Randomness-Dependent Message Security

Eleanor Birrell Kai-Min Chung Rafael Pass Sidharth Telang

December 28, 2012

Abstract

Traditional definitions of the security of encryption schemes assume that the messages encrypted are chosen independently of the randomness used by the encryption scheme. Recent works, implicitly by Myers and Shelat (FOCS'09) and Bellare et al (AsiaCrypt'09), and explicitly by Hemmenway and Ostrovsky (ECCC'10), consider *randomness-dependent message (RDM) security* of encryption schemes, where the message to be encrypted may be selected as a function—referred to as the RDM function—of the randomness used to encrypt this particular message, or other messages, but in a circular way. We carry out a systematic study of this notion.

Our main results demonstrate the following:

- *Full RDM security*—where the RDM function may be an arbitrary polynomial-size circuit—is not possible.
- Any secure encryption scheme can be slightly modified, by just performing some *pre-processing to the randomness*, to satisfy *bounded-RDM* security, where the RDM function is restricted to be a circuit of *a priori* bounded polynomial size. The scheme, however, requires the randomness r needed to encrypt a message m to be slightly longer than the length of m (i.e., $|r| > |m| + \omega(\log k)$ where k is the security parameter).
- We present a black-box provability barriers to compilations of *arbitrary* public-key encryption into RDM-secure ones using just pre-processing of the randomness whenever $|m| > |r| + \omega(\log k)$. On the other hand, under the DDH assumption, we demonstrate the existence of bounded-RDM secure schemes that can encrypt arbitrarily "long" messages using "short" randomness.

We finally note that the existence of public-key encryption schemes imply the existence of a fully RDM-secure encryption scheme in an "ultra-weak" Random-Oracle Model—where the security reduction need not "program" the oracle, or see the queries made by the adversary to the oracle; combined with our impossibility result, this yields the first example of a cryptographic task that has a secure implementation in such a weak Random-Oracle Model, but does not have a secure implementation without random oracles.

Our constructions of RDM secure encryption scheme borrow techniques from Hemmenway and Ostrovksy, and Bellare et al, but our analyses are different. In particular, to analyze our schemes, we develop several new tools regarding t-wise independent hash function, mirroring deterministic extraction lemmas for computationally bounded sources by Trevisan and Vadhan (STOC'00), that may be of independent interest.

1 Introduction

Traditional definitions of secure encryption, including semantic (or CPA) security and CCA security, address the problem of how to securely communicate a message in the presence of a polynomially-bounded adversary that observes encrypted messages. In the standard approach, it is assumed that the message, the keys, and the randomness used to encrypt the message, are all chosen independently.

More recently, new definitions have emerged that relax some of these independence assumptions. Most notably, a line of work initiated independently by Camenisch and Lysyanskaya [20] and by Black, Rogaway, and Shrimpton [13] addresses the problem of "key-dependent" messages (KDM): namely, they consider the security of a public-key encryption scheme in a setting where the message to be encrypted may (adversarially) depend on the secret-key. A variant of this notion instead considers "circular" security: here, the adversary may observe a "cycle" of q messages $\overrightarrow{m}$ encrypted using different keys $(\overrightarrow{pk}, \overrightarrow{sk})$, but where m_i may depend on the depends on the secret-key $sk_{(i+1 \mod q)}$. One motivation for studying key-dependence arises in the context of hard-drive encryption: you want to encrypt your hard-drive, on which your secret-key is also found. Circular security arises naturally in a situation when two parties want to share their secret keys with each other (but not with the rest of the world): a natural solution to the problem would be for player 1 to send an encrypted version of his secret key using player 2's public key, and vice versa. For this protocol to be secure, circular security is needed. More recently, circular security has found important applications in the context of fully-homomorphic encryptions (indeed, to date, all known FHE schemes rely on the assumption that some underlying encryption scheme is circularly secure).

We here focus on an alternative relaxation of the classic independence assumptions, first implicitly considered by Myers and Shelat [38] and Bellare et al [10], and explicitly by Hemmenway and Ostrovsky [33]: We study of the security of encryption schemes in a scenario where the message to be encrypted may be selected as a function—referred to as the *RDM function*—of the randomness used to encrypt this particular message, or other messages, but in a circular way. More precisely, in analogy with KDM security and circular security, we consider two notions of randomness dependent message security.

- *Randomness-dependent message (RDM) security*: roughly speaking, a public-key encryption scheme is said to be RDM-secure if indistinguishability of ciphertexts holds even if the encrypted messages are chosen as a function of the randomness used to encrypt *this particular* message.

- *Circular randomness-dependent (circular-RDM) security*: roughly speaking, a public-key encryption scheme is said to be circular RDM-secure if indistinguishability of ciphertexts holds even if the encrypted messages are chosen as a function of the randomness used to encrypt *other* messages, but in a circular way. More precisely, we consider a scenario where q messages $\overrightarrow{m}$ are encrypted using randomness $\overrightarrow{r}$, where m_1 is chosen as a function of r_q and each other message m_i is chosen as a function of r_{i-1} and the "previous" ciphertext $c_{i-1} = \mathrm{Enc}_{pk}(m_{i-1}, r_{i-1})$.

Why care about Randomness-Dependent Message Security We consider two reasons to study RDM security:

1. *involuntary RDM attacks:* Implementations of secure protocols are prone to programming mistakes; attacks exploiting such programming mistakes (e.g., buffer overflow attacks) have

been demonstrated on secure protocols. Attacks of this type may allow an attacker to see encryptions of randomness dependent messages, even if the original protocol chooses messages independently of the randomness used to encrypt it. RDM security would block such "involuntary" RDM attacks.

To prevent against these we need to be able to handle sufficiently general classes of RDM functions that may be produced by the attackers.

2. *voluntary RDM attacks* As shown in the beautiful work by Myers and Shelat [38], the possibility of encrypting the randomness used in other encryptions, in a circular way, leads to new powerful techniques in the design of encryption schemes. This techniques was further refined in a recent work by Hohenberger, Lewko and Waters [35]. Another application is found in the work of Hemmenway and Ostrovsky [33], that explicitly considers a notion of circular randomness dependent "one-wayness" and show its usefulness for constructing injective trapdoor functions. In this context, the protocol designer is "voluntarily" creating a (circular-)RDM attack. The above-mentioned works either implicitly (as in [38] and [35]), or explicitly (as in [33]) consider and design encryption schemes that are circular-RDM secure for the specific randomness-dependent messages selected by their protocols. Although for this particular application it suffices to consider specific RDM functions, having general-purpose RDM-secure encryption schemes simplifies the design and the security analysis of protocols.

   Another motivation stems from non-black-box simulation techniques pioneered in the work by Barak [5]; in a variant of Barak's simulation technique due to [43], the simulator commits to its own code (that, in particular, contains the randomness used for the commitment, and thus circularity arises). In this particular application, the circularity could be broken, but having general techniques for dealing with RDM security may simplify future applications.

Before explaining our result, let us also point out that RDM secure encryption is very related to *hedged encryption schemes* introduced by Bellare et al [10]—encryption schemes that remain secure as long as the joint message-randomness distribution comes from a high-entropy source, that is independent of the public-key of the encryption scheme (which in turn are very related to *deterministic encryption* [8, 11, 14]; see [10] for more details). Hedged encryption schemes are RDM-secure if restricting the attacker to using RDM functions that do not depend on the public-key.[1] Our focus here is on notions of RDM security where the RDM function may depend also on the public-key.

1.1 Our results

Full RDM security Our first result shows that if the RDM function may be an arbitrary polynomial-size circuit (chosen by the adversary), then RDM security, as defined by Hemmenway and Ostrovsky [33], is impossible to achieve.

Theorem 1.1 (Informal Statement)**.** *There does not exist an encryption scheme that is (fully) RDM-secure.*

We next show that if there exists some polynomial q such that an encryption scheme is q-circular RDM secure, then the encryption scheme is also RDM secure; thus q-circular RDM security is impossible for all polynomials q.

[1] However, it is not clear in general whether hedged encryption schemes are *circular* RDM secure, even if we restrict to RDM functions that do not depend on the public-key.

Theorem 1.2 (Informal Statement)**.** *There does not exist an encryption scheme that is (fully) q-circular RDM-secure for any polynomial q.*

Bounded RDM security Since "unbounded" RDM security is impossible, we consider RDM security with respect to restricted classes of RDM functions.

Our first positive result demonstrates that if the RDM function is restricted to be a circuit of *a priori* polynomially bounded size, then any secure encryption scheme can be modified to satisfy both RDM and circular-RDM security.

Theorem 1.3 (Informal Statement)**.** *Assume the existence of a secure public key encryption scheme. Then, for every polynomial s, there exists an encryption scheme Π that is RDM secure when restricting the RDM function to be computed by a circuit of size at most $s(k)$ where k is the security parameter. Additionally Π is q-circular RDM secure for every polynomial q under the same restrictions on the RDM function.*

Theorem 1.3 is proven by modifying any secure encryption scheme to first "hash" the randomness using a t-wise independent hash-function. The same transformation was previously used by Hemmenway and Ostrovsky [33] to transform "lossy encryption schemes" [44], that can encrypt messages longer than the randomness, into schemes that satisfy a notion of circular-RDM "one-wayness"[2] (as opposed to semantic security) with respect to a particular circular-RDM function (the identity function).

In order to encrypt a message m, our encryption scheme requires using $|m| + \omega(\log k)$ bits; that is, the randomness used to encrypt a message needs to be sufficiently longer than the message being encrypted (as such, the encryption scheme of Theorem 1.3 does not handle "the identity function" as an RDM function.) Our next positive result strictly strengthens the conclusion of Theorem 1.3 (but under a stronger assumption) and the results of [33]: the existence of *lossy trapdoor functions* [45] implies the existence of both bounded RDM-secure and bounded circular-RDM secure encryption schemes that can encrypt also "long" messages using "short" randomness—the ratio between the message-length and the randomness length is proportional to the lossiness of the trapdoor function. Our construction mirrors a construction of hedged encryption of Bellare et al [10]; roughly, the encryption is done by first "hashing" the message-randomness pair and then applying a lossy trapdoor function to the hashed value. The key difference is that we replace the use of univeral hashing (in the construction of [10]) with t-wise independent hashing.[3]

Theorem 1.4 (Informal Statement)**.** *Assume the existence of "sufficiently" lossy trapdoor functions (the existence of which are implied e.g., by the DDH assumption). Then, for every polynomials s, l, there exists a $l(k)$-bit encryption scheme Π using only k-bits of randomness that is RDM secure (and q-circular RDM secure for every polynomial q), when restricting the RDM function to be computed by a circuit of size at most $s(k)$ where k is the security parameter.*

To prove the above two theorems we develop several new information-theoretic tools regarding t-wise independent hash functions, that may be of independent interests. For instance, with very high probability, a t-wise independent hash functions is a "good" randomness extractor for any min-entropy source with *with computationally-bounded leakage* (mirroring a lemma of Trevisan-Vadhan

[2]The notion of q-circular RDM one-wayness of Hemmenway and Ostrovsky requires that no polynomial-time attacker can recover $r_1, r_2, \ldots r_q$ given $\mathrm{Enc}_{pk}(r_q; r_1), \mathrm{Enc}_{pk}(r_1; r_2), \ldots, \mathrm{Enc}_{pk}(r_{q-1}; r_q)$ except with negligible probability, over the choice of pk and uniform $r_1, \ldots r_q$.

[3]The construction of [10] actually requires universal hash *permutations*. As far as we know, constructions of t-wise independent permutations are not known, which requires us to further modify the scheme to guarantee correctness.

[46]). We also present "crooked" versions of such deterministic extraction lemmas (mirroring the "crooked left-over-hash lemma of [24]).

An interesting question is whether any encryption schemes can be modified by simply performing some pre-processing to the randomness (as in Theorem 1.3) to become bounded RDM secure, but still handle long messages using short randomness. At first sight, it may seem like we could use a pseudorandom generator to "stretch" a small seed into the required long random string for the construction in Theorem 1.3. We have no attack against this construction. However, we show that security reductions that only use the attacker and the RDM function as a black-box—following [29], we refer to such reductions as *strongly black-box*—cannot be used to demonstrate RDM security of encryption schemes with *perfect correctness* and *efficiently recognizable public-keys* that can encrypt long messages using short randomness, based on a falsifiable intractability assumption [39]; for instance, this means that the El-Gamal crypto system cannot be modified (by performing pre-processing to the randomness) to become bounded RDM secure for long messages.

Theorem 1.5 (Informal statement)**.** *Assume the existence of one-way functions secure against subexponential-sized circuits. For every polynomials m and r such that $m(k) \geq r(k)+\omega(\log k)$, there exists a polynomial s such that for every $m(\cdot)$-bit encryption scheme Π with perfect correctness and efficiently recognizable public-keys that uses $r(\cdot)$ bits of randomness to encrypt a message, s-bounded security of Π cannot be based on any falsifiable assumption using a strongly black-box reduction, unless the assumption is false.*

Let us point out that the reason Theorem 1.5 does not contradict Theorem 1.4 is that in the construction used to prove Theorem 1.4, valid ("injective") public-keys are indistinguisbale from invalid ("lossy") public-keys, and thus the schemes does not have efficiently recognizable public-keys.

RDM security beyond encryption We note that the notion of RDM security applies not only to encryption but makes sense also in the context of more general cryptographic protocols. For instance, the notion of RDM security directly extends to commitments—just as in the case of encryption, we here let the RDM function select the messages to be committed to as a function of the committer's randomness. We remark that Theorem 1.1 readily extends also to rule out (even computationally binding and computationally-hiding) RDM-secure commitments. Additionally, Theorem 1.3 extends to show that any commitment scheme in the CRS model can be turned into a bounded RDM secure commitment scheme in the CRS model. However, Theorem 1.5 does not extend to the setting to commitments—using a collision-resistant hash function, any RDM secure commitment for short messages can be turned into a RDM-secure commitment for long messages. The above results for commitment schemes can be found in the full version of this work.

