From Known-Plaintext to Chosen-Ciphertext Security*

Ueli Maurer and Johan Sjödin

Department of Computer Science, ETH Zurich, CH-8092 Zurich, Switzerland
{maurer, sjoedin}@inf.ethz.ch

Abstract. Motivated by the quest of reducing assumptions in security proofs in cryptography, this paper is concerned with designing efficient symmetric encryption and authentication schemes based on weak pseudo-random functions (WPRF), which can potentially be much more efficiently implemented than PRFs. Damgård and Nielsen (Crypto '02) showed how to construct a symmetric encryption scheme based on any WPRF that is provably secure under a chosen-plaintext attack. The main ingredient is a construction of a variable-output-length WPRF from any WPRF.

The results of this paper are three-fold. First, we optimize the Damgård-Nielsen encryption method by constructing a more efficient variable-output-length WPRF from any WPRF. Our construction is optimal for a large and natural class of reductions. Second, we propose an efficient PRF from any WPRF. Third, these two results imply the first efficient symmetric encryption scheme based on any WPRF that is provably secure under a chosen-ciphertext-attack, and they also solve open questions posed by Naor and Reingold (Crypto '98) and by Damgård and Nielsen.

Keywords: Weak Pseudo-Random Function, Symmetric Encryption, Known-Plaintext-Attack, Chosen-Ciphertext-Attack

1 Introduction

1.1 Weakening of Cryptographic Assumptions

A general goal in cryptography is to prove the security of cryptographic systems under assumptions that are as weak as possible. Provably secure encryption and authentication schemes based on a *pseudo-random function* (PRF) [13] have been studied extensively in the literature [12]. Informally, a PRF is a function with a secret key that cannot be efficiently distinguished from a uniform random function even when it can be queried adaptively (i.e., under a chosen-plaintext-attack (CPA)).

The notion of a PRF is very strong and, indeed, it is unclear whether functions such as block ciphers proposed in the literature have this very strong security property.[1] When designing cryptographic schemes, it is prudent to postulate weaker properties

* This work was partially supported by the Zurich Information Security Center. It represents the views of the authors.

[1] For example, the design criteria for AES did not include a requirement that a candidate proposal be a PRF, only that it be secure as a block cipher in certain modes of operation, against certain types of attacks.

as this makes it more likely that a certain function has this property or, equivalently, there are potentially much more efficient implementations for the weaker requirement compared to the stronger.

A very promising weaker notion of security, a *weak* PRF (WPRF), was recently proposed by Naor and Reingold [19] (see also [10]) and has already found several applications [20, 1, 10, 21, 17]. Informally, a WPRF is a function with a secret key that cannot be efficiently distinguished from a uniform random function when given a sequence of *random*, as opposed to adaptively chosen, inputs and the corresponding outputs (i.e., under a *known-plaintext-attack* (KPA)). Highly efficient candidates for WPRFs are described in [8] (cf. [20]), although these are not targeted at this particular security notion explicitly.

While the design of WPRFs has not been studied as extensively as PRFs, a concrete argument showing that the WPRF notion is substantially weaker than the PRF notion is that WPRFs can have rather strong structural properties which are known to be devastating for PRFs. For instance, if $\mathcal{G}$ is a group of prime order p in which the Decisional Diffie-Hellman (DDH) [11] assumption holds, then

$$F : \{1, \ldots, p\} \times \mathcal{G} \to \mathcal{G}, \text{ defined by } F_k(x) := F(k, x) = x^k, \tag{1}$$

is a WPRF [10] that commutes (i.e., $F_k(F_{k'}(x)) = F_{k'}(F_k(x))$). A WPRF can also be self inverse (i.e., $F_k(F_k(x)) = x$), have a small fraction of bad points (e.g., $F_k(x) = x$ or $F_k(x) = k$), and have related outputs (e.g., $F_k(x\|1) = F_k(x\|0)$ for all x). Such structural flaws make most encryption and authentication schemes based on PRFs completely insecure (for examples, see [10]).

In this paper we propose provably secure encryption and authentication schemes, for the strongest security notion, under the sole assumption of a WPRF. Of course, the security could also be based on even weaker assumptions like the one-wayness of a certain function, since a PRF can be obtained from any one-way function [14, 13]. However, such schemes are not of practical interest due to their inefficiency. More efficient methods exist [12] based on pseudo-random generators (PRG) [9, 25], but the drawback is that these schemes are all stateful, a typically undesirable property.

1.2 Contributions and Related Work

Our work is motivated by Damgård and Nielsen's elegant work on WPRFs [10]. In their paper, they propose the Pseudorandom Tree (PRT) construction for transforming any WPRF $F : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^\ell$ into a variable-output-length[2] (VOL) WPRF

$$\mathrm{PRT}^F : \{0,1\}^{\kappa'} \times \{0,1\}^n \times \mathbb{N} \to \{0,1\}^*.$$

[2] A variable output length function family $V : \{0,1\}^\kappa \times \{0,1\}^n \times \mathbb{N} \to \{0,1\}^*$ satisfies $|V_k(x,l)| = l$ for all k, x, and l.

Furthermore, they show how to construct an efficient probabilistic symmetric encryption scheme from F that is provably secure under a CPA. They achieve this by encrypting a message $m \in \{0,1\}^*$, under a key $k \in \{0,1\}^{\kappa'}$ and some auxiliary uniform randomness $r \in \{0,1\}^n$, as

$$(k, r, m) \mapsto \Big(r, \mathrm{PRT}_k^F(r, |m|) \oplus m\Big). \tag{2}$$

The efficiency of these constructions are measured by the number of applications to F.

In this paper we propose the following:

1. *The Increasing Chain Tree (ICT) construction of a VOL-WPRF from any WPRF.* Recently, two other VOL-WPRF constructions, the Expanded PRT (ERT) and the Factorial Tree (FCT), were introduced in [18]. The former is slightly more efficient than PRT, whereas the latter is flawed (see Sect. 3.2). Our ICT-construction is more efficient and uses a shorter key than ERT and PRT. Indeed, we show that ICT is optimal for a large and natural class of constructions (and hence also the corresponding encryption scheme, defined by replacing PRT by ICT in (2)).

2. *The Increasing Chain (IC) construction of a PRF from any WPRF.* The IC-construction is more efficient and uses a shorter key than Naor and Reingold's transformation of a WPRF into a PRF[3] [20]. This solves their open question in [19, p. 278]. IC is similar in nature to Goldreich, Goldwasser, and Micali's (GGM) construction of a PRF from any PRG [13], but it is more efficient than first transforming the WPRF into a PRG and then applying the GGM-construction. In particular, IC^F with F as defined in (1) implies a non-trivial reduction of the key material of Naor and Reingold's construction in [21] of a PRF based on the DDH assumption (i.e., the key is not replaced by the output of a PRG based on F).

3. *A CCA-secure encryption scheme from any WPRF.* Our results, combined with a Wegman-Carter [24] based message authentication code (MAC) and the encrypt-then-MAC method [16, 6], imply the first efficient (i.e., practical) encryption scheme from any WPRF that is secure under a *chosen-ciphertext-attack* (CCA). This solves an open question posed by Naor and Reingold in [19, p. 279], and by Damgård and Nielsen in [10, p. 464].

4. *A non-adaptive*[4] *CCA-secure encryption scheme from any WPRF and WMAC*[5]*.* Although this type of security may be unsatisfactory in practice, we consider our mode of operation to be of possible theoretical interest. Formulating the exact re-

[3] I.e., via the construction of a pseudo-random synthesizer.

[4] CCA security formalizes an adversary A's inability, given access to an encryption and decryption oracle, to distinguish between two plaintexts given the encryption c of one of them. *Non-adaptivity* means that A does not have access to the oracles after c is presented.

[5] A *weak message authentication code* (WMAC), implicitly given in [19], is a KPA-secure MAC.

quirements for achieving standard security notions are interesting in its own right, and might motivate further research on basing strong primitives on weak assumptions. Non-adaptive CCA-security has been studied under stronger assumptions in [19].

