\mathcal{A}$ up to the η^{th} query of $\mathcal{A}[\eta]$ using $\mathcal{E}_{\mathsf{k}}$ for answering a query to $\mathcal{O}_b$ and using $\mathcal{D}_{\mathsf{k}}$ for answering a query to $\mathcal{O}_b$. According to the type of the η^{th} query of $\mathcal{A}[\eta]$ (an encryption query or a decryption

query), $\mathcal{B}_1$ or $\mathcal{C}_1$ outputs this query as one of its two chosen messages (the other is chosen randomly) and then $\mathcal{B}_1$ or $\mathcal{C}_1$ gives its received challenge as the answer to the η^{th} query of $\mathcal{A}$. $\mathcal{B}_2$ or $\mathcal{C}_2$ then outputs its guess according to the guess of $\mathcal{A}$ without making any query.

Applying $n + m$ times this lemma, we obtain the expected result. □

From these two theorems, we see that a cipher is a super pseudo-random permutation if and only if itself and its inverse achieve semantic security against any lunchtime adversary (IND-P1-C1). In other words, under the conjecture that a cipher and its inverse achieve a similar security level secure against any lunchtime adversary, SPRP and IND-P1-C1 are equivalent with a linear-cost reduction.

The more intuitive equivalence, between IND-P2-C2 and SPRP, can be obtained under a weaker condition: if π^{-1} is just IND-P1-C0. This result is given in details in the full version [9].

Acknowledgement

The work described in this paper has been supported in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT. The information in this document reflects only the authors' views, is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

References

1. M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A Concrete Security Treatment of Symmetric Encryption: Analysis of the DES Modes of Operation. In *Proc. of the 38th FOCS*. IEEE, New York, 1997.
2. A. Desai and S. Miner. Concrete Security Characterization of PRFs and PRPs: Reduction and Applications. In *Asiacrypt '00*, LNCS 1976, pages 503–516. Springer-Verlag, Berlin, 2000.
3. O. Goldreich, S. Goldwasser, and S. Micali. On The Cryptographic Applications of Random Functions. In *Crypto '84*, LNCS 196. Springer-Verlag, Berlin, 1985.
4. S. Goldwasser and S. Micali. Probabilistic Encryption. *Journal of Computer and System Sciences*, 28:270–299, 1984.
5. S. Halevi and P. Rogaway. A Tweakable Enciphering Mode. In *Crypto '03*, LNCS 2729, pages 482–499. Springer-Verlag, Berlin, 2003.
6. J. Katz and M. Yung. Complete Characterization of Security Notions for Probabilistic Private-Key Encryption. In *Proc. of the 32nd STOC*. ACM Press, New York, 2000.
7. M. Luby and Ch. Rackoff. How to Construct Pseudorandom Permutations from Pseudorandom Functions. *SIAM Journal of Computing*, 17(2):373–386, 1988.
8. M. Naor and M. Yung. Public-Key Cryptosystems Provably Secure against Chosen Ciphertext Attacks. In *Proc. of the 22nd STOC*, pages 427–437. ACM Press, New York, 1990.

9. D. H. Phan and D. Pointcheval. About the Security of Ciphers (Semantic Security and Pseudo-Random Permutations). In *SAC '04*. Springer-Verlag, Berlin, 2004. Full version available from `http://www.di.ens.fr/users/pointche/`.
10. C. Rackoff and D. R. Simon. Non-Interactive Zero-Knowledge Proof of Knowledge and Chosen Ciphertext Attack. In *Crypto '91*, LNCS 576, pages 433–444. Springer-Verlag, Berlin, 1992.
11. C. E. Shannon. Communication Theory of Secrecy Systems. *Bell System Technical Journal*, 28(4):656–715, 1949.