We leave an exploration of RDM security for other tasks (e.g., zero-knowledge and witness indistinguishability—where the RDM function may select the statement and witness to the proved as a function of the prover's randomness, or secure computation—where the RDM function may select a player's input as a function of his randomness) for future work.

On the Soundness of the Random-Oracle Methodology Starting with the work of Canetti, Goldreich and Halevi [21, 22], there are several "uninstantiability results" for the random oracle model [7], showing schemes that are secure in the random oracle model, but where every instantiation of random oracle with a concrete (efficient) function leads to an insecure protocol (see e.g., [5, 28, 36]). Another vein of work shows *tasks* (as opposed to schemes) that can be securely implemented in the random oracle model, but for which there are no secure implementations in the

standard model (see e.g., [41, 42, 9]). As far as we know, all these separations for tasks, however, make a relatively strong use of the random oracle model; [41, 9] rely on the security reduction "programming the random oracle", and [42] relies on the security reduction "seeing all the queries to the random oracle". Thus, it is conceivable that a weaker usage of random oracles may circumvent these uninstantiability results. For instance, Unruh [47] introduced a weaker random oracle model where the adversary may get an (inefficient) non-uniform advice about the random oracle, and suggested that proofs of security in this weaker random oracle model may still be "sound". We here address this question using RDM-secure encryption as a task.

We show that in the random-oracle model the existence of public-key encryption schemes imply the existence of "fully" RDM secure encryption schemes (i.e., without restricting the RDM function); our scheme is essentially identical to the hedged encryption scheme of [10] (but the analysis is quite different given the different security goals).[4] Our use of the random oracle model is extremely weak: we do not need to "program it", or "see queries to it", and security holds even the attacker may get any inefficient non-uniform advice about it (as in the model of [47]). (The only property we need of the random oracle is that it acts as a $k^{\log k}$-wise indepedent hash function.) We refer to such a model as the "ultra-weak" Random Oracle Model.

Theorem 1.6 (Informal Statement)**.** *Assume the existence of a secure public key encryption scheme. Then, there exists a encryption scheme* Π *that is "fully" RDM secure in the "ultra-weak" Random Oracle Model.*

Theorem 1.6, combined with our impossibility result (Theorem 1.1), thus yields an example of an arguably natural task (i.e., RDM-secure encryption) that can be securely implemented in the ultra-weak random-oracle model, but not in the standard model. Let us point out that a cruicial aspect of the security proof of our RO-based scheme is that the RDM function is not allowed to query the random oracle; in case we allow it to query the random oracle, our impossibility result still holds.

1.2 Related Work

As mentioned in the introduction, (circular) RDM security was first implicitly considered by Myers and Shelat [38] and explicitly by Hemmenway and Ostrovsky [33]. [38] [35] demonstrate semantic security of encryption schemes of a specific type of circular RDM attack, but do not formally introduce a notion of RDM security. Hemmenway and Ostrovsky [33] provide the first formal definition of RDM-secure encryption schemes, but only investigate, and provide constructions of, schemes satisfying the weaker notion of "circular-RDM one-wayness". As far as we know, we are the first to explicitely study the feasibility of satisfying (circular-)RDM *semantic* security (as opposed to one-wayness). As mentioned above, Bellare et al [10] study *hedged encryption schemes* that are closely related to RDM-secure encryption schemes; such encryption schemes are RDM secure if restricting the attacker to using RDM functions that do not depend on the public-key. Nevertheless, as mentioned, the constructions of both Bellare et al and Hemmenway and Ostrovsky are very useful to us.

As mentioned in the introduction, the related notion of key-dependent message (KDM) security was first introduced by Black, Rogaway, and Shrimpton in 2002 [13], who demonstrated the possibility of achieving their definition in the random-oracle model. The related notion of circular security (in which there exists a cycle of ciphertexts where each message depends on the previous

[4] Hedged encryption exists also in the plain model so we cannot hope to get a separation by directly appealing to the results of [10].

secret key) was independently and concurrently introduced by Camenisch and Lysyanskaya [20], who also showed constructions in the random-oracle model. Follow-up work considered message-dependent PRFs [31] and symmetric encryption [34, 4] in the standard model. In [30] barriers to constructing KDM secure schemes for general classes of key-dependencies. In 2008, Boneh, Halevi, Hamburg, and Ostrovsky presented the first KDM-secure public-key encryption scheme [15]; their construction was based on the DDH assumption. Subsequent work developed schemes that were KDM secure and CCA2 secure [19], KDM secure and resilient to leakage on the secret key [6], circular secure under alternative assumptions [16], and circular secure against larger classes of functions [17]. Recent work has also shown that there exist schemes that are secure under standard definitions but which are not 2-circular secure [1, 23].

A separate, but related line of related work focuses on leakage-resilient encryption (see e.g., [37, 25, 2, 3, 40, 18]). In a sense, RDM security can be viewed as a CPA security game where the attacker gets to see some leakage on the encryptor's randomness before selecting the messages; indeed, in our positive results, this view will be instrumental.

2 Preliminaries

For a distribution S, $s \leftarrow S$ means that s is chosen according to distribution S. For a set S, $s \leftarrow S$ means that s is chosen uniformly from the set S. U_n denotes the uniform distribution over n-bit strings. For a probabilistic algorithm A, $A(x;r)$ denotes the output of A running on input x with randomness r; $A(x)$ denotes the output of A on input x with uniformly chosen randomness. All logarithms are base 2 unless otherwise specified. We say that a function $\varepsilon : \mathbb{N} \to [0,1]$ is negligible if for every constant $c \in \mathbb{N}$, $\varepsilon(n) < k^{-c}$ for sufficiently large k.

2.1 Statistical Distance and Computational Indistinguishability

Definition 2.1 (Statistical Difference)**.** The *statistical difference* between two probability distributions X, Y is defined by $\Delta(X,Y) = (1/2) \cdot \sum_x |\Pr[x \leftarrow X] - \Pr[x \leftarrow Y]|$. X and Y are ε-close if $\Delta(X,Y) \leq \varepsilon$.

The *statistical difference* between two ensembles $\{X_k\}_k$ and $\{Y_k\}_k$ is a function δ defined by $\delta(k) = \Delta(X_k, Y_k)$. Two probability ensembles are said to be *statistically close* if their statistical difference is negligible. We also say X_k and Y_k are statistically close if $\Delta(X_k, Y_k) \leq \epsilon(k)$ for some negligible function ϵ.

Definition 2.2 (Computational Indistinguishability)**.** Two ensembles $\{X_k\}, \{Y_k\}$ are *computationally indistinguishable* if for every PPT distinguisher D, there exists a negligible function μ such that for every $k \in \mathbb{N}$,

$$|\Pr[D(1^k, X_k) = 1] - \Pr[D(1^k, Y_k) = 1]| \leq \mu(k).$$

2.2 Entropy

Definition 2.3. [Min-Entropy] The *min-entropy* of a random variable X, denoted $H_\infty(X)$ is defined by $H_\infty(X) = -\log(\max_x \Pr[x \leftarrow X])$.

Definition 2.4 (k-source)**.** A random variable X is a *k-source* if $H_\infty(X) \geq k$.

2.3 Hash Functions and Extractors

We use the standard t-wise independent hash functions and randomness extractors.

Definition 2.5 (t-wise Independent Hash Functions)**.** A family of hash functions $\mathcal{H} = \{h : S_1 \to S_2\}$ is *t-wise independent* if the following two conditions hold:

1. $\forall x \in S_1$, the random variable $h(x)$ is uniformly distributed over S_2, where $h \leftarrow \mathcal{H}$.
2. $\forall x_1 \neq \cdots \neq x_t \in S_1$, the random variables $h(x_1), \ldots, h(x_t)$ are independent, where $h \leftarrow \mathcal{H}$.

Definition 2.6 (Strong Randomness Extractor)**.** A function $\mathrm{Ext}\{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is a *strong (k, ε)-extractor* if for every k-source X over $\{0,1\}^n$, $(U_d, \mathrm{Ext}(X, U_d))$ is ε-close to (U_d, U_m).

2.4 Public-key Encryption Schemes

We now recall the formal definitions of public-key encryption schemes and its standard CPA security.

Definition 2.7 (Public-Key Encryption)**.** An *l-bit public-key encryption scheme* consists of a triple $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ of PPT algorithms where (i) Gen takes a security parameter 1^k as input and generates a pair of public and secret key $(pk, sk) \leftarrow \mathrm{Gen}(1^k)$, (ii) Enc takes a public key pk and a message m in a message space $\{0,1\}^{l(k)}$ as input and generates a ciphertext $c \leftarrow \mathrm{Enc}_{pk}(m)$, (iii) Dec is a deterministic algorithm that takes a secret key sk and a ciphertext c as input and outputs $m' = \mathrm{Dec}_{sk}(c)$, and (iv) there exists a negligible function μ such that for every $k \in \mathbb{N}$, for random $(pk, sk) \leftarrow \mathrm{Gen}(1^k)$,

$$\Pr\left[\exists m \in \{0,1\}^{l(k)} \mathrm{s.t.} \mathrm{Dec}_{sk}(\mathrm{Enc}_{pk}(m)) \neq m\right] \leq \mu(k),$$

where the probability is taken over the randomness of Gen and the randomness of the encryption. We say that Π has *perfect correctness* if the above condition holds for $\mu(k) = 0$.

We recall the standard definitions of CPA and CCA security.

Definition 2.8 (CPA and CCA Security)**.** An l-bit public-key encryption scheme $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is *CPA-secure* if for every probabilistic polynomial time adversary $A = (A_1, A_2)$, the ensembles $\{\mathrm{IND}_0^\Pi(A, k)\}_k$ and $\{\mathrm{IND}_1^\Pi(A, k)\}_k$ are computationally indistinguishable, where

$$\begin{aligned} \mathrm{IND}_b^\Pi(A, k) := \quad & (pk, sk) \leftarrow \mathrm{Gen}(1^k) \\ & (m_0, m_1, \mathsf{state}) \leftarrow A_1(1^k, pk) \\ & c \leftarrow \mathrm{Enc}_{pk}(m_b) \\ & o \leftarrow A_2(c, \mathsf{state}) \\ & \text{Output } o \end{aligned}$$

We say Π is *CCA-secure* if the above holds when A_2 has access to a decryption oracle but is not allowed to query the decryption oracle with the challenge ciphertext c.

Remark. In the above definition and for essentially all the results in this paper, we consider a *uniform* polynomial-time attacker A. In case security holds against also *non-uniform* polynomial-time attackers, we refer to the scheme as being non-uniformly CPA/CCA secure. As is often the case, all our constructions in uniform setting directly extend also to the case of non-uniform security (if assuming that the underlying schemes are non-uniformly secure).

Note that the above definition assumes that messages encrypted are chosen independently of the randomness used by the encryption algorithm.

3 Definitions

In this section, we formally define two notions of randomness-dependent message security for encryption schemes.

Our first definition is essentially equivalent to the definition of RDM security due to Hemmenway and Ostrovsky [33]. In this definition, messages are adversarially chosen functions (after seeing the public key) of the randomness used for encryption: we say the encryption scheme is secure if the adversary cannot distinguish between encryptions of different functions of the randomness.

Definition 3.1. [RDM-Security] An l-bit public-key encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ is *randomness-dependent message secure (RDM-secure)* if for every PPT adversary $A = (A_1, A_2)$, the ensembles $\{\text{RDM}_0^\Pi(A,k)\}_{k\in\mathbb{N}}$ and $\{\text{RDM}_1^\Pi(A,k)\}_{k\in\mathbb{N}}$ are computationally indistinguishable where

$$\begin{aligned}\text{RDM}_b^\Pi(A,k) := \quad & (pk, sk) \leftarrow \text{Gen}(1^k)\\ & (f_0, f_1, \mathsf{state}) \leftarrow A_1(1^k, pk)\\ & r \leftarrow U_R\\ & c \leftarrow \text{Enc}_{pk}(f_b(r); r)\\ & o \leftarrow A_2(c, \mathsf{state})\\ & \text{Output } o\end{aligned}$$

and R is the encryption randomness length of Π. The RDM functions f_b are represented as circuits from $\{0,1\}^{|r|}$ to $\{0,1\}^{l(k)}$ We say Π is RDM-CCA-secure if the above holds when A_2 has access to a decryption oracle but is not allowed to query the decryption oracle with the challenge ciphertext c.

We remark that by a standard hybrid argument, we can assume without loss of generality that the adversary A_1 always choose f_1 to be a constant function $f_1 = 0$. As mentioned, Definition 3.1 is essentially identical to the notion of RDM security defined by Hemmenway and Ostrovsky [33]: the definition of [33] is a multi-message version of Definition 3.1 where the attacker gets to see a sequence of encrypted messages (that may depend in a correlated way on the randomness used to encrypt them), and thus the definition of [33] implies Definition 3.1. (Looking forward, since we are proving an impossibility result regarding Definition 3.1, considering a weaker definition makes our results stronger.)

Consider a sequence of encryptions where messages are functions of the previous (but most recent) encryption randomness and ciphertext. Security in this setting is guaranteed by CPA security, since encryption randomness is still independent of the messages. However if this dependency is circular, it is unclear whether or not we have security. We now formally introduce this notion of circular randomness dependent message security.