2 Preliminaries

2.1 Notation and Definitions

Let $s \overset{\$}{\leftarrow} S$ denote the operation of selecting s uniformly at random from the set S. If $\mathcal{D}$ is a probability distributions over S then $s \leftarrow \mathcal{D}$ denotes the operation of selecting s at random according to $\mathcal{D}$. Let $\mathcal{U}_n$ denote the uniform distribution over $\{0,1\}^n$. Furthermore, $\mathcal{R}_{L,\ell}$ and $\mathcal{R}_{\leq L,\ell}$ denote random functions with range $\{0,1\}^\ell$, and domain $\{0,1\}^L$ and $\{0,1\}^{\leq L} := \cup_{i=1}^{L}\{0,1\}^i$, respectively. $A^{\mathcal{O}}$ denotes an algorithm A with oracle access to $\mathcal{O}$, and $\Pr[\Pi : \mathcal{E}]$ is the probability that event $\mathcal{E}$ occurs in random experiment Π. For two functions f and g, let $f \circ g\,(x) := f(g(x))$. If x and y are two bitstrings, $x\|y$ denotes their concatenation, $x[i]$ the i^{th} bit of x, and $x[i,j] := x[i]\|x[i+1]\|\cdots\|x[j]$ with $i < j$.

2.2 Cryptographic Functions

We state our results in the concrete security framework, introduced by Bellare, Kilian, and Rogaway in [5] and used in many subsequent works [3, 4, 6]. Let $\mathcal{O}^f$ denote an oracle which, if invoked, returns $(r, f(r))$, where f is a function and r a uniformly at random chosen input of f. Let $F : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^\ell$ be a function family and $g : \{0,1\}^\kappa \to \{0,1\}^n$ a function (with $\kappa < n$). The $\mathbf{w}$-*advantage* of adversary A, for $\mathbf{w} \in \{\mathbf{prf}, \mathbf{wprf}, \mathbf{mac}, \mathbf{wmac}, \mathbf{prg}\}$, is defined as

$$\begin{aligned}
\mathbf{Adv}^{\mathbf{prf}}_{F,A} &:= \Pr\big[k \leftarrow \mathcal{U}_\kappa, b \leftarrow A^{F_k} : b = 1\big] - \Pr\big[R \leftarrow \mathcal{R}_{n,\ell}, b \leftarrow A^{R} : b = 1\big] \\
\mathbf{Adv}^{\mathbf{wprf}}_{F,A} &:= \Pr\Big[k \leftarrow \mathcal{U}_\kappa, b \leftarrow A^{\mathcal{O}^{F_k}} : b = 1\Big] - \Pr\Big[R \leftarrow \mathcal{R}_{n,\ell}, b \leftarrow A^{\mathcal{O}^{R}} : b = 1\Big] \\
\mathbf{Adv}^{\mathbf{mac}}_{F,A} &:= \Pr\left[k \leftarrow \mathcal{U}_\kappa, (m,\tau) \leftarrow A^{F_k}, b = \begin{cases} 1 & \text{if } \tau = F_k(m),\ m \text{ ``new''} \\ 0 & \text{otherwise} \end{cases} : b = 1\right] \\
\mathbf{Adv}^{\mathbf{wmac}}_{F,A} &:= \Pr\left[k \leftarrow \mathcal{U}_\kappa, (m,\tau) \leftarrow A^{\mathcal{O}^{F_k}}, b = \begin{cases} 1 & \text{if } \tau = F_k(m),\ m \text{ ``new''} \\ 0 & \text{otherwise} \end{cases} : b = 1\right] \\
\mathbf{Adv}^{\mathbf{prg}}_{g,A} &:= \Pr[k \leftarrow \mathcal{U}_\kappa, b \leftarrow A(g(k)) : b = 1] - \Pr[r \leftarrow \mathcal{U}_n, b \leftarrow A(r) : b = 1]
\end{aligned}$$

and the corresponding maximal advantages as

$$\mathbf{Adv}^{\mathbf{w}}_{F}(t,q) := \max_{A}\{\mathbf{Adv}^{\mathbf{w}}_{F,A}\} \quad \text{and} \quad \mathbf{Adv}^{\mathbf{prg}}_{g}(t) := \max_{A}\{\mathbf{Adv}^{\mathbf{prg}}_{g,A}\},$$

where the maximum is taken over all A restricted to q (respectively $q-1$ in case of the **mac** or **wmac** notion) invocations of its oracle and the standard time-complexity t.[6]

Next, we describe both the notion of a variable-input-length (VIL) function family and a variable-output-length (VOL) function family.

Definition 1. *A function family $F : \{0,1\}^\kappa \times \{0,1\}^{\leq N} \to \{0,1\}^n$ is referred to as having* variable-input-length *(VIL).*

The **vil-w**-advantage $\mathbf{Adv}_{F,A}^{\textbf{vil-w}}$ of adversary A, for $\mathbf{w} \in \{\mathbf{prf}, \mathbf{wprf}, \mathbf{mac}, \mathbf{wmac}\}$, is defined as $\mathbf{Adv}_{F,A}^{\mathbf{w}}$ but with $R \leftarrow \mathcal{R}_{*,\ell}$, and the maximal **vil-w**-advantage as

$$\mathbf{Adv}_{F}^{\textbf{vil-w}}(t,q,\mu) := \max_{A}\{\mathbf{Adv}_{F,A}^{\textbf{vil-w}}\},$$

where the maximum is taken over all A restricted to time-complexity t and at most q (respectively $q-1$ in case of the **mac** or the **wmac** notion) oracle invocations of F for which the total length of the inputs to F is at most μ bits (including the forgery message in case of the **mac** and **wmac** notion).

Definition 2. *A function family $F : \{0,1\}^\kappa \times \{0,1\}^n \times \mathbb{N} \to \{0,1\}^*$ is referred to as having* variable-output-length *(VOL) if $|F_k(x,l)| = l$, for all k, x, and l.*

The security notion of a VIL-WPRF is defined as follows [10]. Let $\mathcal{R}_{n,*}$ denote the following probabilistic function. On input (x,l) check whether a string o_x is defined; if not define it to be the empty string. Then check whether o_x has length at least l; if not append to o_x a uniformly at random chosen string from $\{0,1\}^{l-|o_x|}$. Then output $o_x[1,l]$. Let $\mathcal{O}_{\mathbf{vol}}^{f}$ be an oracle that on input l outputs $(x, f(x,l))$ for uniformly at random chosen x. The **vol-wprf**-advantage of adversary A in attacking F is

$$\mathbf{Adv}_{F,A}^{\textbf{vol-wprf}} := \Pr\Big[k \leftarrow \mathcal{U}_\kappa, b \leftarrow A^{\mathcal{O}_{\mathbf{vol}}^{F_k}} : b=1\Big] - \Pr\Big[k \leftarrow \mathcal{U}_\kappa, b \leftarrow A^{\mathcal{O}_{\mathbf{vol}}^{\mathcal{R}_{n,*}}} : b=1\Big]$$

and the maximal **vol-wprf**-advantage as

$$\mathbf{Adv}_{F}^{\textbf{vol-wprf}}(t,q,\mu) := \max_{A}\{\mathbf{Adv}_{F,A}^{\textbf{vol-wprf}}\},$$

where the maximum is taken over all A restricted to time-complexity t and at most q samples for which the output lengths of F total at most μ bits.

[6] I.e., we distinguish between the case when the advantage of an adversary A is a difference between the probability that an event (i.e., $b = 1$ in the above cases) occurs in two different random experiments and the case when the advantage is simply the probability that an event occurs in a single experiment. In the former case, t is the maximum of the worst-case total running time of the different experiments, and in the latter case, t is the worst-case total running time of the experiment (in some fix RAM model of computation). We also adopt the convention that t includes the length of the RAM program describing A.

3 The IC and ICT Constructions

In this section we introduce the IC-construction, for transforming a WPRF into a PRF, and the ICT-construction, for transforming a WPRF into a VOL-WPRF. Throughout this section, let $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ be a function family and t_F the worst-case running time for computing F.

3.1 A PRF from any WPRF

The IC-construction transforms $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ into

$$\mathrm{IC}^F : (\{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n) \times \{0,1\}^N \to \{0,1\}^n,$$

for some fix N, and is defined by the following algorithm for computing $\mathrm{IC}^F_{k_0,r,\tau_0}(x)$:

```
for $i = 1$ to $|x|-1$ do
  $k_i = F_{k_{i-1}}(r)$
for $i = 1$ to $|x|$ do
  if $x[i] = 1$ then
    $\tau_i = F_{k_{i-1}}(\tau_{i-1})$
  else
    $\tau_i = \tau_{i-1}$
return $\tau_{|x|}$
```

Note that F is invoked at most $2N - 1$ times. However, the first $N - 1$ invocations can be preprocessed and cached, and hence at most N invocations are necessary or, to be precise, as many invocations as there are ones in the input.

The following theorem is proved in in Appendix A.[7]

Theorem 1. *For any t, q, and input length N of* IC^F

$$\mathbf{Adv}^{\mathbf{prf}}_{\mathrm{IC}^F}(t,q) \leq N \cdot \left(\mathbf{Adv}^{\mathbf{wprf}}_{F}(t,q) + \frac{q(q+1)}{2^{n+1}} \right).$$

Remark 1. In [21], Naor and Reingold presented an efficient construction of a PRF based on the DDH assumption. It is easy to verify, that IC^F with F as defined in (1) is the same construction but with a non-trivial reduction of key material by a factor of roughly N (recall that N is the input length of IC^F). To be more precise, the first for-loop (above) generates a sequence $k_0, \ldots, k_{N-1}$ of keys from the initial key (k_0, r, τ_0) and the second for-loop exactly corresponds to the Naor-Reingold construction with $k_0, \ldots, k_{N-1}$ as its key. The reduction of key-material is non-trivial because, interestingly, $k_0, \ldots, k_{N-1}$ is not pseudo-random. For instance $F^{-1}_{k_1}(k_2) = F^{-1}_{k_2}(k_3)$ holds which can easily be verified given k_1, k_2, k_3. Furthermore, it can be shown that the second argument (i.e., r) of the initial key (k_0, r, τ_0), need not be kept secret.

[7] We refer to [10] for constructing an n-bit block WPRF F from any WPRF.

3.2 A VOL-WPRF from any WPRF

The ICT-construction, illustrated in Fig. 1(a) in Appendix D, is defined as

$$\begin{aligned} \mathrm{ICT}^F : \{0,1\}^{2n} \times \{0,1\}^n \times \mathbb{N} &\to \{0,1\}^* \\ (k\|r, x, l) &\mapsto \left(\mathrm{IC}^F_{k,r,x}(\langle 1\rangle)\| \cdots \| \,\mathrm{IC}^F_{k,r,x}(\langle \lceil l/n\rceil\rangle)\right)[1, l], \end{aligned}$$

where $\langle i\rangle$ denotes the standard bit encoding of the integer i. It is easy to verify that for computing $\mathrm{ICT}^F(k\|r, x, l)$, we need $d = \lfloor \log_2(\lceil l/n\rceil)\rfloor$ calls to F for generating (or pre computing) the keys $k_0, \ldots, k_d$ in the above IC-algorithm and further $\lceil l/n\rceil$ calls (at the most) for computing the output. In other words, the overhead applications, i.e., the number of applications of F minus the number of n-bit output blocks, are at most $\lfloor \log_2(\lceil l/n\rceil)\rfloor$. A proof of the following theorem is provided in Appendix C.

Theorem 2. *For any t, q, and μ*

$$\mathbf{Adv}^{\textbf{vol-wprf}}_{\mathrm{ICT}^F}(t, q, \mu) \le d_{\max} \cdot \mathbf{Adv}^{\textbf{wprf}}_F(t', q2^{d_{\max}-1} + 1) + \frac{4^{d_{\max}} \cdot q^2 - q}{2^{n+1}},$$

where $t' = t + \mathcal{O}(\frac{q \cdot l_{\max}}{n} \cdot t_F)$, $d_{\max} = \lfloor \log_2(\lceil l_{\max}/n\rceil)\rfloor + 1$, and $l_{\max} \le \mu$ is the maximum allowed output length of ICT^F.

It can be shown that the second half of the key (i.e., r) of ICT need not be kept secret and can hence be publicly generated.

Comparisons with other constructions. The idea behind ICT (of this paper), PRT of [10], and ERT of [18] is to first generate keys $k_0, \ldots, k_d$ from the initial key and then derive the output sequentially as follows. Each n-bit output block is derived by invoking F_{k_i} (with $i \in \{0, \ldots, d\}$) to the input or a previously computed output block.

ICT is superior to PRT and ERT for three reasons, see Fig. 1 in Appendix D. First, the initial key of ICT is n bits (plus n bits which may be publicly known) verses $3n$ bits for PRT and ERT. Second, ICT needs d invocations of F to generate the keys $k_0, \ldots, k_d$ whereas PRT and ERT needs $2d + 1$. Third, the maximal output size using $k_0, \ldots, k_d$ is $(2^{d+1} - 1)n$ for ICT, roughly $(3^{\frac{d+1}{2}} - 1)n$ for ERT, and roughly $(2^{\frac{d+1}{2}+1} - 2)n$ for PRT.[8]

Finally, we point out that the security proof of the FCT-construction (in [18]) is flawed. Moreover, it is obvious (to see from the definition) that FCT is insecure for WPRFs that commutes. Since such WPRFs exist under the DDH assumption (see (1)), a fix of the security proof would contradict the assumption and thus be a major breakthrough in number theory.

[8] The latter two values are exact if d is odd. Otherwise $(2 \cdot 3^{\frac{d}{2}} - 1)n$ and $(3 \cdot 2^{\frac{d}{2}} - 2)n$ are exact, respectively.

Optimality of the ICT-construction. The value $\mathrm{ICT}^F_{k_0,r}(x,l)$, where $l = 3n$, is computed by first computing $k_1 = F_{k_0}(r)$ and then returning

$$y := F_{k_0}(x) \| F_{k_1}(x) \| F_{k_1} \circ F_{k_0}(x).$$

For $l = 7n$, an extra key $k_2 = F_{k_1}(r)$ is derived and then the following is returned

$$y \| F_{k_2}(x) \| F_{k_2} \circ F_{k_0}(x) \| F_{k_2} \circ F_{k_1}(x) \| F_{k_2} \circ F_{k_1} \circ F_{k_0}(x).$$

Note that for each n-bit output block, the key indices in the evaluations of F occurs in increasing order. A natural question is whether more can be output before a new key needs to be generated, i.e., can we, using an extra call to F, output more than ICT^F maximally can for a fixed number of generated keys. The answer turns out to be "no" unless the Inverse Decisional Diffie Hellman (IDDH) assumption [2] is false, since (under this assumption) there is a WPRF F, described below, which with high probability both commutes, i.e., $\Pr_{k,k',x}[F_k \circ F_{k'}(x) = F_{k'} \circ F_k(x)] \approx \frac{1}{2}$, and is self inverse, i.e., $\Pr_{k,x}[F_k \circ F_k(x) = x] \approx \frac{1}{2}$. If F is used and more is output at least two output blocks will have the same value with high probability (which is unlikely the case for $\mathcal{R}_{n,*}$). As a consequence ICT is optimal for constructions of this type.

$$F : \{1, \ldots, p\} \times \mathcal{G} \to \mathcal{G} \quad \text{is defined by} \quad F_k(x) := \begin{cases} x^k & \text{if } x \in P_1 \\ x^{k^{-1}} & \text{if } x \in P_2 \end{cases}, \tag{3}$$

where $\mathcal{G}$ is a group of prime order p, k^{-1} satisfies $k \cdot k^{-1} = 1 \pmod{p}$, and P_1 and P_2 are two partitions of roughly equal size.[9] The proof that F is a WPRF if the IDDH assumption holds in $\mathcal{G}$ can easily be derived using results from [2].

4 Applications

By applying our results, we optimize Damgård and Nielsen's CPA-secure encryption scheme (see (2)) and use well-known techniques for achieving CCA-security from any WPRF. We also present a new mode of operation which we think is of theoretical interest.

4.1 Symmetric Encryption

A symmetric encryption scheme $\mathcal{SE} = (E, D)$ consists of two efficient algorithms. The (randomized) encryption algorithm E maps a key k and a message m to a ciphertext $c = E_k(m)$, and the deterministic decryption algorithm D maps a key k and a ciphertext $c = E_k(m)$ to the message $m = D_k(c)$. There are several notions for privacy and integrity of $\mathcal{SE}$ (for an overview, we refer to [6, 15, 3]).

We consider the privacy notion IND-PX-CY for $\mathsf{X},\mathsf{Y} \in \{0, 1, 2\}$ as introduced in [15]. The following is a concrete version of their definition.