Definition 3.2 (q-circular RDM Security)**.** Let $q : \mathbb{N} \to \mathbb{N}$ be efficiently computable. An l-bit public-key encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ is *q-circular RDM secure* if for every PPT adversary $A = (A_1, A_2)$, the following two ensembles $\{\text{CIR}_0^\Pi(A,k)\}_{k\in\mathbb{N}}$ and $\{\text{CIR}_1^\Pi(A,k)\}_{k\in\mathbb{N}}$ are computationally indistinguishable, where

$$
\begin{aligned}
\mathrm{CIR}_b^{\Pi}(A,k) := \quad & (pk, sk) \leftarrow \mathrm{Gen}(1^k) \\
& (f_0^1, f_0^2, \ldots, f_0^{q(k)}, f_1^1, f_1^2, \ldots, f_1^{q(k)}, \mathsf{state}) \leftarrow A_1(1^k, pk) \\
& r^1, r^2, \ldots, r^{q(k)} \leftarrow U_R^{q(k)} \\
& c^1 \leftarrow \mathrm{Enc}_{pk}(f_b^1(r^q); r^1) \\
& \text{for } i = 2, \ldots, q \\
& \quad c^i \leftarrow \mathrm{Enc}_{pk}(f_b^i(r^{i-1}, c^{i-1}); r^i) \\
& o \leftarrow A_2(\bar{c}, \mathsf{state}) \\
& \text{Output } o
\end{aligned}
$$

and R is the encryption randomness length of Π. The RDM functions f_b^i are represented as circuits as defined in Definition 3.1. $\bar{c}$ denotes the vector $(c^1, c^2, \ldots c^n)$. Furthermore, Π is *circular RDM secure* if Π is k^c-circular RDM secure for every constant c. *q-circular-CCA* and *circular-CCA* RDM security are defined in analogous way.

Remark. Note that by a hybrid argument, we can assume without loss of generality that A always choose $f_1^i = 0$ for every $i \in [q]$. We will use this observation later in the proof of Theorem 5.6.

We also define relaxations of RDM security and circular RDM security where we restrict the RDM function to be computable by circuits of *a priori* bounded size.

Definition 3.3. Let $s : \mathbb{N} \to \mathbb{N}$ be efficiently computable. An l-bit public key encryption scheme Π is *s-bounded RDM secure* (resp., *s-bounded (q-)circular RDM secure*) if Π is RDM secure (resp., (q-)circular RDM secure) under the additional restriction that in the corresponding security game, the adversary A_1 can only output RDM functions computable by circuits of size bounded by $s(k)$. CCA security is defined analogously.

4 RDM Security and Our Impossibility Results

In this section we prove that both RDM-security and q-circular security are impossible to achieve. Throughout this section, we focus on bit-encryption schemes; this only makes our results stronger. We first establish the impossibility result on the RDM-secure encryption schemes; our techniques (of using pairwise independent hashfunctions to signal a message) are similar to those used by Bellare and Keelveedhi [12] in a different context.

Theorem 4.1. *For every* 1*-bit encryption scheme* $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$*,* Π *is not RDM-secure.*

Proof. Let $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ be a 1-bit encryption scheme. We construct a PPT adversary $A = (A_1, A_2)$ that breaks the RDM security of Π. The idea is to use f_b to signal the bit b in the RDM_b^{Π} experiment by pairwise independent hash functions.

Fix a security parameter $k \in \mathbb{N}$. Let C denotes the ciphertext space of Π for the corresponding security parameter k, and let $\mathcal{H} = \{h : C \to \{0,1\}\}$ be a pairwise independent hash function family that hashes ciphertexts to a bit. Our adversary A uses $h \leftarrow \mathcal{H}$ to construct functions $f_{b,h}$ for $b \in \{0,1\}$ that signals the bit b as follows.

- $A_1(1^k, pk)$: A_1 samples $h \leftarrow \mathcal{H}$ and outputs $(f_{0,h}, f_{1,h}, h)$, where for $b \in \{0,1\}$, $f_{b,h}$ on input r, outputs a message $m \in \{0,1\}$ such that $h(\mathrm{Enc}_{pk}(m, r)) = b$ if such an m exists; otherwise $f_{b,h}$ outputs $m = 0$.
- $A_2(c, h)$: A_2 simply outputs one bit $h(c)$.

To show that A breaks the RDM security of Π, it suffices to show the following claim, which clearly implies $\mathrm{RDM}_0^\Pi(A,k)$ and $\mathrm{RDM}_1^\Pi(A,k)$ are distinguishable.

Claim 4.2. $\Pr[RDM_b^\Pi(A,k)=b] \geq 3/4 - \mathrm{negl}(k)$ *for* $b \in \{0,1\}$.

Proof. Note that the output of $\mathrm{RDM}_b^\Pi(A,k)$ is simply $h(\mathrm{Enc}_{pk}(f_b(r),r))$ where $(pk,sk) \leftarrow \mathrm{Gen}(1^k), r \leftarrow U_{|r|}$, and $h \leftarrow \mathcal{H}$. The correctness of Π implies that,

$$\Pr_{pk,r}[\mathrm{Enc}_{pk}(0,r) \neq \mathrm{Enc}_{pk}(1,r)] \geq 1 - \mathrm{negl}(k). \tag{1}$$

When this is the case, by the pairwise independence,

$$\Pr_h[\exists\ m \text{ s.t. } h(\mathrm{Enc}_{pk}(m,r)) = b] = 3/4.$$

It follows by an union bound that

$$\begin{aligned}
&\Pr[\mathrm{RDM}_b^\Pi(A,k)=b] \\
&\geq \Pr_{pk,r,h}[(\mathrm{Enc}_{pk}(0,r) \neq \mathrm{Enc}_{pk}(1,r)) \wedge (\exists\ m \text{ s.t. } h(\mathrm{Enc}_{pk}(m,r)) = b)] \\
&\geq 3/4 - \mathrm{negl}(k).
\end{aligned}$$

□

□

We proceed to establish the impossibility result on the circular RDM-secure encryption schemes.

Theorem 4.3. *For every* 1*-bit encryption scheme* $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$*,* Π *is not* q*-circular RDM-secure for every efficiently computable and polynomially bounded* q*.*

We prove Theorem 4.3 by showing that in fact, circular RDM security implies RDM security.

Theorem 4.4. *Let* $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ *be a* 1*-bit encryption scheme, and* $q : \mathbb{N} \to \mathbb{N}$ *be efficiently computable and polynomially bounded. If* Π *is* q*-circular RDM-secure, then* Π *is RDM-secure.*

Theorem 4.3 follows by combining Theorem 4.1 and 4.4. Before presenting the formal proof of Theorem 4.4, we first discuss the proof for the special case that Π has perfect correctness and that $q = 2$, to illustrate the idea behind the proof. Suppose there exists a PPT adversary A that breaks the RDM security of Π, we want to construct a PPT adversary B that breaks the 2-circular security of Π.

The idea is to let B simulate the attack of A in the circular RDM security game using the second message (in general, using the last message). More precisely, recall that in the RDM security game RDM_b^Π, A generates RDM functions f_0 and f_1, and receives $c = \mathrm{Enc}_{pk}(f_b(r);r)$. To simulate the attack of A in CIR_b^π, B generates $f_0^1, f_0^2, f_1^1, f_1^2$ in a way so that B will receive $\bar{c} = (c^1, c^2)$ with $c^2 = \mathrm{Enc}_{pk}(f_b(r^2);r^2)$. Then B can output whatever A_2 outputs on input c^2, and break the circular RDM security with the same advantage as A.

Now, the key observation is that the RDM function $f_b^2(r^1, c^1)$ can in fact decrypt c^1 to get the message $f_b^1(r^2)$ by checking whether c^1 equals to $\mathrm{Enc}_{pk}(0,r^1)$ or $\mathrm{Enc}_{pk}(1,r^1)$ (the perfect correctness implies $\mathrm{Enc}_{pk}(0,r^1) \neq \mathrm{Enc}_{pk}(1,r^1)$ and the decryption will be always correct). Thus, B can let $f_b^1 = f_b$ and let $f_b^2(r^1,c^1) = f_b^1(r^2)$, and by doing so B will receive $c^2 = \mathrm{Enc}_{pk}(f_b^2(r^1,c^1),r^2) = \mathrm{Enc}_{pk}(f_b(r^2),r^2)$, as desired. This completes the proof of the special case.

We can readily extend the proof to the general q-circular RDM security, by letting B set $f_b^1 = f_b$ and $f_b^{i+1}(r^i, c^i) = f_b^i(r^{i-1}, c^{i-1})$ for $i = 1, \ldots, q-1$. On the other hand, the imperfect correctness only causes negligible probability of decryption errors, and reduce the advantage of B by a negligible amount. We proceed to present a formal proof.

Proof. For the sake of contradiction, suppose that there exists a PPT adversary A that breaks the RDM security of Π. Namely, there exists a PPT distinguisher D and a non-negligible function $\varepsilon(\cdot)$ such that for every $k \in \mathbb{N}$,

$$|\Pr[D(\mathrm{RDM}_0^{\Pi}(A,k)) = 1] - \Pr[D(\mathrm{RDM}_1^{\Pi}(A,k)) = 1]| \geq \varepsilon(k).$$

We construct a PPT adversary B that breaks the q-circular RDM security of Π as follows.

- $B_1(1^k, pk)$:
  - B_1 runs $A_1(1^k, pk)$ to obtain $(f_0, f_1, \mathsf{state})$.
  - B_1 sets $f_0^1 = f_0$ and $f_1^1 = f_1$.
  - For $i = 2, \ldots, q$ and $b \in \{0,1\}$, B constructs f_b^i as follows. On input (r^{i-1}, c^{i-1}), f_b^i outputs 0 if $\mathrm{Enc}_{pk}(0, r^{i-1}) = c^{i-1}$, and outputs 1 otherwise.
  - Finally, B_1 outputs $(f_0^1, \ldots, f_0^q, f_1^1, \ldots, f_1^q, \mathsf{state})$.
- $B_2(\bar{c}, \mathsf{state})$: B_2 simply outputs $A_2(c^q, \mathsf{state})$.

We claim that for $b \in \{0,1\}$,

$$\Delta(\mathrm{RDM}_b^{\Pi}(A,k)), \mathrm{CIR}_b^{\Pi}(B,k)) \leq \mathrm{negl}(k),$$

which implies

$$|\Pr[D(\mathrm{CIR}_0^{\Pi}(B,k)) = 1] - \Pr[D(\mathrm{CIR}_1^{\Pi}(B,k)) = 1]| \geq \varepsilon(k) - \mathrm{negl}(k),$$

and completes the proof. To prove the claim, we note that the correctness of Π implies that there exist negligible functions μ and η such that for at least $1-\mu(k)$ fraction of pk, $\Pr_r[\mathrm{Enc}_{pk}(0,r) \neq \mathrm{Enc}_{pk}(1,r)] \geq 1-\eta(k)$. Hence, with such a pk, $\Pr[f_b^2(r^1, c^1) = f_b^1(r^q)] \geq 1-\eta(k)$ and also for every $i \in \{3, \ldots, q\}$,

$$\Pr[f_b^i(r^{i-1}, c^{i-1}) = f_b^{i-1}(r^{i-2}, c^{i-2})] \geq 1 - \eta(k).$$

By a union bound,

$$\Pr[f_b^q(r^{q-1}, c^{q-1}) = f_b^1(r^q) = f_b(r^q)] \geq 1 - q(k)\eta(k).$$

Therefore, in $\mathrm{CIR}_b^{\Pi}(B,k)$, except for negligible probability, B forwards the same distribution $c^q = \mathrm{Enc}_{pk}(f_b(r^q), r^q)$ to A_2 as in $\mathrm{RDM}_b^{\Pi}(A,k)$, and thus $\Delta(\mathrm{RDM}_b^{\Pi}(A,k)), \mathrm{CIR}_b^{\Pi}(B,k)) \leq \mathrm{negl}(k)$. □

5 Positive Results

5.1 Bounded RDM Security

In the previous sections we have seen that RDM security and circular RDM security are impossible to achieve. In this section we see how we can achieve the weaker notions of bounded RDM security

and bounded circular RDM security. In fact we achieve a stronger notion of RDM security which implies both of the above.

This strong RDM security is in fact security in the presence of randomness leakage (such that the leakage function size is *a priori* bounded by a polynomial) which is available to the adversary when it chooses the messages to encrypt.

Definition 5.1. For every $s, p : \mathbb{N} \to \mathbb{N}$ an l-bit public-key encryption scheme $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ is *s-bounded p-strong RDM secure (BSRDM-secure)* if for every PPT adversary $A = (A_1, A_2)$, the ensembles $\{\mathrm{BSRDM}_0^{\Pi}(A,k)\}_{k\in\mathbb{N}}$ and $\{\mathrm{BSRDM}_1^{\Pi}(A,k)\}_{k\in\mathbb{N}}$ are computationally indistinguishable where

$$\begin{aligned}\mathrm{BSRDM}_b^{\Pi}(A,k) := \quad & (pk, sk) \leftarrow \mathrm{Gen}(1^k)\\ & r \leftarrow U_R\\ & (f, \mathsf{state}_1) \leftarrow A_1(1^k, pk)\\ & (m_0, m_1, \mathsf{state}_2) \leftarrow A_2(f(r), \mathsf{state}_1)\\ & c \leftarrow \mathrm{Enc}_{pk}(m_b; r)\\ & o \leftarrow A_3(c, \mathsf{state}_2)\\ & \text{Output } o,\end{aligned}$$

R is the encryption randomness length of Π and $f : \{0,1\}^{|r|} \to \{0,1\}^{p(k)}$ is a function computed by a circuit of size at most $s(k)$. CCA security is defined analogously.

We show that any secure encryption scheme can be compiled to a bounded strong RDM-secure encryption scheme (with "long" encryption randomness).

Theorem 5.2. *Assume the existence of a CPA (resp., CCA) secure public key encryption scheme. Then, there exists a l-bit s-bounded p-strong RDM-secure (resp., RDM-CCA-secure) encryption scheme for every polynomial l, s and p.*

We start by providing a construction that converts any secure encryption scheme to bounded strong RDM secure encryption scheme. The main idea is that though leakage degrades the randomness, the randomness is long enough to have enough residual min-entropy so that the random bits necessary for encryption can be extracted from it. The problem with this is that the extractor seed will have to be part of the public key, and the adversary can choose a leakage function after seeing the public key. Hence the leakage could be such that the seed always fails to extract randomness from the source. This is where we exploit the fact that the set of possible leakage functions is bounded: using a union bound, we show that if the randomness used by the encryption scheme is long enough, then with overwhelming probability a random seed can extract randomness from the source resulting from any leakage function. The following lemma captures the above idea.