[9] In addition it must be efficient to decide whether $x \in P_1$ (or not).

Definition 3 (IND-PX-CY). *Let $\mathcal{M}$ and $\mathcal{K}$ denote the message- and key-space, respectively, of $\mathcal{SE} = (E, D)$. The* **ind**-**p**x-**c**y*-advantage of an adversary A for $\mathcal{SE}$ and $x, y \in \{0, 1, 2\}$ is defined as follows (where $\perp$ denotes no oracle).*

$$\mathbf{Adv}^{\mathbf{ind\text{-}p}x\text{-}\mathbf{c}y}_{\mathcal{SE},A} := 2 \cdot \Pr\left[k \xleftarrow{\$} \mathcal{K}, (m_0, m_1) \leftarrow A^{\mathcal{O}_1, \mathcal{O}_2}, b \xleftarrow{\$} \{0,1\}, \hat{b} \leftarrow A^{\mathcal{O}_1', \mathcal{O}_2'}(E_k(m_b)) : \hat{b} = b\right] - 1,$$

where $(\mathcal{O}_1, \mathcal{O}_1') = \begin{cases} (\perp, \perp) & \text{if } x = 0 \\ (E_k, \perp) & \text{if } x = 1 \\ (E_k, E_k) & \text{if } x = 2 \end{cases}$, $(\mathcal{O}_2, \mathcal{O}_2') = \begin{cases} (\perp, \perp) & \text{if } y = 0 \\ (D_k, \perp) & \text{if } y = 1 \\ (D_k, D_k) & \text{if } y = 2 \end{cases}$,

$m_0, m_1 \in \mathcal{M}$ *with* $|m_0| = |m_1|$, *and A does not query $\mathcal{O}_2'$ with c. Furthermore, let*

$$\mathbf{Adv}^{\mathbf{ind\text{-}p}x\text{-}\mathbf{c}y}_{\mathcal{SE}}(t, q, \mu, q', \mu') := \max_A \{\mathbf{Adv}^{\mathbf{ind\text{-}p}x\text{-}\mathbf{c}y}_{\mathcal{SE},A}\},$$

where the maximum is taken over all A restricted to time-complexity t, at most $q-1$ encryption queries of total length at most $(\mu - |m_0|)$ bits, and q' decryption queries of total length at most μ' bits.[10]

The IND-P2-C2 and IND-P1-C1 notions are often referred to as adaptive IND-CCA and non-adaptive IND-CCA, respectively.[11]

The strongest integrity notion is *integrity of ciphertexts* (INT-CTXT) [6]:

Definition 4 (INT-CTXT). [6] *Let $\mathcal{K}$ denote the keyspace of $\mathcal{SE} = (E, D)$ and D_k^* an algorithm that on input c outputs 1 iff c is a valid ciphertext under the key k. Furthermore, let $x_1, \ldots, x_q$ and $y_1, \ldots, y_{q'}$ denote adversary A's oracle queries to E_k and D_k^*, respectively. Then*

$$\mathbf{Adv}^{\mathbf{int\text{-}ctxt}}_{\mathcal{SE},A} := \Pr\left[k \xleftarrow{\$} \mathcal{K}, A^{E_k, D_k^*}, b := \begin{cases} 1 & \text{If } \exists i\ \forall j : D_k^*(y_i) = 1 \wedge y_i \neq E_k(x_j) \\ 0 & \text{otherwise} \end{cases} : b = 1\right]$$

and $\mathbf{Adv}^{\mathbf{int\text{-}ctxt}}_{\mathcal{SE}}(t, q, \mu, q', \mu') := \max_A \{\mathbf{Adv}^{\mathbf{int\text{-}ctxt}}_{\mathcal{SE},A}\}$ *for all t, q, μ, q', and μ', where the maximum is taken over all A restricted to time-complexity t, at most q queries to E_k of total length at most μ bits, and at most q' queries to D_k^* of total length at most μ' bits.*

4.2 The Damgård-Nielsen Mode of Operation

In [10], Damgård and Nielsen introduced the following mode of operation, for constructing a IND-P2-C0 secure encryption scheme based on any VOL-WPRF

[10] The parameters (q, μ) and (q', μ') are omitted when $x = 0$ and $y = 0$, respectively.

[11] As shown in [15], IND-P1-CY implies IND-P2-CY for $\mathsf{Y} \in \{0, 1, 2\}$. Furthermore, IND-P2-C0 and IND-P2-C2 are equivalent notions to FTG-CPA and FTG-CCA, respectively, and FTG implies the ROR, LOR, and SEM notions [3].

$V : \{0,1\}^\kappa \times \{0,1\}^n \times \mathbb{N} \to \{0,1\}^n$. Let $\mathcal{SE}_1$ denote the symmetric encryption scheme defined by encrypting a message $m \in \{0,1\}^*$, under the key $k \in \{0,1\}^\kappa$ and some auxiliary uniform randomness $r \in \{0,1\}^n$, as

$$(k, r, m) \mapsto \Big(r, V_k(r, |m|) \oplus m\Big). \qquad (\mathcal{SE}_1) \qquad (4)$$

The following proposition, similarly given in [10], states that $\mathcal{SE}_1$ is IND-P2-C0-secure if V is a VOL-WPRF. The proof is provided in Appendix B.

Proposition 1. *For any t, q, and μ*

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1}(t, q, \mu) \leq 2 \cdot \mathbf{Adv}^{\textbf{vol-wprf}}_{V}(t, q, \mu) + \frac{q-1}{2^{n-1}}.$$

We optimize the above scheme by using ICT as VOL-WPRF (as opposed to PRT, cf. (2)), i.e.,

$$(k, r, m) \mapsto \Big(r, \mathrm{ICT}_k(r, |m|) \oplus m\Big). \qquad (\mathcal{SE}'_1) \qquad (5)$$

The security proof follows directly from Proposition 1 and Theorem 2.

Corollary 1. *For any t, q and μ,*

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}'_1}(t, q, \mu) \leq 2 \cdot d_{\max} \cdot \mathbf{Adv}^{\textbf{wprf}}_{F}(t', q2^{d_{\max}-1}) + \frac{4^{d_{\max}} \cdot q^2 + q}{2^n},$$

where $t' = t + \mathcal{O}(\frac{q \cdot l_{\max}}{n} \cdot t_F)$, $d_{\max} := \lfloor \log_2(\lceil l_{\max}/n \rceil) \rfloor + 1$, and $l_{\max} \leq \mu$ is the maximum allowed output length of ICT^F*.*

4.3 Encrypt-then-MAC

A well-known mode of operation for constructing an IND-P2-C2-secure encryption scheme which also assures INT-CTXT, from an IND-P2-C0-secure encryption scheme $\mathcal{SE} = (E, D)$ and a VIL-MAC M, is to simply encrypt with E and then authenticate the ciphertext using M [16, 6].[12] More precisely, let $\mathcal{SE}_2$ be the encryption scheme defined by encrypting a message m, under the key k and k', as

$$(k, r, m) \mapsto \Big(E_k(m), M_{k'}(E_k(m))\Big). \qquad (\mathcal{SE}_2) \qquad (6)$$

The following proposition originates from [6]. It states that if $\mathcal{SE}$ is IND-P2-C0-secure and M is a secure VIL-MAC, then $\mathcal{SE}_2$ is IND-P2-C2-secure and assures INT-CTXT.

[12] The domain of M must contain the ciphertext-space of $\mathcal{SE}$.

Proposition 2 (Encrypt-then-MAC). [6] *For any t, q, q', μ, and μ'*

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_2}(t,q,\mu) \leq \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}}(t,q,\mu),$$

$$\mathbf{Adv}^{\textbf{int-ctxt}}_{\mathcal{SE}_2}(t,q,\mu,q',\mu') \leq q' \cdot \mathbf{Adv}^{\textbf{vil-mac}}_{M}(t,q,\mu+q\delta+\mu'), \text{ and}$$

$$\mathbf{Adv}^{\textbf{ind-p2-c2}}_{\mathcal{SE}_2}(t,q,\mu,q',\mu') \leq q' \cdot \mathbf{Adv}^{\textbf{vil-mac}}_{M}(t,q,\mu+q\delta+\mu') + \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}}(t,q,\mu),$$

where we assume that ciphertexts are δ bits longer than the corresponding plaintexts.