Lemma 5.3 (Deterministic Extraction From Bounded Leakage Sources)**.** *Let $\mathcal{F} = \{f : \{0,1\}^n \to \{0,1\}^\ell\}$ be a class of (leakage) functions. Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ be a t-wise independent hash function family. If*

$$\begin{cases} t \geq 2(m + \ell + \log|\mathcal{F}| + \log(1/\delta) + 3), \\ m \leq n - \ell - 3\log(1/\varepsilon) - \log t - 5, \end{cases}$$

then with probability at least $(1-\delta)$ over $h \leftarrow \mathcal{H}$, it holds that for every $f \in \mathcal{F}$,

$$\Delta((f(U_n), h(U_n)), (f(U_n), U_m)) \leq \varepsilon.$$

The proof of the lemma can be found in Appendix A, and relies on the ideas similar to those used by [46] to demonstrate deterministic extraction from sources computable by bounded size circuits. We now see how we can get a bounded-SRDM-secure encryption scheme from any secure encryption scheme. The following transformation is essentially identical to the one used in [33] but using different parameters.[5]

Definition 5.4. For every polynomial s and p and encryption scheme $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$, define a new encryption scheme $\Pi' = (\mathrm{Gen}', \mathrm{Enc}', \mathrm{Dec}')$ as follows:

- $\mathrm{Gen}'(1^k) : (pk, sk) \leftarrow \mathrm{Gen}(1^k), h_k \leftarrow \mathcal{H}_k$ where $\mathcal{H}_k = \{h_k : \{0,1\}^{R'(k)} \to \{0,1\}^{R(k)}\}$ is a $t(k)$-wise independent family of hash functions where $R(.)$ is the length of the randomness of Enc, $R'(.)$ is the length of the randomness of Enc',

  $$t(k) \geq 2(R(k) + k + s(k) + p(k) + 3)$$

  and

  $$R'(k) = p(k) + R(k) + 3k + \log t(k) + 5$$

  Output $((pk, h_k), sk)$.

- $\mathrm{Enc}'_{(pk,h_k)}(m) : r \leftarrow U_{R'(k)}$; output $\mathrm{Enc}_{pk}(m; h_k(r))$.

- $\mathrm{Dec}'_{sk}(c)$: output $\mathrm{Dec}_{sk}(c)$.

We now show that the above construction transforms a CPA (resp., CCA) secure scheme to a bounded strong RDM (resp., CCA-RDM) secure scheme (which implies Theorem 5.2).

Lemma 5.5. *Let s, p be polynomials. Let Π be a CPA (resp., CCA) secure public key encryption scheme, and Π' be the transformed encryption scheme obtained from Definition 5.4. Then, Π' is s-bounded p-strong RDM (resp., CCA-RDM) secure.*

Proof. We start by proving the lemma for the case of CPA security. Assume for contradiction that there exists a probabilistic polynomial time adversary A that breaks the s-bounded p-strong RDM security of Π'. That is, there exists a probabilistic polynomial time distinguisher D and non-negligible function ϵ such that for all $k \in \mathbb{N}$

$$|\Pr[D(\mathrm{BSRDM}_0^{\Pi'}(A,k)) = 1] - \Pr[D(\mathrm{BSRDM}_1^{\Pi'}(A,k)) = 1]| \geq \varepsilon(k)$$

We construct an adversary B that breaks the CPA security of Π. On input $(pk, 1^k)$, B samples $h_k \leftarrow \mathcal{H}_k$ and runs $(f, \mathsf{state}_1) \leftarrow A_1((pk, h_k), 1^k)$. B then samples $r_1 \leftarrow \{0,1\}^{R'(k)}$, runs $(m_0, m_1, \mathsf{state}_2) \leftarrow A_2(f(r_1), \mathsf{state}_1)$ and outputs $(m_0, m_1, \mathsf{state}_2)$. On input (c, state_2), B outputs $A_3(c, \mathsf{state}_2)$.

Note that the only difference in the experiments $\mathrm{IND}_b^{\Pi}(B,k)$ and $\mathrm{BSRDM}_b^{\Pi'}(A,k)$ is that in $\mathrm{BSRDM}_b^{\Pi'}(A,k)$, A receives a randomness leakage $f(r)$ and a ciphertext encrypted using $h_k(r)$, whereas in $\mathrm{IND}_b^{\Pi}(B,k)$, A receives $f(r_1)$ but the ciphertext is encrypted using a fresh independent randomness r. Applying Lemma 5.3 with $\mathcal{F}$ being the class of s-bounded RDM functions and note that $|\mathcal{F}| \leq 2^{s(k)}$, with probability at least $1 - 2^{-k}$ over the choice of $h_k \leftarrow \mathcal{H}_k$, the statistical

[5] As mentioned in Section 1.2, the results of [33] however require the underlying encryption schemes to satisfy additional properties (e.g., "lossiness") and the results established about the resulting encryption scheme are very different.

distance between the two experiments conditioned on this h_k is at most 2^{-k}. Therefore, the overall statistical distance is at most 2^{-k+1}. This implies that

$$|Pr[D(\mathrm{IND}_0^{\Pi}(B,k))=1] - Pr[D(\mathrm{IND}_1^{\Pi}(B,k))=1]| \geq \varepsilon - 2^{-k+1} \geq \varepsilon/2$$

which is a contradiction.

The above proof extends to the case of CCA security if we simply let the adversary B forward the decryption oracle queries of A, and forward back the answers.

□

It is clear that bounded strong RDM security implies RDM security. Additionally, bounded strong RDM security is strong enough to also imply bounded circular RDM security.

Theorem 5.6. *For all l-bit public key encryption schemes $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$, if Π is s-bounded l-strong RDM secure (resp., CCA-RDM secure) then Π is s-bounded circular RDM secure (resp., CCA-RDM secure).*

Proof. We start by proving the theorem for the case of CPA security. Assume for contradiction there exists a polynomial q and probabilistic polynomial time adversary A that breaks the s-bounded q-circular RDM security of Π. Namely there exists a probabilistic polynomial time distinguisher D and non-negligible function ε such that for all $k \in \mathbb{N}$

$$|Pr[D(\mathrm{BCIR}_0^{\Pi}(A,k))=1] - Pr[D(\mathrm{BCIR}_1^{\Pi}(A,k))=1]| \geq \varepsilon(k)$$

Define for $1 \leq j \leq q(k)+1$, $\mathrm{H}_j^{\Pi}(A,k)$ as follows

$$\begin{array}{ll} \mathrm{H}_j^{\Pi}(A,k) := & (pk,sk) \leftarrow \mathrm{Gen}(1^k) \\ & f^1, f^2, \ldots, f^{q(k)}, \mathsf{state} \leftarrow A_1(1^k, pk) \\ & r^1, r^2 \ldots r^{q(k)} \leftarrow \{0,1\}^{q(k)|r|} \\ & \text{for } i = 1, 2 \ldots, q(k) \\ & \quad \text{if } i < j \text{ then } c^i \leftarrow \mathrm{Enc}_{pk}(f^i(r^{i-1}, c^{i-1}); r^i) \\ & \quad \text{else } c^i \leftarrow \mathrm{Enc}_{pk}(0; r^i) \\ & o \leftarrow A_2(\bar{c}, \mathsf{state}) \\ & \text{Output } o \end{array}$$

where $r^0 = r^{q(k)}$, c^0 is the empty string and $\bar{c} = c^1, c^2, \ldots c^{q(k)}$

We define an adversary B that breaks the s-bounded l-strong RDM security of Π. On input $(1^k, pk)$, B first runs $f^1, f^2, \ldots, f^{q(k)}, \mathsf{state} \leftarrow A_1(1^k, pk)$ and outputs f^1. Note that f^1 is computed by a circuit of size at most $s(k)$ and $f^1 : \{0,1\}^{|r|} \rightarrow \{0,1\}^{l(k)}$. On input $f^1(r)$, B first samples $j \leftarrow \{1, 2, \ldots q(k)\}$ and $\bar{r}^{-j} \leftarrow \{0,1\}^{|r|(q(k)-1)}$. For $i = 1, \ldots j-1$, B computes $c^i \leftarrow \mathrm{Enc}_{pk}(f^i(r^{i-1}, c^{i-1}); r^i)$ where if $j = q(k)$ then $f^1(r)$ is used as $f^1(r^0, c^0)$ to compute c^1 ($f^1(r)$ is not used otherwise). B outputs $(f^j(r^{j-1}, c^{j-1}), 0)$ as (m_0, m_1). On input c, B sets $c^j = c$ and computes for $i = j+1, \ldots q(k)$, $c^i \leftarrow \mathrm{Enc}_{pk}(0; r^i)$. Finally B outputs $A_2(\bar{c}, \mathsf{state})$.

B is a probabilistic polynomial time adversary and for all $k \in \mathbb{N}$, denoting B running with a fixed value of j as B_j, we have, $\mathrm{BSRDM}_1^{\Pi}(B_j, k) = \mathrm{H}_j^{\Pi}(A,k)$ and $\mathrm{BSRDM}_0^{\Pi}(B_j, k) = \mathrm{H}_{j+1}^{\Pi}(A,k)$. Since j is uniform in $\{1, 2, \ldots q(k)\}$ we have,

$$\Pr[D(\mathrm{BSRDM}_1^{\Pi}(B,k))=1] = (1/q(k))\Sigma_{j=1}^{q(k)} \Pr[D(\mathrm{H}_j^{\Pi}(A,k))=1]$$

and

$$\Pr[D(\mathrm{BSRDM}_0^{\Pi}(B,k))=1] = (1/q(k))\Sigma_{j=1}^{q(k)} \Pr[D(\mathrm{H}_{j+1}^{\Pi}(A,k))=1].$$

Since $\mathrm{H}_1^{\Pi}(A,k) = \mathrm{BCIR}_1^{\Pi}(A,k)$ and $\mathrm{H}_{q(k)+1}^{\Pi}(A,k) = \mathrm{BCIR}_0^{\Pi}(A,k)$, we have

$$|\Pr[D(\mathrm{BSRDM}_1^{\Pi}(B,k)) = 1] - \Pr[D(\mathrm{BSRDM}_0^{\Pi}(B,k)) = 1]| \geq \frac{\varepsilon(k)}{q(k)} = \varepsilon'(k)$$

where ε' is also non-negligible, which is a contradiction.

The above proof extends to the case of CCA security if we simply let the adversary B forward the decryption oracle queries of A, and forward back the answers.

□

5.2 Bounded RDM Secure Encryption Schemes with Short Randomness

The above construction yields strong bounded RDM-secure encryption schemes where the length of the randomness is longer than the length of the message. We now provide a construction of a bounded RDM-secure and bounded circular RDM-secure encryption scheme that can encrypt arbitrarily long messages using "short" randomness. This construction, however, relies on stronger cryptographic assumption—namely, we require the existence of "lossy" trapdoor functions.

Definition 5.7 ([45])**.** A tuple $(\mathrm{GenLossy}, \mathrm{GenInj}, F, \mathrm{invert})$ is an (n,u)-lossy trapdoor function if the following holds:

- (Injection mode) For every $k \in \mathbb{N}$, $\Pr[(pk,sk) \leftarrow \mathrm{GenInj}(1^k) : x \leftarrow U_{n(k)} : \mathrm{invert}_{sk}(F_{pk}(x)) = x] = 1$
- (Lossy mode) For every $k \in \mathbb{N}$ and $pk \leftarrow \mathrm{GenLossy}(1^k)$, the size of the range of $F_{pk}(.)$ (which takes as input strings of length $n(k)$) is at most $2^{u(k)}$.
- The following ensembles are computationally indistinguishable

$$\{(pk,sk) \leftarrow \mathrm{GenInj}(1^k) : pk\}_{k\in\mathbb{N}}$$

$$\{pk \leftarrow \mathrm{GenLossy}(1^k) : pk\}_{k\in\mathbb{N}}$$

We turn to providing our construction of a bounded-RDM secure encryption scheme that can encrypt also "long" messages using "short" randomness—the ratio between the message-length and the randomness length is proportional to the lossiness of the trapdoor function. Formally, we establish the following theorem.

Theorem 5.8. *Let l and R be the message length and randomness length parameters with $R(k) \geq k$. Assuming the existence of (n,u)-lossy trapdoor functions with $n \geq 3(l+R)$ and $u \leq R/8$, then for every polynomial s, there exist a l-bit s-bounded circular RDM secure encryption scheme with randomness length R.*

In particular, assuming the DDH assumption holds, for every polynomial l, R, s with $R(k) \geq k$, there exist a l-bit s-bounded circular RDM secure encryption scheme with randomness length R.

We mention that the "in particular" part of the theorem follows by the DDH-based construction of lossy trapdoor functions in [45]. Our construction is closely related to the "pad-then-deterministic" construction of hedged encryption schemes of Bellare et al [10], where the encryption is done by first applying a *invertible* universal hash *permutation* h to the message-randomness pair $(m||r)$ and then applying a lossy trapdoor function F_{pk} to the hashed value. Recall that hedged encryption scheme already satisfy a notion of RDM security when restricting to RDM functions

that do not depend on the public-key. To deal with RDM functions that depend on the public key, our key modification to their scheme is to replace the use of univeral hashing with t-wise independent hashing. However, since constructions of t-wise independent *permutations* are not known, to deal with arbitrary t-wise independent hash functions, we further modify the scheme to "pad" the message-randomness pair with a sufficiently long sequence of 0's.

Recall that the standard construction of t-wise independent hash functions is a degree $t-1$ univariate polynomial over a prime field, which is invertible by the Berlekamp algorithm [**?**].

Definition 5.9. Let l, R, and s be the message length, randomness length, and size parameters with $R(k) \geq k$. Let $(\text{GenLossy}, \text{GenInj}, F, \text{invert})$ be an (n,u)-lossy trapdoor function with public-key length v such that $u \leq R/8$ and $n = 3(l+R)$. Let $t = 8(s+u+v+R)$ and $\mathcal{H}_n = \{h : \{0,1\}^n \to \{0,1\}^n\}$ be an *invertible* family of $t(\cdot)$-wise independent hash functions. Define an l-bit s-bounded (circular) RDM-secure encryption scheme $\Pi = (\text{Gen}, \text{Enc}, \text{Dec})$ with randomness length R as follows[6]:

- $\text{Gen}(1^k) : (pk, sk) \leftarrow \text{GenInj}(1^k), h \leftarrow \mathcal{H}_n$; output $((pk,h),(sk,h))$.
- $\text{Enc}_{(pk,h)}(m) : r \leftarrow U_{R(k)}$; output $c = F_{pk}(h(m||r||0^{2(l+R)}))$.
- $\text{Dec}_{(sk,h)}(c)$: output the first $l(k)$ bits of $h^{-1}(\text{invert}_{sk}(c))$.

While our construction is bounded circular RDM secure, it is instructive to first focus on the bounded RDM security. Recall the security of the [10] scheme (which relies on a construction of deteministic encryption from [14]) relies on a "crooked" version of leftover hash lemma [24], which asserts that when F_{pk} has small range size (which is the case in the lossy mode) and the source $(m||r)$ has sufficient min-entropy and is independent of h, then $F_{pk}(h(m||r))$ is statistically close to the "crooked" distribution $F_{pk}(U_{|m|+|r|})$.

In our context, however, the adversary selects a s-bounded RDM function f after seeing the public key, and thus the source $(f(r)||r||0^{2(l+R)})$ may be correlated with the hash function h (and also F_{pk}). We overcome this issue by using t-wise independent hashing and proving a crooked version of the deterministic extraction lemma from computationally bounded source of Trevisan and Vadhan [46]. The lemma asserts that with overwhelming probability over $h \leftarrow \mathcal{H}$, the encryption $F_{pk}(f(r)||r||0^{2(l+R)})$ is statistically close to a corresponding crooked distribution $F_{pk}(U_n)$ for *every lossy* function F_{pk} and *every* s-bounded RDM function f. Therefore, the s-bounded RMD security follows by switching to the lossy mode and applying the crooked deterministic extraction lemma. We proceed to state the crooked deterministic extraction lemma and prove the s-bounded RDM security of our scheme. The proof of Lemma 5.10 is deferred to Appendix B and follows similar techniques to those used by [46].

Lemma 5.10 (Crooked Deterministic Extraction)**.** *Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^n\}$ be a t-wise independent hash function family. Let $\mathcal{F} = \{f : \{0,1\}^n \to R_f\}$ be a family of functions where each $f \in \mathcal{F}$ has range R_f of size $|R_f| \leq 2^m$. Let $\mathcal{C}$ be a family of distributions over $\{0,1\}^n$ such that every $X \in \mathcal{C}$ has min-entropy $H_\infty(X) \geq k$. If*

$$\begin{cases} t \geq 2(m + \log|\mathcal{F}| + \log|\mathcal{C}| + \log(1/\delta) + 3), \\ m \leq k - 2\log(1/\varepsilon) - \log t - 2, \end{cases}$$

[6] In fact, to achieve only bounded RDM security (as opposed to circular RDM security), it suffices to, say, satisfy $u \leq R/5$ and set $t = 4(s+u+v)$. We do not optimize the parameters here.

then with probability at least $(1-\delta)$ *over* $h \leftarrow \mathcal{H}$*, it holds that for every* $f \in \mathcal{F}$ *and every* $X \in \mathcal{C}$*,*

$$\Delta(f(h(X)), f(U_n)) \leq \varepsilon.$$

Lemma 5.11. *The* l*-bit encryption scheme* $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ *constructed in Definition 5.9 is correct and* s*-bounded RDM secure.*

Proof. We first show that Π is correct. Let $S = \{x||0^{2(l+R)} : x \in \{0,1\}^{l+R}\}$. Note that for $((pk,h),(sk,h)) \leftarrow \mathrm{Gen}(1^k)$, if $h|_S : S \to \{0,1\}^n$ is injective, then for every message $m \in \{0,1\}^l$, $\mathrm{Dec}_{(sk,h)}(\mathrm{Enc}_{(pk,h)}(m)) = m$ with probability 1. Thus, it suffices to show that $h|_S$ is injective except with negligible probability, which follows directly by Lemma B.5.

For proving RDM security, we show that for every PPT adversary A, the view of A in both BRDM_0^{Π} and BRDM_1^{Π} are computationally indistinguishable to the same distribution (that is independent of the message and randomness used in the encryption).

Let $b \in \{0,1\}$. Recall that in the experiment $\mathrm{BRDM}_b^{\Pi}(A,k)$, the adversary A receives (i) a public key (pk,h), where $(pk,sk) \leftarrow \mathrm{GenInj}(1^k)$ and $h \leftarrow \mathcal{H}_n$, and (ii) a ciphtertext $c = \mathrm{Enc}_{(pk,h)}(f_b(r);r) = F_{pk}(h(f_b(r)||r||0^{2(l+R)}))$, where $r \leftarrow U_{R(k)}$ and f_b is a size-s RDM function generated by $(f_0, f_1, \mathsf{state}) \leftarrow A(1^k, pk)$. Thus, the view of A in $\mathrm{BRDM}_b^{\Pi}(A,k)$ can be described by (pk,h,c,σ), where σ is the internal randomness of A.

By the indistinguishability of the lossy and injective public keys of the lossy trapdoor functions, the view of A is computationally indistinguishable to the experiment where pk is instead generated in the lossy mode. Specifically, we refer to the view of A in the lossy mode as (pk,h,c,σ), where $pk \leftarrow \mathrm{GenLossy}(1^k)$ and h c, σ are generated as above.

We now argue that the view of A in the lossy mode is statistically close to $(pk,h,F_{pk}(U_n),\sigma)$. Recall that in the lossy mode, F_{pk} has range of size at most 2^u. Applying Lemma 5.10 with $\varepsilon = \delta = 2^{-u}$, $\mathcal{F} = \{F_{pk} : \{0,1\}^n \to R\}_{pk \leftarrow \mathrm{GenLossy}(1^k)}$, and $\mathcal{C} = \{(f(r)||r||0^{2(l+R)}) : r \leftarrow U_{R(k)}, f \in \text{ size-}s \text{ RDM function}\}$, we have with probability at least $(1-2^{-u})$ over $h \leftarrow \mathcal{H}$, it holds that for every $pk \leftarrow \mathrm{GenLossy}(1^k)$ and every size-s RDM function f, $\Delta(F_{pk}(h(f(r)||r||0^{2(l+R)})), F_{pk}(U_n)) \leq 2^{-u}$. Note that for such good h and for every lossy public-key pk, the statistical distance between the ciphertext c in the view of A and $F_{pk}(U_n)$ is at most 2^{-u}. It follows that the statistical distance between the view of A in the lossy mode and $(pk,h,F_{pk}(U_n),\sigma)$ is at most $2^{-u} + 2^{-u} \in \mathrm{negl}(k)$.

Putting things together, we showed that for every PPT adversary A and $b \in \{0,1\}$, the view of A in BRDM_b^{Π} is computationally indistinguishable to $(pk,h,F_{pk}(U_n),\sigma)$ where $pk \leftarrow \mathrm{GenLossy}(1^k)$ and $h \leftarrow \mathcal{H}_n$ and σ being the internal randomness of A, which clearly implies $\{\mathrm{BRDM}_0^{\Pi}(A,k)\}_k$ and $\{\mathrm{BRDM}_1^{\Pi}(A,k)\}_k$ are computationally indistinguishable. □

We now turn to prove also circular RDM security of our scheme. To do this, we require the use of a generalized form of the above crooked deterministic extraction lemma that also deals with leakage (just as our "plain" deterministic extraction of leakage-source lemma, lemma 5.3), whose proof can be found in Appendix B.

Lemma 5.12 (Crooked Deterministic Extraction from Bounded Leakage Sources)**.** *Let* $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^n\}$ *be a* t*-wise independent hash function family. Let* $\mathcal{F} = \{f : \{0,1\}^n \to R_f\}$ *be a family of functions where each* $f \in \mathcal{F}$ *has range* R_f *of size* $|R_f| \leq 2^m$*. Let* $\mathcal{G} = \{g : \{0,1\}^n \to \{0,1\}^n\}$ *be a family of functions. Let* $\mathcal{C}$ *be a family of distributions over* $\{0,1\}^n$ *such that every* $X \in \mathcal{C}$ *has min-entropy* $H_\infty(X) \geq k$*. If*

$$\begin{cases} t \geq 2(2m + \log|\mathcal{F}| + \log|\mathcal{G}| + \log|\mathcal{C}| + \log(1/\delta) + 3), \\ m \leq (k - 3\log(1/\varepsilon) - \log t - 5)/2, \end{cases}$$

then with probability at least $(1-\delta)$ *over* $h \leftarrow \mathcal{H}$*, it holds that for every* $f \in \mathcal{F}$*,* $g \in \mathcal{G}$*, and* $X \in \mathcal{C}$*,*

$$\Delta((f(g(X)), f(h(X))), (f(g(X)), f(U_n))) \leq \varepsilon.$$

Lemma 5.13. *The l-bit encryption scheme* $\Pi = (\mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec})$ *constructed in Definition 5.9 is s-bounded circular RDM secure.*

Proof. The proof of circular RDM security follows the same line as the proof of Theorem 5.6 (recall that Theorem 5.6 shows that bounded strong RDM security implies bounded circular RDM security). Note that in the proof of Theorem 5.6, the strong RDM security is only used to show the indistinguishability of the last two hybrids, where the last message $f^q(r_{q-1}, c_{q-1})$ is replaced by 0 and the strong RDM security is used to simulate the first encryption $\mathrm{Enc}_{pk}(f^1(r_q); r_1)$ using randomness leakage $f^1(r_q)$ (whereas the indistinguishability of remaining hybrids relies only on the standard CPA security). In the current context, the message length may be longer than the randomness length so r_q may have no randomness left after leaking $f^1(r_q)$. However, note that in the lossy mode, F_{pk} has a small range and so r_q has sufficient entropy left after conditioning on the first encryption $F_{pk}(h(f^1(r_q)||r_1||0^{2(l+R)}))$. Indeed, indistinguishability of these two hybrids follows in exactly the same way as in the proof of Theorem 5.11, but by relying on Lemma 5.12 instead of Lemma 5.10 (using $F_{pk}(h(f^1(r_q)||r_1||0^{2(l+R)}))$ as leakage). □

5.3 RDM security in the Random Oracle model

In this section we see how we can achieve (full) RDM security in the Random oracle model. Our use of the random oracle model is extremely weak: we do not need to "program it", or "see queries to it". Additionally, security holds even if the attacker may get any inefficient non-uniform advice about the random oracle (as in the model of [47]). We refer to such a model as the "ultra-weak" random oracle model, and omit a formal definition.

Definition 5.14. For every public key encryption scheme $\Pi = (Gen, Enc, Dec)$, define a new encryption scheme $\Pi' = (Gen', Enc', Dec')$ using a random oracle $RO : \{0,1\}^{R'(k)} \to \{0,1\}^{R(k)}$, where R is the length of randomness of Enc, m is the length of messages of Enc, and $R'(k) = R(k) + m(k) + 3k + \log^2 k + 5$ is the length of randomness of Enc'.

- $\mathrm{Gen}'(1^k) := (pk, sk) \leftarrow \mathrm{Gen}(1^k)$. Output (pk, sk).
- $\mathrm{Enc}'(pk, m) :=$ Sample $r \leftarrow U_{R'(k)}$. Output $\mathrm{Enc}(pk, m; RO(r))$.
- $\mathrm{Dec}'(sk, c) :=$ Output $\mathrm{Dec}(sk, c)$.

Theorem 5.15. *For every non-uniformly secure CPA secure public key encryption scheme* Π*,* Π' *as defined in Definition 5.14 is RDM secure in the ultra-weak random oracle model.*

Proof. The theorem follows by essentially identical arguments to the proof of Theorem 5.5, by noting that (1) the above transformation is essentially the transformation defined in Definition 5.4 except that instead of applying h_k, here we apply the random oracle RO for randomness extraction, and (2) a random oracle is a $k^{\log k}$-wise independent family of hash functions, and for any polynomial p and sufficiently large k

$$k^{\log k} > 2(k + p(k) + m(k)).$$

There is just a single point that needs to be addressed: how to deal with attackers that receive (inefficient) non-uniform advice about the random oracle. Note that since the randomness extraction

lemma (i.e., Lemma 5.3) holds in a statistical sense, even conditioned on the "whole" hashfunction (which is the random oracle in this case), it thus also holds conditioned on a function of it (i.e., the non-uniform advice the attacker may get). The only difference is that since we start off with non-uniform attacker, we need to rely on the underlying scheme being non-uniformly secure.

□

Theorem 5.15, combined with Theorem 4.1, show the existence of a task—RDM secure encryption—that can be achieved in the ultra-weak random oracle model (assuming the existence of CPA secure encryption schemes), but cannot be achieved in the plain model. As far as we know, this is the first separation between tasks achievable in such a weak random oracle model, and the plain model.

6 Barriers for RDM Security with Long Messages

An interesting question is whether any encryption schemes can be modified by just performing pre-processing to the randomness (as in Theorem 5.5) to become bounded RDM secure but still handle long messages using short randomness. In this section we demonstrate barriers to using black-box proofs of security to show RDM security of encryption scheme that handle messages that are longer than randomness, based on a computational intractability assumption. More specifically, we present black-box barriers to encrypt schemes with *perfect correctness* and *efficiently recognizable public-keys* that can encrypt long messages using short randomness. This shows that simply performing some processing on the randomness of an encryption scheme (as in the construction used in Theorem 5.5) does not suffice to make *every* encryption scheme resilient to bounded-RDM attacks if requiring that long message can be encrypted using short randomness. For instance, this means that the El-Gamal crypto system [] cannot be modified (by performing pre-processing to the randomness) to become bounded RDM secure for long messages.

Let us start by defining a falsifiable security game (following [39, 30, 26]).

Definition 6.1. A *falsifiable security game* is an efficient random system Γ that on security parameter k interacts with an adversary A and outputs a bit, which we denote by $A(1^k) \leftrightarrow \Gamma(1^k)$. We say Γ is *secure* if for all PPT A, $Pr[A(1^k) \leftrightarrow \Gamma(1^k) = 1]$ is negligible in k where the probability is over the randomness of A and Γ.