The following standard method for constructing a VIL-MAC M originates in [24] and is used in many other papers (e.g. [22, 7]). The idea is to first hash the message using an ε-almost universal hash function $H : \{0,1\}^{\kappa'} \times \{0,1\}^* \to \{0,1\}^N$ (see definition below) and then apply a PRF $Q : \{0,1\}^{\kappa} \times \{0,1\}^N \to \{0,1\}^n$ to the result. More precisely, $M_{k\|k'}(x) := H_{k'} \circ Q_k(x)$. This method is appealing for several reasons: it is stateless, H exists, and the cryptographic function Q is only invoked on short inputs.

Definition 5. [23] *A hash function family $H : \{0,1\}^{\kappa'} \times \{0,1\}^* \to \{0,1\}^N$ is referred to as an ε-almost universal (ε-AU) if for all distinct $m, m' \in \{0,1\}^*$*

$$\Pr\big[k \leftarrow \mathcal{U}_{\kappa'} : H_k(m) = H_k(m')\big] \leq \varepsilon.$$

We do not describe an ε-AU hash function as it is not within the scope of this paper.[13] However, for completeness we state the above described method concretely dependent on ε.[14]

Proposition 3. [7] *For a function family $Q : \{0,1\}^{\kappa} \times \{0,1\}^N \to \{0,1\}^n$ and an ε-AU hash function $H : \{0,1\}^{\kappa'} \times \{0,1\}^* \to \{0,1\}^N$,*

$$\mathbf{Adv}^{\textbf{vil-mac}}_{H\circ Q}(t,q,\mu) \leq \mathbf{Adv}^{\textbf{prf}}_{Q}(t,q) + \frac{q(q-1)}{2} \cdot \varepsilon + \frac{1}{2^n},$$

where ε typically depends on μ (or on the maximal allowed input length to H).

By combining Proposition 1, 2, and 3, Corollary 1, and Theorem 1, the first efficient encryption scheme based on any WPRF that is secure under a CCA follows:

Corollary 2. *For any t, q, μ, q', μ', ε-AU $H : \{0,1\}^{\kappa'} \times \{0,1\}^* \to \{0,1\}^N$, and* $\mathrm{IC}^F : (\{0,1\}^n)^3 \times \{0,1\}^N \to \{0,1\}^n$*. Let $\mathcal{SE}'_2$ denote the encryption scheme defined by*

[13] There are techniques [24] for constructing a $\frac{2}{2^N}$-AU hash function H, with $\kappa' \leq 4N$, for all $l'_{\max} \leq 2^N$ where $l'_{\max}$ is the maximal allowed length of the input to H.

[14] More generally, one can show that $H \circ Q$ is a VIL-PRF.

encrypting with $\mathcal{SE}_1'$ (cf. (5)) and then authenticating with $H \circ \mathrm{IC}^F$ (cf. (6)), then

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_2'}(t,q,\mu) \leq 2\,d_{\max} \cdot \mathbf{Adv}^{\textbf{wprf}}_F(t', q2^{d_{\max}-1}+1) + \frac{4^{d_{\max}} \cdot q^2 + q}{2^n},$$

$$\mathbf{Adv}^{\textbf{int-ctxt}}_{\mathcal{SE}_2'}(t,q,\mu,q',\mu') \leq q'N \cdot \mathbf{Adv}^{\textbf{wprf}}_F(t,q) + q'\left(N\frac{(q+1)^2}{2^{n+1}} + \frac{q(q-1)}{2}\varepsilon + \frac{1}{2^n}\right),$$

$$\mathbf{Adv}^{\textbf{ind-p2-c2}}_{\mathcal{SE}_2'}(t,q,\mu,q',\mu') \leq \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_2'}(t,q,\mu) + \mathbf{Adv}^{\textbf{int-ctxt}}_{\mathcal{SE}_2'}(t,q,\mu,q',\mu'),$$

where $t' = t + \mathcal{O}(\frac{q \cdot l_{\max}}{n} \cdot t_F)$, $d_{\max} = \lfloor \log_2(\lceil l_{\max}/n \rceil) \rfloor + 1$ and $l_{\max} \leq \mu$ is the maximum allowed output length of ICT^F.

Remark 2. In [10], Damgård and Nielsen also proposed to use the encrypt-then-mac method for achieving CCA-security. However, their approach for constructing the VIL-MAC from the sole assumption of a WPRF introduces a too large overhead for the solution to be practical (the number of applications of the WPRF per evaluation is in the order of l, where l is the length of the message). Our construction of the VIL-MAC, from any n-bit block WPRF, is much more efficient (using at most n applications of the WPRF independently of the message length).

4.4 Encrypt-and-WMAC

In this section, we present a new mode of operation for constructing a IND-P2-C1-secure encryption scheme based on any VOL-WPRF $V : \{0,1\}^\kappa \times \{0,1\}^n \times \mathbb{N} \to \{0,1\}^*$ and a WMAC $W : \{0,1\}^{\kappa'} \times \{0,1\}^n \to \{0,1\}^\ell$. Let $\mathcal{SE}_3$ denote the encryption scheme defined by encrypting a message $m \in \{0,1\}^*$ under the key $(k,k') \in \{0,1\}^\kappa \times \{0,1\}^{\kappa'}$ and some auxiliary uniform randomness $r \in \{0,1\}^n$ as

$$(k,r,m) \mapsto \Big(r, V_k(r,|m|) \oplus m, W_{k'}(r)\Big). \qquad (\mathcal{SE}_3) \qquad (7)$$

Theorem 3. *For any t, q, μ, q', and μ'*

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_3}(t,q,\mu) \leq \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1}(t,q,\mu), \text{ and}$$

$$\mathbf{Adv}^{\textbf{ind-p2-c1}}_{\mathcal{SE}_3}(t,q,\mu,q',\mu') \leq q' \cdot \mathbf{Adv}^{\textbf{wmac}}_W(t,q) + \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1}(t,q,\mu).$$

The proof is given in Appendix B.

Remark 3. Recall that if $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ is a WPRF, then IC^F is as PRF and hence also a (W)MAC. Therefore, by using IC^F as WMAC and ICT^F as VOL-WPRF in (7) results in a IND-P2-C1 secure encryption scheme from any WPRF. Moreover, since IC^F is invoked on random inputs, $n/2$ invocations of F will be needed for each encryption on average. It is an open question whether even more efficient constructions of WMACs from any WPRF exists.

References

1. W. Aiello, S. Rajagopalan, and R. Venkatesan. High-speed pseudorandom number generation with small memory. In *Fast Software Encryption*, volume 1636 of *LNCS*, pages 290–304. Springer, 1999.
2. F. Bao, R. H. Deng, and H. Zhu. Variations of Diffie-Hellman problem. In *ICICS '03*, volume 2836 of *LNCS*, pages 301–312. Springer, 2003.
3. M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A concrete security treatment of symmetric encryption. In *Proc. of the 38th Symposium on Foundations of Computer Science*, pages 394–403. IEEE, 1997.
4. M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway. Relations among notions of security for public-key encryption schemes. In *Advances in Cryptology - CRYPTO '98*, volume 1462 of *LNCS*, pages 26–45. Springer, 1998.
5. M. Bellare, J. Kilian, and P. Rogaway. The security of cipher block chaining. In *Advances of Cryptology - CRYPTO '94*, volume 839 of *LNCS*, pages 341–358. Springer, 1994.
6. M. Bellare and C. Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In *Advances in Cryptology - ASIACRYPT '00*, volume 1976 of *LNCS*, pages 531–545. Springer, 2000.
7. J. Black, S. Halevi, H. Krawczyk, T. Krovetz, and P. Rogaway. Umac: Fast and secure message authentication. In *Advances in Cryptology - CRYPTO '99*, volume 1666 of *LNCS*, pages 313–328. Springer, 1999.
8. Avrim Blum, Merrick L. Furst, Michael J. Kearns, and Richard J. Lipton. Cryptographic primitives based on hard learning problems. In *Advances of Cryptology - CRYPTO '93*, volume 773 of *LNCS*, pages 278–291. Springer, 1993.
9. M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM J. Comput.*, 13(4):850–864, 1984.
10. I. Damgård and J. B. Nielsen. Expanding pseudorandom functions; or: from known-plaintext security to chosen-plaintext security. In *Advances of Cryptology - CRYPTO '02*, volume 2442 of *LNCS*, pages 449–464. Springer, 2002.
11. W. Diffie and M. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.
12. O. Goldreich. *Foundations of Cryptography*, volume II Basic Applications. Cambridge University Press, 2004.
13. O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *J. of the ACM*, 33(4):792–807, 1986.
14. J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.
15. J. Katz and M. Yung. Complete characterization of security notions for probabilistic private-key encryption. In *Proc. of the 32nd Annual Symposium on Theory of Computing*, pages 245–254. ACM, 2000.
16. S. Kent and R. Atkinson. IP encapsulating security payload (ESP). Request for Comments 2406, November 1998.
17. U. Maurer, Y. A. Oswald, K. Pietrzak, and J. Sjödin. Luby-rackoff ciphers with weak round functions. In *Advances in Cryptology - EUROCRYPT '06*, LNCS. Springer, 2006.
18. K. Minematsu and Y. Tsunoo. Expanding weak PRF with small key size. In *ICISC '05*, LNCS. Springer, 2005.
19. M. Naor and O. Reingold. From unpredictability to indistinguishability: A simple construction of pseudo-random functions from MACs. In *Advances of Cryptology - CRYPTO '98*, volume 1462 of *LNCS*, pages 267–282. Springer, 1998.

20. M. Naor and O. Reingold. Synthesizers and their application to the parallel construction of pseudo-random functions. *J. Comput. Syst. Sci*, 58(2):336–375, 1999.
21. M. Naor and O. Reingold. Number-theoretic constructions of efficient pseudo-random functions. *J. of the ACM*, 51(2):231–262, 2004.
22. V. Shoup. On fast and provably secure message authentication based on universal hashing. In *Advances in Cryptology - CRYPTO '96*, LNCS, pages 313–328. Springer, 1996.
23. D. R. Stinson. Universal hashing and authentication codes. In *Advances of Cryptology - CRYPTO '91*, volume 576 of *LNCS*, pages 74–85. Springer, 1992.
24. M. N. Wegman and J. L. Carter. New hash functions and their use in authentication and set equality. *J. of Computer and System Sciences*, 22:265–279, 1981.
25. A. C. Yao. Theory and applications of trapdoor functions. In *Proc. of the 23rd Annual IEEE Symposium on Foundations of Computer Science*, pages 80–91, 1982.

A Proof of Theorem 1

Proof (of Theorem 1). Let $\mathit{\Pi}_0$ denote the following game

$$(k,r,r') \overset{\$}{\leftarrow} \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^n,\ b \leftarrow A^{\mathrm{IC}^F_{k,r,r'}}.$$

For $j = \{1,\dots,N\}$, let $\mathit{\Pi}_{2j-1}$ be the same game as $\mathit{\Pi}_{2j-2}$, except that F_{k_j} is replaced by a random function, and let $\mathit{\Pi}_{2j}$ be the same game as $\mathit{\Pi}_{2j-1}$, except that k_j is replaced by a random string and for each query x to IC^F the intermediate value τ_j is replaced by a random string unless $x[1,j]$ is not a prefix of an earlier query. Finally, note that $\mathit{\Pi}_{2N}$ is equivalent to

$$R \leftarrow \mathcal{R}_{N,n},\ b \leftarrow A^R.$$

Let S_j denote the event that $b = 1$ in $\mathit{\Pi}_j$ (for $j \in \{0,\dots,2N\}$). Then

$$\begin{aligned}\mathbf{Adv}^{\mathbf{prf}}_{\mathrm{IC}^F}(t,q) := \Pr[S_0]-\Pr[S_{2N}] &= \sum_{j=1}^{N} (\Pr[S_{2j-2}]-\Pr[S_{2j-1}]) + (\Pr[S_{2j-1}]-\Pr[S_{2j}]) \\ &\le \sum_{j=1}^{N} \mathbf{Adv}^{\mathbf{wprf}}_F\big(t, \min\{q+1, 2^{j-1}+1\}\big) + \min\left\{\frac{(q+1)q}{2^{n+1}}, \frac{(2^{j-1}+1)2^{j-1}}{2^{n+1}}\right\} \\ &\le N \cdot \left(\mathbf{Adv}^{\mathbf{wprf}}_F(t, q+1) + \frac{q(q+1)}{2^{n+1}}\right),\end{aligned}$$

where the first inequality follows from the triangle inequality, the easily verified fact that there is a distinguisher (restricted to time-complexity t and most $\min(q+1, 2^{j-1}+1)$ oracle invocations) with **wprf**-advantage at least $\Pr[S_{2j-2}] - \Pr[S_{2j-1}]$, the fact that $\mathit{\Pi}_{2j-1}$ and $\mathit{\Pi}_{2j}$ (for all $j \in \{1,\dots,N\}$) are equivalent games as long as r and the assignments of the variable τ_{j-1} in the algorithm are all distinct. Since these values are at most $\min\{q+1, 2^{j-1}+1\}$ and completely random, the probability of this event is upper bounded by $\min\left\{\frac{(q+1)q}{2^{n+1}}, \frac{(2^{j-1}+1)2^{j-1}}{2^{n+1}}\right\}$. □

B Proof of Proposition 1 and Theorem 3

Proof (of Proposition 1). Let E and D denote the encryption and decryption algorithm of $\mathcal{SE}_1$, respectively. Recall that $E_k(m) := (r, V_k(r, |m|) \oplus m)$, where V is a VOL function family. Let $\mathit{\Pi}_0$ denote the original IND-P2-C0 game, i.e.,

$$k \xleftarrow{\$} \{0,1\}^\kappa; (x_0, x_1) \leftarrow A^{E_k}; b \xleftarrow{\$} \{0,1\}; y \leftarrow E_k(x_b); \hat{b} \leftarrow A^{E_k}(y).$$

Furthermore, let $\mathit{\Pi}_1$ be the same game as $\mathit{\Pi}_0$, except that V_k is replaced by $\mathcal{R}_{\mathrm{n},*}$. Let $\mathit{\Pi}_2$ be the same game as $\mathit{\Pi}_1$, except that the input y to the adversary is replaced by a truly random string y' of the same length. For $i \in \{0,1,2\}$ let S_i denote the event that A outputs b, i.e., $\hat{b} = b$, in game $\mathit{\Pi}_i$. It follows that

$$\begin{aligned}\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1}(t,q,\mu) &:= 2 \cdot \Pr[S_0] - 1 = 2\left(\Pr[S_2] + \sum_{i=0}^{1} \Pr[S_i] - \Pr[S_{i+1}]\right) - 1 \\ &\leq 2 \cdot \left(\frac{1}{2} + \mathbf{Adv}^{\textbf{vol-wprf}}_{V}(t,q,\mu) + \frac{q-1}{2^n}\right) - 1,\end{aligned}$$

where the inequality follows from the triangle inequality, the fact that a distinguisher for V with **wprf**-advantage at least $\Pr[S_0] - \Pr[S_1]$ using resources t, q, and μ can be constructed,[15] the fact that $\mathit{\Pi}_1$ and $\mathit{\Pi}_2$ are equivalent games as long as the input to $\mathcal{R}_{\mathrm{n},*}$ in the computation of y is different from the other inputs to $\mathcal{R}_{\mathrm{n},*}$, and that $\Pr[S_2] = 1/2$ since b is independent of y'. □

Proof (of Theorem 3). Let E and D denote the encryption and decryption algorithm of $\mathcal{SE}_3$, respectively. Recall that $E_{k\|k'}(m) := (r, V_k(r, |m|) \oplus m, W_{k'}(r))$, where V is a VOL-WPRF and W a WMAC. We prove the second inequality (the proof of the first inequality is straight forward and therefore omitted). Let $\mathit{\Pi}_0$ denote the original IND-P2-C1 game, i.e.,

$$\begin{aligned}&(k,k') \xleftarrow{\$} \{0,1\}^\kappa \times \{0,1\}^{\kappa'}, \\ &(x_0,x_1) \leftarrow A^{E_{k\|k'}, D_{k\|k'}}, \\ &b \xleftarrow{\$} \{0,1\}, y \leftarrow E_{k\|k'}(x_b), \\ &\hat{b} \leftarrow A^{E_{k\|k'}}(y).\end{aligned}$$