Our separation result applies to any black-box reduction (i.e., the reduction only accesses the attacker A as a black-box) that accesses the RDM function f as a black-box. Following the terminology used by [29] (in the context of key-dependent message security), we refer to such reductions as *strongly black-box* reductions. We point out that our proof of security in Section 5.1 is indeed strongly black-box.

Definition 6.2. [29] A *strongly black-box* reduction from the p-bounded RDM security of a public key encryption scheme Π to the security of a falsifiable security game Γ is a PPT oracle aided machine $R^{(\cdot)}$ such that for any adversary A that breaks the p-bounded RDM security of Π, R^A breaks the security of Γ. Additionally, R treats the challenge RDM functions provided by A as a black-box.

More concretely, for any adversary A such that there exists a PPT distinguisher D, polynomial p and infinitely many k such that

$$|Pr[D(BRDM_0^{\Pi}(A,k)) = 1] - Pr[D(BRDM_1^{\Pi}(A,k)) = 1]| > \frac{1}{p(k)}$$

there exists a polynomial p' and infinitely many k' such that

$$Pr[R^A(1^k) \leftrightarrow \Gamma(1^k) = 1] > \frac{1}{p'(k')}$$

Our lower bound relies on the existence of a pseudorandom function family secure against subexponential-size circuits.

Definition 6.3. Let m, r be polynomials. We say that a family of efficient functions $\mathcal{F}_k^{m,r} = \{f_s : \{0,1\}^{r(k)} \to \{0,1\}^{m(k)}; s \in \{0,1\}^k\}_{k\in\mathbb{N}}$ is a *strong pseudorandom function family* if there exists a constant $\varepsilon \in [0,1]$ such that for every distinguisher A of size at most 2^{k^ε}, there exists a negligible function μ such that for every $k \in \mathbb{N}$,

$$|\Pr[s \leftarrow \{0,1\}^k : A^{f_s}(1^k) = 1] - \Pr[A^{RF}(1^k) = 1]| \leq \text{negl}(k)$$

where RF is a random oracle.

It is well-known that the existence of strong pseudorandom function families can be based on the existence of one-way functions with subexponential security [27, 32]

We now have the following theorem.

Theorem 6.4. *Assume the existence of an efficient strong pseudorandom function family* $\mathcal{F}_k^{m,r} = \{f_s : \{0,1\}^{r(k)} \to \{0,1\}^{m(k)}; s \in \{0,1\}^k\}_{k\in\mathbb{N}}$ *for polynomials* m *and* r *such that*

$$m(k) \geq r(k) + \omega(\log k).$$

Then there exists a polynomial p *such that the following holds. Consider any public-key encryption scheme* Π *with perfect correctness and efficiently recognizable public-keys, message length* m *and encryption randomness length* r*, and a secure falsifiable security game* Γ*. Then, there does not exist a* strongly-black-box *reduction from the* p*-bounded RDM security of* Π *to the security of* Γ*.*

We now provide an overview of the proof. Assume we have a reduction R such that, for any adversary A that breaks the p-bounded RDM security of Π, R^A breaks the security of the falsifiable security game Γ. For every *valid public-key*, due to the *perfect correctness* property of Π, we can always construct an inefficient adversary A that breaks the p-bounded RDM security of Π with probability 1 (and that simply outputs $\perp$ in case the ciphertext it receives is invalid, or in case the public-key is invalid); we thus have that R^A, though inefficient, breaks the security of Γ. Furthermore, we are guaranteed that R not only uses A as a black-box but also that the RDM function specified by A is used as a black-box. Now if the the RDM function chosen by A behaves like a random function then its range is "elusive"—that is, the only way to compute any image of the function is to query the RDM function in its pre-image. This allows us to *efficiently* simulate R^A. Given that the reduction must treat the RDM function as a black-box, such queries can be seen by the simulator and knowing these queries allows it to circumvent the inefficiency of A; to efficiently simulate A correctly, we need to be able to efficiently recognize what public-keys are invalid (since such queries are answered $\perp$ by A, even if R has correctly encrypted some message).[7] Now, our simulator is an efficient machine that breaks the security of Γ, and we have a contradiction. However, we require that the RDM function behaves as a random function, and a random function

[7]The reason that we have A answer $\perp$ to queries on invalid public keys is that the perfect correctness property only holds for valid public-keys; without the perfect correctness property we are not guaranteed that the message recoverd by our simulator is the same as the one recoverd by the inefficient attacker.

itself does not have a polynomial bound on its size. To this end, we use a pseudorandom function as the RDM function. Still, A is inefficient and could be used by the reduction to distinguish the pseudorandom function from a random function, in which case our simulation fails. To circumvent this, we use a pseudorandom function secure against subexponential sized adversaries and simply use this function with a sufficiently long seed to ensure security against A.

Proof. We start by choosing the polynomial p. Let $t(\cdot)$ be a polynomial time bound on the complexity of $\mathcal{F}_k^{m,r}$, and $\varepsilon \in (0,1)$ be a constant such that the indistinguishability of $\mathcal{F}_k = \mathcal{F}_k^{m,r}$ holds against size 2^{k^ε} distinguishers. Let $k' = r^{2/\varepsilon}$, and set $p(k) \triangleq t(k')$.

Now, assume for contradiction that there exists a public-key encryption scheme Π with perfect correctness, message length m, and encryption randomness length r, a secure falsifiable security game Γ, and a strongly-black-box reduction R such that R reduces the p-bounded RDM security of Π to the security of Γ. Namely, for every adversary A that breaks the p-bounded RDM security of Π, R^A breaks the security of Γ while using both A and the RDM functions chosen by A as black-boxes. In this case, we shall show that the security of Γ can be broken by an efficient adversary, a contradiction to the assumed security of Γ.

Towards this goal, we first define an (inefficient) "canonical" adversary A that breaks the p-bounded RDM security of Π . It follows that R^A breaks the security of Γ with non-negligible probability. We then construct an *efficient* "meta-reduction" M that internally emulates R^A efficiently and breaks the security of Γ with essentially the same probability, which completes the proof.

We now construct the (inefficient) canonical adversary A for breaking the p-bounded RDM security of Π. For every $s \in \{0,1\}^{k'}$ define A_s as follows:

- On input $(1^k, pk)$, A_s checks if pk is a valid public-key given the security parameter k; if not it outputs $\perp$. Otherwise, it chooses the RDM function $f = f_s \in \mathcal{F}_{k'}$ and output (f, state) where $\mathsf{state} = (f, pk)$.[8]
- On input the challenge ciphertext (c^*, state) (for a valid public-key pk), A_s checks for all $r \in \{0,1\}^{r(k)}$ whether $c^* = \mathrm{Enc}_{pk}(f(r); r)$. If there is such an r then output a bit $b' = 0$ else output $b' = 1$.

It is easy to see that for every s, A_s breaks the p-bounded RDM security of Π. Hence for every s, R^{A_s} breaks the security of Γ. Hence, this also holds true for a random s. We define R^A as R^{A_s} where $s \leftarrow \{0,1\}^k$ and we have that R^A breaks the security of Γ. More concretely, there exists a polynomial p' such that for infinitely many k

$$\Pr[\Gamma(1^k) \leftrightarrow R^A(1^k) = 1] > \frac{1}{p'(k)}$$

We proceed to construct an *efficient* meta-reduction M that interacts with Γ by internally emulating R^A efficiently and forwarding R's messages to Γ and back. Notice that emulating R^A efficiently is easy except that in order to generate the message b', we needs to enumerate over all randomness r, which takes time $O(2^r)$ and is inefficient. Instead, M maintains a list of queries L made by R to the corresponding RDM function f (recall that R treats f as an oracle, and thus M sees that queries of R to f) and on input the challenge ciphertext c^* generates b' by checking for all $q \in L$ whether $\mathrm{Enc}_{pk}(f(q); q) = c^*$. If there is such a q, M sets $b' = 0$ else it sets $b' = 1$. It is

[8] Technically, f maps $r(k')$-bit strings to $m(k')$-bit strings, but we can simply shrink the domain and the range to the proper size.

not hard to see that by simulating b' this way, M can emulate the execution of R^A in polynomial time.

To show a contradiction it remains to show that M breaks the security of Γ with non-negligible probability. Let $\mathsf{Hybrid}_{\mathsf{ineff}} = \Gamma(1^k) \leftrightarrow R^A(1^k)$ and $\mathsf{Hybrid}_{\mathsf{eff}} = \Gamma(1^k) \leftrightarrow M(1^k)$. We prove that, for every k, $|Pr[\mathsf{Hybrid}_{\mathsf{ineff}} = 1] - Pr[\mathsf{Hybrid}_{\mathsf{eff}} = 1]| \leq \epsilon(k)$ for some negligible ϵ via a hybrid argument.

Let A'_f be the adversary obtained by replacing the f_s in A_s with the function f. More precisely, A'_f behaves just as A_s except that on input $(pk, 1^k)$, A'_f outputs f as the RDM function. We define $R^{A'}$ as $R^{A'_{RF}}$ here RF is a random function from $\{0,1\}^{r(k)}$ to $\{0,1\}^{m(k)}$. Similarly M' behaves just as M except that it uses a random function RF instead of a pseudorandom function. We define hybrids $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}} = \Gamma(1^k) \leftrightarrow R^{A'}(1^k)$ and $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}} = \Gamma(1^k) \leftrightarrow M'(1^k)$

Claim 6.5. $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ *and* $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$ *are statistically close.*

Proof. In both $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ and $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$, the RDM function f is a random function RF. Note that in $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ a ciphertext query c^* is answered by using the canonical attacker A' by finding the unique (by perfect correctness) message $m = f(r)$ encrypted as $c^* = \mathrm{Enc}(f(r); r)$, and is answered as 1 if such an r exists, and as 0 otherwise. On the other hand, in $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$, the meta-reduction M' provides the correct message $m = f(r)$ if and only if R has queried the RDM function on an input r such that $c^* = \mathrm{Enc}(f(r); r)$; if $c \neq \mathrm{Enc}(f(r), r)$ for any r, then M' outputs 0 just as A' does. To show that $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ and $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$ are statistically close, it suffices to show the probability that R asks some ciphertext query $c^* = \mathrm{Enc}(f(r); r)$ without first querying the RDM function on r, is negligible. Since R makes at most a polynomial number of ciphertext queries, by a union bound, it suffices to bound the probability that R makes a single query $c^* = \mathrm{Enc}(f(r); r)$ without having queried the RDM function on r. By perfect correctness, we have that for any fixed value c, and any fixed random string r, the probability over the random function RF that $c = \mathrm{Enc}(f(r); r))$ is upper bounded by $2^{-m(k)}$. It follows by a union bound that for every fixed string c^*, the probability over the choice of the random function RF that there exists some r such that $c = \mathrm{Enc}(f(r); r)$ is upper bounded by $2^r/2^m$, which is negligible in k. The same holds for any choice of random variable c^*. This concludes the proof of the claim. □

Claim 6.6. $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$ *and* $\mathsf{Hybrid}_{\mathsf{eff}}$ *are statistically close.*

Proof. Both $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{eff}}$ and $\mathsf{Hybrid}_{\mathsf{eff}}$ are PPT and differ only in the usage of a random function RF in one and a function $f \leftarrow \mathcal{F}_k$ in the other. Hence the claim follows directly from the pseudorandomness property of $\mathcal{F}_k$ and the fact that Γ outputs a single bit. □

Claim 6.7. $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ *and* $\mathsf{Hybrid}_{\mathsf{ineff}}$ *are statistically close.*

Proof. Again $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ and $\mathsf{Hybrid}_{\mathsf{ineff}}$ differ only in the usage of a random function RF in one and a function $f \leftarrow \mathcal{F}_k$ in the other. However both run in time at most $p(k)2^{r(k)}$ which is at most $2^{k'^{\epsilon}}$. By the strong pseudorandomness of $\mathcal{F}_k$ and the fact that Γ outputs a single bit we have that $\mathsf{Hybrid}^{\mathsf{RF}}_{\mathsf{ineff}}$ and $\mathsf{Hybrid}_{\mathsf{ineff}}$ are statistically close.. □

The above three claims complete the hybrid argument and consequently the proof. □

7 Randomness Dependent Security in Commitment Schemes

In this section we consider RDM security for commitment schemes. To contrast our result with those for encryption schemes, we consider a setting as close as possible to the encryption case; that is, we consider non-interactive commitment schemes in the Common Reference String (CRS) model. In essence, non-interactive commitment schemes in the CRS model differ from encryption schemes in two ways: 1) for commitment schemes, there is not necessarily an efficient way to decrypt messages, and 2) for commitment schemes, on the other hand, we require a binding property.

We proceed to formally defining commitment schemes in the CRS model.

Definition 7.1. An *l-bit (non-interactive) commitment scheme* in the CRS model consists of a tuple $(\mathrm{Gen}, \mathrm{Com})$ of PPT algorithms where (i) Gen takes a security parameter 1^k as input and generates a common reference string $crs \leftarrow \mathrm{Gen}(1^k)$, (ii) Com takes crs and a message m in a message space $\{0,1\}^{l(k)}$ as input and generates a commitment $c \leftarrow \mathrm{Com}(crs, m)$.

Definition 7.2. We say an l-bit non-interactive commitment scheme in the CRS model $(\mathrm{Gen}, \mathrm{Com})$ is *secure* if the following conditions hold:

- *Computational Hiding*: For every probabilistic polynomial time adversary $A = (A_1, A_2)$, the ensembles $\{\mathrm{IND}_0(A,k)\}_k$ and $\{\mathrm{IND}_1(A,k)\}_k$ are computationally indistinguishable, where

$$\begin{aligned}\mathrm{IND}_b(A,k) := \quad & crs \leftarrow \mathrm{Gen}(1^k)\\ & (m_0, m_1, \mathsf{state}) \leftarrow A_1(1^k, crs)\\ & c \leftarrow \mathrm{Com}(crs, m_b)\\ & o \leftarrow A_2(c, \mathsf{state})\\ & \text{Output } o\end{aligned}$$

- *Computational Binding*: For every probabilistic polynomial time adversary A, there exists a negligible function μ such that

$$\begin{aligned}\Pr[crs \leftarrow \mathrm{Gen}(1^k), (m_0, m_1, r_0, r_1) \leftarrow A(1^k, crs) : m_0 \neq m_1\\ \wedge \mathrm{Com}(crs, m_0; r_0) = \mathrm{Com}(crs, m_1; r_1) = c] \leq \mu(k)\end{aligned}$$

We may now define RDM security in exactly the same way as RDM security for encryption schemes.

Definition 7.3. A secure l-bit commitments scheme in the CRS model $(\mathrm{Gen}, \mathrm{Com})$ is *randomness-dependent message secure (RDM-secure))* if for every PPT adversary $A = (A_1, A_2)$, the ensembles $\{\mathrm{RDM}_0(A,k)\}_{k\in\mathbb{N}}$ and $\{\mathrm{RDM}_1(A,k)\}_{k\in\mathbb{N}}$ are computationally indistinguishable where

$$\begin{aligned}\mathrm{RDM}_b(A,k) := \quad & crs \leftarrow \mathrm{Gen}(1^k)\\ & (f_0, f_1, \mathsf{state}) \leftarrow A_1(1^k, crs)\\ & r \leftarrow U_{|r|}\\ & c \leftarrow \mathrm{Com}(crs, f_b(r); r)\\ & o \leftarrow A_2(c, \mathsf{state})\\ & \text{Output } o\end{aligned}$$

The RDM functions f_b are represented as circuits from $\{0,1\}^{|r|}$ to $\{0,1\}^{l(k)}$

7.1 Lower bound

We first remark that our impossibility result for "full" RDM security for encryption schemes directly extends also to commitment schemes.

Theorem 7.4. *No 1-bit commitment scheme in the CRS model* $(\mathrm{Gen}, \mathrm{Com})$ *is RDM-secure.*

Proof. Note that in the proof of Theorem 4.1, the only property of the encryption scheme that is used to violate RDM security is Equation 1; that is;

$$\Pr_{pk,r}[\mathrm{Enc}_{pk}(0,r) \neq \mathrm{Enc}_{pk}(1,r)] \geq 1 - \mathrm{negl}(k). \tag{2}$$

Let us now argue that the analog of this property also holds for commitment schemes; that is

$$\Pr_{crs,r}[\mathrm{Com}(crs,0;r) \neq \mathrm{Com}(crs,1;r)] \geq 1 - \mathrm{negl}(k)$$

This holds since if committing to 0 and 1 (using the same randomness) results in identical commitments, with non-negligible probability, then even the honest sender can later decommit to both 0 and 1 with non-negligible probability and break the computational binding property. □

Remark. We remark that the proof also extends to *interactive* commitments. Recall that in the non-interactive case, RDM attacker applied a uniformly selected hashfunction to the commitment and selected the bit v to commit to so as to bias the output of the hashfunction. In the interactive setting, the RDM attacker acts as an honest receiver using some uniformly selected randomness r_R, and next picks RDM functions that emulate an interaction between the honest sender (using his actual randomness) and the honest receiver using randomness r_R, applies a hashfunction to the *complete transcript* of the interaction, and finally picks the bit b so as to bias the output of the hashfunction. It follows using the same argument as above that such an attacker violates RDM security, unless computational binding of the commitment is broken (with non-negligible probability over the choice of the CRS, and the receiver's randomness).

7.2 Upper bounds

Since "full" RDM security is impossible to achieve, as with encryption schemes, we define RDM security with respect to *a priori* bounded RDM functions.

Definition 7.5. For every $s : \mathbb{N} \to \mathbb{N}$ a secure l-bit (non-interactive) commitment scheme in the CRS model $(\mathrm{Gen}, \mathrm{Com})$ is *s-bounded RDM secure* if for every PPT adversary $A = (A_1, A_2)$, the ensembles $\{\mathrm{BRDM}_0(A,k)\}_{k\in\mathbb{N}}$ and $\{\mathrm{BRDM}_1(A,k)\}_{k\in\mathbb{N}}$ are computationally indistinguishable where

$$\begin{aligned}
\mathrm{BRDM}_b(A,k) := \quad & crs \leftarrow \mathrm{Gen}(1^k) \\
& (f_0, f_1, \mathsf{state}) \leftarrow A_1(1^k, crs) \\
& r \leftarrow U_{|r|} \\
& c \leftarrow \mathrm{Com}(crs, f_b(r); r) \\
& o \leftarrow A_2(c, \mathsf{state}) \\
& \text{Output } o
\end{aligned}$$

and $f : \{0,1\}^{|r|} \to \{0,1\}^{l(k)}$ is a function computed by a circuit of size at most $s(k)$.

We note that exactly same construction as that for bounded RDM security for encryption schemes works for also non-interactive commitments in the CRS model.

Definition 7.6. For every polynomial s and (non-interactive) l-bit commitment scheme in the CRS model (Gen, Com), define a new l-bit commitment scheme in the CRS model $(\text{Gen}', \text{Com}')$ as

- $\text{Gen}'(1^k) := crs \leftarrow \text{Gen}(1^k), h_k \leftarrow \mathcal{H}_k$ where $\mathcal{H}_k = \{h_k : \{0,1\}^{R'(k)} \rightarrow \{0,1\}^{R(k)}\}$ is a $t(k)$-wise independent family of hash functions where $R(.)$ is the length of the randomness of Com, $R'(.)$ is the length of the randomness of Com$'$,

  $$t(k) \geq 2(k + s(k) + l(k))$$

  and

  $$R'(k) = l(k) + R(k) + 3k + \log t(k) + 5$$

  Output (crs, h_k).

- $\text{Com}'((crs, h_k), m) := r \leftarrow U_{R'(k)}$. Output $\text{Com}(crs, m; h_k(r))$.

Theorem 7.7. *Let s be a polynomial. Let (Gen, Com) be a secure l-bit commitment scheme in the CRS model, and $(\text{Gen}', \text{Com}')$ be the transformed commitment scheme obtained from Definition 7.6. Then, $(\text{Gen}', \text{Com}')$ is s-bounded RDM secure.*

Proof. Note that the above construction for bounded RDM-secure commitment schemes is the exactly the same as the one we used for encryption schemes, except that public key is replaced by the CRS and the encryption scheme is replaced by a commitment commitment scheme (and thus, there may not exist an efficient decryption procedure). However the proof of Theorem 5.5 does not rely on the existence of an efficient decryption procedure, and thus exactly the same proof applies also for commitments schemes. That is, it follows that the transformed commitment scheme is bounded RDM-secure.

It remains to show the computational binding property of $(\text{Gen}', \text{Com}')$. Note that an opening (m_0, r_0, m_1, r_1) for $(\text{Gen}', \text{Com}')$ constitutes an opening $(m_0, h_k(r_0), m_1, h_k(r_1))$ for the original scheme. Hence computational binding of $(\text{Gen}', \text{Com}')$ follows from the computational binding property of of (Gen, Com). □

We finally note that (in contrast to Theorem 6.4), for the case of commitments, it is easy to construct a bounded RDM secure commitments schemes that use "short" randomness to commit to "long" messages.

Definition 7.8. For every polynomial m and l-bit commitment scheme in the CRS model (Gen, Com) define a m-bit commitment scheme in the CRS model $(\text{Gen}', \text{Com}')$ as follows

- We define Gen$'$ as sampling $h \leftarrow \mathcal{H}_k$ and appending h to the common reference string generated by running Gen, where $\{\mathcal{H}_k\}_k$ is a Collision Resistant Hash function family with $\mathcal{H}_k$ from $\{0,1\}^{m(k)}$ to $\{0,1\}^{l(k)}$.

- We define Com$'$ as just applying h its m-bit long message and using Com on the shortened l-bit message.

Theorem 7.9. *Let s and m be polynomials. Let (Gen, Com) be an l-bit commitment scheme in the CRS model. Assuming there exists a Collision Resistant Hash function family $\{\mathcal{H}_k\}_k$ where $\mathcal{H}_k$ is from $\{0,1\}^{m(k)}$ to $\{0,1\}^{l(k)}$ and $h \in \mathcal{H}_k$ is computed by a circuit of size $p(k)$, the m-bit commitment scheme in the CRS model $(\text{Gen}', \text{Com}')$ obtained from Definition 7.8 is s-bounded RDM secure if (Gen, Com) is $(s + p)$-bounded RDM secure.*

Proof. It follows by the collision resistant property of the hash function that the transformed scheme remains computational binding. More formally, assume for contradiction there exists a PPT A such that for infinitely many k

$$Pr[crs \leftarrow \mathrm{Gen}'(1^k), (m_0, m_1, r_0, r_1) \leftarrow A(1^k, crs) : m_0 \neq m_1 \\ \wedge \mathrm{Com}'(crs, m_0; r_0) = \mathrm{Com}$$

We construct PPT B that contradicts the computational binding property of $(\mathrm{Gen}, \mathrm{Com})$ as follows: On input $(1^k, crs)$, B samples $h \leftarrow \mathcal{H}_k$, runs $(m_0, m_1, r_0, r_1) \leftarrow A(1^k, (crs, h))$ and outputs $(h(m_0), h(m_1), r_0, , r_1)$. By the collision resistance of $\mathcal{H}_k$ we have that the probability $h(m_0) = h(m_1)$ is negligible, otherwise B has efficiently found a collision on a $h \leftarrow \mathcal{H}_k$. Hence we have that B breaks the computational binding property of $(\mathrm{Gen}, \mathrm{Com})$ with non-negligible probability and we have a contradiction.

The RDM hiding property of $(\mathrm{Gen}', \mathrm{Com}')$ follows by that of $(\mathrm{Gen}, \mathrm{Com})$. Any RDM query $f(.)$ on the new scheme Com' can be simulated by the slightly more "complex" RDM query $h(f(.))$ on the original scheme Com. If $f(.)$ was evaluated by a circuit of size at most $s(k)$ then $h(f(.))$ is evaluated by a circuit of size at most $s(k) + p(k)$. Hence assuming a PPT attacker that breaks the s-bounded RDM hiding property of $(\mathrm{Gen}', \mathrm{Com}')$ we can construct a PPT attacker that breaks the $s + p$-bounded RDM hiding property of $(\mathrm{Gen}, \mathrm{Com})$. Com' uses randomness of the same length as Com. This concludes the proof. □

References

[1] Tolga Acar, Mira Belenkiy, Mihir Bellare, and David Cash. Cryptographic agility and its relation to circular encryption. In *EUROCRYPT*, pages 403–422, 2010.

[2] Adi Akavia, Shafi Goldwasser, and Vinod Vaikuntanathan. Simultaneous hardcore bits and cryptography against memory attacks. In *TCC*, pages 474–495, 2009.

[3] Joël Alwen, Yevgeniy Dodis, and Daniel Wichs. Leakage-resilient public-key cryptography in the bounded-retrieval model. In *CRYPTO*, pages 36–54, 2009.

[4] Benny Applebaum, Danny Harnik, and Yuval Ishai. Semantic security under related-key attacks and applications. In *ICS*, pages 45–60, 2011.

[5] Boaz Barak. How to go beyond the black-box simulation barrier. In *FOCS*, pages 106–115, 2001.

[6] Boaz Barak, Iftach Haitner, Dennis Hofheinz, and Yuval Ishai. Bounded key-dependent message security. In *EUROCRYPT*, pages 423–444, 2010.

[7] M. Bellare and J. Rompel. Randomness-efficient oblivious sampling. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 276–287, 1994.

[8] Mihir Bellare, Alexandra Boldyreva, and Adam O'Neill. Deterministic and efficiently searchable encryption. In *CRYPTO*, pages 535–552, 2007.

[9] Mihir Bellare, Alexandra Boldyreva, and Adriana Palacio. An uninstantiable random-oracle-model scheme for a hybrid-encryption problem. In *EUROCRYPT*, pages 171–188, 2004.

[10] Mihir Bellare, Zvika Brakerski, Moni Naor, Thomas Ristenpart, Gil Segev, Hovav Shacham, and Scott Yilek. Hedged public-key encryption: How to protect against bad randomness. In *ASIACRYPT*, pages 232–249, 2009.

[11] Mihir Bellare, Marc Fischlin, Adam O'Neill, and Thomas Ristenpart. Deterministic encryption: Definitional equivalences and constructions without random oracles. In *CRYPTO*, pages 360–378, 2008.

[12] Mihir Bellare and Sriram Keelveedhi. Authenticated and misuse-resistant encryption of key-dependent data. In *CRYPTO*, pages 610–629, 2011.

[13] John Black, Phillip Rogaway, and Thomas Shrimpton. Encryption-scheme security in the presence of key-dependent messages. In *Revised Papers from the 9th Annual International Workshop on Selected Areas in Cryptography*, SAC '02, pages 62–75, 2003.

[14] Alexandra Boldyreva, Serge Fehr, and Adam O'Neill. On notions of security for deterministic encryption, and efficient constructions without random oracles. In *CRYPTO*, pages 335–359, 2008.

[15] Dan Boneh, Shai Halevi, Mike Hamburg, and Rafail Ostrovsky. Circular-secure encryption from decision diffie-hellman. In *Proceedings of the 28th Annual conference on Cryptology: Advances in Cryptology*, CRYPTO 2008, pages 108–125, 2008.

[16] Zvika Brakerski and Shafi Goldwasser. Circular and leakage resilient public-key encryption under subgroup indistinguishability. In *Proceedings of the 30th annual conference on Advances in cryptology*, CRYPTO'10, pages 1–20, 2010.

[17] Zvika Brakerski, Shafi Goldwasser, and Yael Tauman Kalai. Black-box circular-secure encryption beyond affine functions. In *Proceedings of the 8th conference on Theory of cryptography*, TCC'11, pages 201–218, 2011.