Furthermore, let $\mathit{\Pi}_1$ be the same game as $\mathit{\Pi}_0$, except that all decryption queries, for which the auxiliary random part r is distinct from the auxiliary random parts $r_1, r_2, \ldots$

[15] The distinguisher simply runs the adversary A, which maximizes $\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1,A}$ with resources (t,q,μ), answering its oracle queries with help of its own oracle, and outputs 1 if A is successful, i.e., when $b = \hat{b}$, and else 0.

of the ciphertexts received from the encryption oracle, are rejected. Furthermore, let S_i, for $i \in \{0, 1\}$, denote the event that A outputs b, i.e., $\hat{b} = b$, in $\mathit{\Pi}_i$. It follows that

$$\begin{aligned}\mathbf{Adv}^{\textbf{ind-p2-c1}}_{\mathcal{SE}_3}(t, q, \mu, q', \mu') &:= \Pr[S_0] = \Big(\Pr[S_0] - \Pr[S_1]\Big) + \Pr[S_1] \\ &\leq q' \cdot \mathbf{Adv}^{\textbf{wmac}}_{W}(t, q) + \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_2}(t, q, \mu).\end{aligned}$$

The inequality follows from the triangle inequality, the fact that $\mathit{\Pi}_0$ and $\mathit{\Pi}_1$ are equivalent games unless the MAC is forged (using q' tries), and that the decryption queries do not help adversary in $\mathit{\Pi}_1$ since the adversary can simulate the decryption oracle itself. It remains to show that

$$\mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_3}(t, q, \mu) \leq \mathbf{Adv}^{\textbf{ind-p2-c0}}_{\mathcal{SE}_1}(t, q, \mu),$$

We show this by showing that any IND-P2-C0 adversary A for $\mathcal{SE}_3 = (E, D)$ can be transformed into an IND-P2-C0 adversary A' for $\mathcal{SE}_1 = (E', D')$ using the same resources. A' chooses a key k' uniformly at random from the keyspace of W and simply runs A answering its oracle queries $m_1, \ldots, m_q$ with help of its own oracle E_k and $W_{k'}$ (for each query m_i, A' simply invokes $E_k(\cdot)$ on input m_i and receives (r_i, c_i). It then returns $(r, c_i, W_{k'}(r))$ to A). A' transforms the challenge ciphertext the same way before returning it to A. Finally, A' decides the same way as A. □

C Proof of Theorem 2 (and two PRGs from any WPRF)

For any function $f : \{0,1\}^n \to \{0,1\}^n$ and bitstring $\mathbf{x} = x_1 \| \cdots \| x_s \in \{0,1\}^{sn}$, where $x_1, \ldots, x_s$ are n-bit blocks, let $f(\mathbf{x}) := f(x_1) \| \cdots \| f(x_q)$. If $\mathcal{D}_1$ and $\mathcal{D}_2$ are two probability distributions over S, then let the distinguishing advantage of a distinguisher A be defined as

$$\mathbf{Adv}^{\mathcal{D}_1, \mathcal{D}_2}_{A} := \Pr[s \leftarrow \mathcal{D}_1, b \leftarrow A(s) : b = 1] - \Pr[s \leftarrow \mathcal{D}_2, b \leftarrow A(s) : b = 1],$$

and the maximal distinguishing advantage as

$$\mathbf{Adv}^{\mathcal{D}_1, \mathcal{D}_2}(t) := \max_{A}\{\mathbf{Adv}^{\mathcal{D}_1, \mathcal{D}_2}_{A}\},$$

where the maximum is taken over all D with time-complexity t.

Before we present the proof of Theorem 2, let us consider the following two constructions of PRGs from any WPRF.

For all $0 < s,\ d \in \mathbb{N}$, let

$$\begin{aligned} G^F_s &: \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^{sn} \to \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^{2sn} \\ &(k, r, \mathbf{x}) \mapsto (F_k(r), r, \mathbf{x} \| F_k(\mathbf{x})) \end{aligned} \tag{8}$$

and

$$\begin{aligned} G^F_{s,d} &: \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^{sn} \to \{0,1\}^n \times \{0,1\}^n \times \{0,1\}^{2^d sn} \\ &(k, r, \mathbf{x}) \mapsto G^F_{s2^0} \circ \cdots \circ G^F_{s2^{d-1}} \circ G^F_{s2^{(d-2)}}(k, r, \mathbf{x}). \end{aligned} \tag{9}$$

The following two lemmata state that the above constructions are PRGs if F is a WPRF.

Lemma 1. *For any t and $s > 0$*

$$\mathbf{Adv}^{\mathbf{prg}}_{G^F_s}(t) \le \mathbf{Adv}^{\mathbf{wprf}}_F(t, s+1) + \frac{s(s+1)}{2^{n+1}}.$$

Proof. Consider the following three distributions:

$$\mathcal{D}_1 := F_k(r) \,\|\, r \,\|\, \mathbf{x} \,\|\, F_k(\mathbf{x}), \quad \mathcal{D}_2 := R(r) \,\|\, r \,\|\, \mathbf{x} \,\|\, R(\mathbf{x}), \quad \text{and} \quad \mathcal{D}_3 := \mathcal{U}_{2(s+1)n},$$

where $r\|\mathbf{x} \leftarrow \mathcal{U}_{(s+1)n}$ and $R \leftarrow \mathcal{R}_{n,n}$. Then

$$\begin{aligned} \mathbf{Adv}^{\mathbf{prg}}_{G^F_s}(t) &= \mathbf{Adv}^{\mathcal{D}_1,\mathcal{D}_3}(t) \le \mathbf{Adv}^{\mathcal{D}_1,\mathcal{D}_2}(t) + \mathbf{Adv}^{\mathcal{D}_2,\mathcal{D}_3}(t) \\ &\le \mathbf{Adv}^{\mathbf{wprf}}_F(t, s+1) + \frac{s(s+1)}{2^{n+1}}, \end{aligned}$$

which follows directly from the triangle inequality, the fact that $\mathcal{D}_2$ and $\mathcal{D}_3$ are the same distributions as long as the values $r, x_1, \ldots, x_s$ are all distinct, and the trivial fact that a distinguisher for distinguishing $\mathcal{D}_1$ from $\mathcal{D}_2$ with time-complexity t can directly be transformed into a **wprf**-adversary for F with the same advantage and time-complexity, and at most $s+1$ oracle invocations. □

Lemma 2. *For any t and $s, d > 0$*

$$\mathbf{Adv}^{\mathbf{prg}}_{G^F_{s,d}}(t) \le d \cdot \mathbf{Adv}^{\mathbf{wprf}}_F(t, s2^{d-1}+1) + \frac{s^2 \cdot (4^d - 1)}{2^{n+1}}.$$

Proof. The proof follows from a simple hybrid argument and Lemma 1.

$$\begin{aligned} \mathbf{Adv}^{\mathbf{prg}}_{G^F_{s,d}}(t) &\le \sum_{i=0}^{d-1} \mathbf{Adv}^{\mathbf{prg}}_{G^F_{s2^i}}(t) \le \sum_{i=0}^{d-1} \left(\mathbf{Adv}^{\mathbf{wprf}}_F(t, s2^i + 1) + \frac{s2^i\,(s2^i+1)}{2^{n+1}} \right) \\ &\le d \cdot \mathbf{Adv}^{\mathbf{wprf}}_F(t, s2^{d-1}+1) + \frac{1}{2^{n+1}} \underbrace{\left(\frac{s^2\,(4^d-1)}{3} + s \cdot (2^d - 1) \right)}_{\le s^2 \cdot (4^d - 1)}. \end{aligned}$$

□

Proof (of Theorem 2). As a consequence (to the above) the ICT-construction can similarly be defined as follows. For any $l \in \mathbb{N}$, let $d = \lfloor \log_2(\lceil l/n \rceil) \rfloor + 1$ and

$$\begin{aligned} \mathrm{ICT}^F : \{0,1\}^{2n} \times \{0,1\}^n \times \mathbb{N} &\to \{0,1\}^* \\ (k\|r, x, l) &\mapsto \tilde{G}^F_{1,d}(k, r, x)[1, l], \end{aligned}$$

where $\tilde{G}^F_{s,d}(k, r, \mathbf{x})$ is defined by $G^F_{s,d}(k, r, \mathbf{x}) = (k', r, \mathbf{x} \| \tilde{G}^F_{s,d}(k, r, \mathbf{x}))$, i.e., the first two output-components of $G^F_{s,d}(k, r, \mathbf{x})$ and $\mathbf{x}$ are not an output.

Let $r, k, x_1, \ldots, x_q \leftarrow \mathcal{U}_n$, $l_i \in \mathbb{N}$ the length of the i^{th} output, $l_{\max}$ the maximal allowed output length of ICT^F, and $d_{\max} := \lfloor \log_2(\lceil l_{\max}/n \rceil) \rfloor + 1$. For $R \leftarrow \mathcal{R}_{n,*}$, consider the following distributions where $A \sqsubseteq B$ indicates that distribution A can be sampled with help of B, by first sampling from B and then removing and reranging bits of the sample.

$$\begin{aligned} \mathcal{D}_1 &:= x_1 \| \cdots \| x_q \| \, R(x_1, l_1) && \| \cdots \| \, R(x_q, l_q) \\ \mathcal{D}_2 &:= x_1 \| \cdots \| x_q \| \, \mathrm{ICT}^F(k\|r, x_1, l_1) && \| \cdots \| \, \mathrm{ICT}^F(k\|r, x_q, l_q) \\ &\;= x_1 \| \cdots \| x_q \| \, \tilde{G}^F_{1,\, d_{\max}}(k, r, x_1)[1, l_1] \, \| \cdots \| \, \tilde{G}^F_{1,\, d_{\max}}(k, r, x_q)[1, l_q] \\ &\sqsubseteq x_1 \| \cdots \| x_q \| \, \tilde{G}^F_{q, d_{\max}}(k, r, x_1 \| \cdots \| x_q) \sqsubseteq G^F_{q, d_{\max}}(k, r, \mathbf{x}) =: \mathcal{D}_2' \\ \mathcal{D}_3 &:= \mathcal{U}_{qn + l_1 + \cdots + l_q} && \sqsubseteq \mathcal{U}_{(2 + q 2^{d_{\max}}) \cdot n} \quad =: \mathcal{D}_3' \,. \end{aligned}$$

It is straight forward, using Lemma 2, to verify that

$$\begin{aligned} \mathbf{Adv}^{\textbf{vol-wprf}}_{\mathrm{ICT}^F}(t, q, \mu) &\le \mathbf{Adv}^{\mathcal{D}_1, \mathcal{D}_3}(t') + \mathbf{Adv}^{\mathcal{D}_2', \mathcal{D}_3'}(t') \le \frac{q(q-1)}{2^{n+1}} + \mathbf{Adv}^{\textbf{prg}}_{G^F_{q, d_{\max}}}(t') \\ &\le \frac{q(q-1)}{2^{n+1}} + d \cdot \mathbf{Adv}^{\textbf{wprf}}_F(t', s 2^{d_{\max}-1} + 1) + \frac{q^2 \cdot (4^{d_{\max}} - 1)}{2^{n+1}}, \end{aligned}$$

where $t' = t + \mathcal{O}(\frac{q \cdot l_{\max}}{n} \cdot t_F)$.[16] □

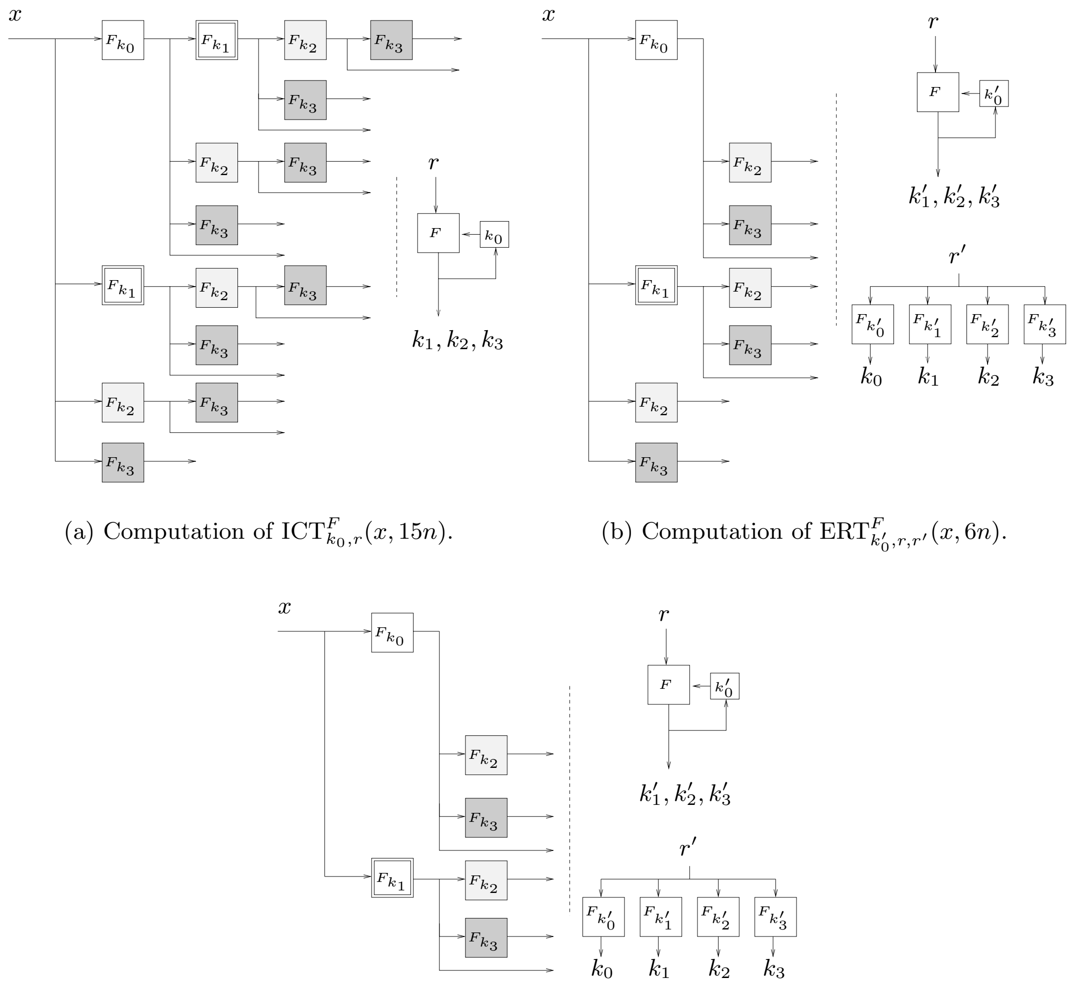
Remark 4 (Reduction of key material). All the keys of our constructions (in this paper) can be generated from a WPRF $F : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$, two publicly random domain points (i.e., two n-bit strings r and r'), and a secret key k of F. To generate t (n-bit) keys $k_1, \ldots, k_t$ one simply computes $k_1 \| \ldots \| k_t := \mathrm{ICT}^F_{k\|r}(r', tn)$. This follows from the fact that

$$r \| r' \| \, \mathrm{ICT}^F_{k\|r}(r', tn) = G^F_{1, \lceil \log_2(t) \rceil}(k, r, r')[n{+}1, (t{+}3)n]$$

which is pseudo-random according to Lemma 2. It can easily be shown that this key generation method is more efficient and uses a shorter key than the method presented in [10].

[16] The worst-case running time for sampling $\mathcal{D}_1$, $\mathcal{D}_2'$, or $\mathcal{D}_3'$ is in $\mathcal{O}(\frac{q \cdot l_{\max}}{n} \cdot t_F)$.

D Figures



(a) Computation of $\text{ICT}^F_{k_0,r}(x, 15n)$.

(b) Computation of $\text{ERT}^F_{k'_0,r,r'}(x, 6n)$.

(c) Computation of $\text{PRT}^F_{k'_0,r,r'}(x, 4n)$.

Fig. 1. The following figure illustrates, (a) the ICT-construction of this paper, (b) the ERT-construction [18], and (c) the PRT-construction [10]. Each figure corresponds to the computation of the output of maximal size using 4 (generated) keys. The key generation is illustrated to the right of the dashed line in each figure. Interestingly, the generated key sequence, i.e., $k_0, \ldots, k_3$, is not pseudo-random in (a) as opposed to (b) and (c) (cf. Remark 1).