[18] Zvika Brakerski, Yael Tauman Kalai, Jonathan Katz, and Vinod Vaikuntanathan. Overcoming the hole in the bucket: Public-key cryptography resilient to continual memory leakage. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 501–510, 2010.

[19] Jan Camenisch, Nishanth Chandran, and Victor Shoup. A public key encryption scheme secure against key dependent chosen plaintext and adaptive chosen ciphertext attacks. In *Proceedings of the 28th Annual International Conference on Advances in Cryptology: the Theory and Applications of Cryptographic Techniques*, EUROCRYPT '09, pages 351–368, 2009.

[20] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques: Advances in Cryptology*, EUROCRYPT '01, pages 93–118, 2001.

[21] Ran Canetti, Oded Goldreich, and Shai Halevi. On the random-oracle methodology as applied to length-restricted signature schemes. In *TCC*, pages 40–57, 2004.

[22] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557–594, 2004.

[23] David Cash, Matthew Green, and Susan Hohenberger. New definitions and separations for circular security. In *Public Key Cryptography*, pages 540–557, 2012.

[24] Yevgeniy Dodis and Adam Smith. Correcting errors without leaking partial information. In *STOC*, pages 654–663, 2005.

[25] Stefan Dziembowski and Krzysztof Pietrzak. Leakage-resilient cryptography. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 293–302, 2008.

[26] Craig Gentry and Daniel Wichs. Separating succinct non-interactive arguments from all falsifiable assumptions. In *STOC*, pages 99–108, 2011.

[27] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions (extended abstract). In *FOCS*, pages 464–479, 1984.

[28] Shafi Goldwasser and Yael Tauman Kalai. On the (in)security of the fiat-shamir paradigm. In *FOCS*, pages 102–113, 2003.

[29] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In *Proceedings of the 6th Theory of Cryptography Conference on Theory of Cryptography*, TCC '09, pages 202–219, Berlin, Heidelberg, 2009. Springer-Verlag.

[30] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In *Proceedings of the 6th Theory of Cryptography Conference on Theory of Cryptography*, TCC '09, pages 202–219, 2009.

[31] Shai Halevi and Hugo Krawczyk. Security under key-dependent inputs. In *Proceedings of the 14th ACM conference on Computer and communications security*, CCS '07, pages 466–475, 2007.

[32] Johan Håstad, Russell Impagliazzo, Leonid Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28:12–24, 1999.

[33] Brett Hemenway and Rafail Ostrovsky. Building injective trapdoor functions from oblivious transfer. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:127, 2010.

[34] Dennis Hofheinz and Dominique Unruh. Towards key-dependent message security in the standard model. In *Proceedings of the theory and applications of cryptographic techniques 27th annual international conference on Advances in cryptology*, EUROCRYPT'08, pages 108–126, 2008.

[35] Susan Hohenberger, Allison B. Lewko, and Brent Waters. Detecting dangerous queries: A new approach for chosen ciphertext security. In *EUROCRYPT*, pages 663–681, 2012.

[36] Ueli M. Maurer, Renato Renner, and Clemens Holenstein. Indifferentiability, impossibility results on reductions, and applications to the random oracle methodology. In *TCC*, pages 21–39, 2004.

[37] Silvio Micali and Leonid Reyzin. Physically observable cryptography (extended abstract). In *TCC*, pages 278–296, 2004.

[38] Steven Myers and Abhi Shelat. Bit encryption is complete. In *FOCS*, pages 607–616, 2009.

[39] Moni Naor. On cryptographic assumptions and challenges. In *CRYPTO*, pages 96–109, 2003.

[40] Moni Naor and Gil Segev. Public-key cryptosystems resilient to key leakage. In *CRYPTO*, pages 18–35, 2009.

[41] Jesper Buus Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In *CRYPTO*, pages 111–126, 2002.

[42] Rafael Pass. On deniability in the common reference string and random oracle model. In *CRYPTO*, pages 316–337, 2003.

[43] Rafael Pass, Alon Rosen, and Wei lung Dustin Tseng. Public-coin parallel zero-knowledge for np, 2011.

[44] Chris Peikert, Vinod Vaikuntanathan, and Brent Waters. A framework for efficient and composable oblivious transfer. In *CRYPTO*, pages 554–571, 2008.

[45] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. *SIAM J. Comput.*, 40(6):1803–1844, 2011.

[46] L. Trevisan and S. Vadhan. Extracting randomness from samplable distributions. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 32–42, 2000.

[47] Dominique Unruh. Random oracles and auxiliary input. In *CRYPTO*, pages 205–223, 2007.

[48] Salil Vadhan. Pseudorandomness. `http://people.seas.harvard.edu/~salil/pseudorandomness/`, 2011.

A Proof of Deterministic Extraction Lemma (Lemma 5.3)

Here we present the proof for Lemma 5.3. We will need the following lemmas. The first is from [46] and is via a standard application of the t-moment method.

Lemma A.1 ([46])**.** *Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ be a t-wise independent hash function family, and let X be a distribution over $\{0,1\}^n$ with min-entropy $H_\infty(X) \geq k$. If $m \leq k - 2\log(1/\varepsilon) - \log t - 2$, then*

$$\Pr_{h \leftarrow \mathcal{H}}[\Delta(h(X), U_m) \geq \varepsilon] \leq 2^{-(t/2-m-3)}.$$

The following is a standard simple lemma and the proof can be found in [48].

Lemma A.2. *Let $X = (X_1, X_2)$ be a distribution over $\{0,1\}^n = \{0,1\}^{n_1} \times \{0,1\}^{n_2}$ with min-entropy $H_\infty(X) \geq n - \Delta$. For every $\varepsilon \in (0,1)$, it holds that with probability at least $1 - \varepsilon$ over $x_1 \leftarrow X_1$, $H_\infty(X_2|_{X_1=x_1}) \geq n_2 - \Delta - \log(1/\varepsilon)$.*

Lemma A.3 (Lemma 5.3 restated)**.** *Let $\mathcal{F} = \{f : \{0,1\}^n \to \{0,1\}^\ell\}$ be a class of (leakage) functions. Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^m\}$ be a t-wise independent hash function family. If*

$$\begin{cases} t \geq 2(m + \ell + \log|\mathcal{F}| + \log(1/\delta) + 3), \\ m \leq n - \ell - 3\log(1/\varepsilon) - \log t - 5, \end{cases}$$

then with probability at least $(1 - \delta)$ over $h \leftarrow \mathcal{H}$, it holds that for every $f \in \mathcal{F}$,

$$\Delta((f(U_n), h(U_n)), (f(U_n), U_m)) \leq \varepsilon.$$

Proof. Let $k = n-\ell-\log(1/\varepsilon)-1$. Consider any $f \in \mathcal{F}$. By applying Lemma A.2 to the distribution $(f(U_n), U_n)$ with the parameter ε set to be $\varepsilon/2$, we obtain that with probability at least $(1-\varepsilon/2)$ over $z \leftarrow f(U_n)$, the distribution $X_{f,z} \triangleq U_n|_{f(U_n)=z}$ has min-entropy $H_\infty(X) \geq k$.

Noting that $m \leq k - 2\log(1/(\varepsilon/2)) - \log t - 2$, Lemma A.1 implies that for every X over $\{0,1\}^n$ with $H_\infty(X) \geq k$, with probability at least $1-2^{-t/2}$ over $h \leftarrow \mathcal{H}$, $\Delta(h(X), U_m) \leq \varepsilon/2$. By an union bound over $f \in \mathcal{F}$ and $z \in \{0,1\}^\ell$ such that the corresponding $H_\infty(X_{f,z}) \geq k$, it follows that with probability at least

$$1 - 2^{-t/2} \cdot |\mathcal{F}| \cdot 2^\ell \geq 1 - \delta$$

over $h \leftarrow \mathcal{H}$, for every $f \in \mathcal{F}$ and $z \in \{0,1\}^\ell$ such that the corresponding $H_\infty(X_{f,z}) \geq k$, we have $\Delta(h(X_{f,z}), U_m) \leq \varepsilon/2$. For such $h \in \mathcal{H}$, we have for every $f \in \mathcal{F}$,

$$\begin{aligned}
\Delta((f(U_n), h(U_n)), (f(U_n), U_m)) &= \mathbb{E}_{z \leftarrow f(U_n)}[\Delta(h(X_{f,z}), U_m)] \\
&\leq \Pr_{z \leftarrow f(U_n)}[H_\infty(X_{f,z}) \geq k] \cdot (\varepsilon/2) + \Pr_{z \leftarrow f(U_n)}[H_\infty(X_{f,z}) < k] \cdot 1 \\
&\leq \varepsilon/2 + \varepsilon/2 = \varepsilon,
\end{aligned}$$

as desired. □

B Proofs of Crooked Version Lemmas

Let us first recall a tail bound for t-wise independent random variables:

Lemma B.1 ([7])**.** *Let $t > 4$ be an even integer. Suppose $Y_1, \ldots, Y_K$ are t-wise independent random variables taking value in $[0,1]$. Let $Y = \sum_i Y_i$, $\mu = \mathbb{E}[Y]/K$, and $A > 0$. Then,*

$$\Pr[|Y - K\mu| > A] \leq 8 \cdot \left(\frac{tK\mu + t^2}{A^2}\right)^{t/2}.$$

The following lemma is a crooked version of Lemma A.1 and is proved in a similar way.

Lemma B.2. *Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^n\}$ be a t-wise independent hash function family, $f : \{0,1\}^n \to R$ be an arbitrary function with range R of size $|R| \leq 2^m$. Let X be a distribution over $\{0,1\}^n$ with min-entropy $H_\infty(X) \geq k$. If $m \leq k - 2\log(1/\varepsilon) - \log t - 2$, then*

$$\Pr_{h \leftarrow \mathcal{H}}[\Delta(f(h(X)), f(U_n)) \geq \varepsilon] \leq 2^{-(t/2-m-3)}.$$

Proof. W.l.o.g., we can assume that $|R| = 2^m$ and X is a flat source, i.e., X is uniform over a subset $S \subset \{0,1\}^n$ of size $|S| = 2^k$. Let $M = 2^m$ and $K = 2^k$. Fix a $z \in R$, and let $\mu_z = \Pr[f(U_n) = z]$. Note that for every $x \in \{0,1\}^n$, $\Pr[f(H(x)) = z] = \mu_z$ and $\{f(H(x))\}_{x \in \{0,1\}^n}$ are t-wise independent. By the tail bound for t-wise independent random variables, with probability at least $1 - 2^{-(t/2-3)}$ over $h \leftarrow \mathcal{H}$,

$$|\Pr[f(h(X)) = z] - \mu_z| \leq \begin{cases} \varepsilon\mu_z & \text{if } \mu_z \geq 1/M, \\ \varepsilon/M & \text{o.w.} \end{cases} \tag{3}$$

Indeed, let Y_x be an indicator random variable such taht $Y_x = 1$ iff $f(H(x)) = z$, and let $Y = \sum_{x \in S} Y_x$. By definition, $\mathbb{E}[Y] = K\mu_z$. If $\mu_z \geq 1/M$, applying Lemma B.1 with $A = \varepsilon K \mu_z$, we have

$$\Pr_{h \leftarrow \mathcal{H}}[|\Pr[f(h(X)) = z] - \mu_z| > \varepsilon\mu_z] = \Pr[|Y - K\mu_z| \geq \varepsilon K\mu_z] \leq 8 \cdot \left(\frac{tK\mu_z + t^2}{(\varepsilon K\mu_z)^2}\right)^{t/2} \leq 2^{-(t/2-3)}.$$

If $\mu_z < 1/M$, Lemma B.1 with $A = \varepsilon K/M$ says

$$\Pr_{h \leftarrow \mathcal{H}}\left[\left|\Pr[f(h(X)) = z] - \mu_z\right| > \frac{\varepsilon}{M}\right] = \Pr\left[|Y - K\mu_z| \geq \frac{\varepsilon K}{M}\right] \leq 8 \cdot \left(\frac{tK\mu_z + t^2}{(\varepsilon K/M)^2}\right)^{t/2} \leq 2^{-(t/2-3)}.$$

By an union bound over $z \in R$, with probability at least $1 - 2^{-(t/2-m-3)}$ over $h \leftarrow \mathcal{H}$, Eq.(3) holds for every $z \in R$. For such good h,

$$\begin{aligned}\Delta(f(h(X)), f(U_n)) &= \frac{1}{2}\sum_{z \in R}\left|\Pr[f(h(X)) = z] - \mu_z\right| \\ &\leq \frac{1}{2}\sum_{z \in R}\left(\varepsilon\mu_z + \frac{\varepsilon}{M}\right) \\ &= \varepsilon\end{aligned}$$

□

Lemma B.3 (Lemma 5.10 restated)**.** *Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^n\}$ be a t-wise independent hash function family. Let $\mathcal{F} = \{f : \{0,1\}^n \to R_f\}$ be a family of functions where each $f \in \mathcal{F}$ has range R_f of size $|R_f| \leq 2^m$. Let $\mathcal{C}$ be a family of distributions over $\{0,1\}^n$ such that every $X \in \mathcal{C}$ has min-entropy $H_\infty(X) \geq k$. If*

$$\begin{cases} t \geq 2(m + \log|\mathcal{F}| + \log|\mathcal{C}| + \log(1/\delta) + 3), \\ m \leq k - 2\log(1/\varepsilon) - \log t - 2, \end{cases}$$

then with probability at least $(1-\delta)$ over $h \leftarrow \mathcal{H}$, it holds that for every $f \in \mathcal{F}$ and every $X \in \mathcal{C}$,

$$\Delta(f(h(X)), f(U_n)) \leq \varepsilon.$$

Proof. (Sketch) The lemma follows immediately by Lemma B.2 and an application of union bound over $\mathcal{F}$ and $\mathcal{C}$. □

Lemma B.4 (Lemma 5.12 restated)**.** *Let $\mathcal{H} = \{h : \{0,1\}^n \to \{0,1\}^n\}$ be a t-wise independent hash function family. Let $\mathcal{F} = \{f : \{0,1\}^n \to R_f\}$ be a family of functions where each $f \in \mathcal{F}$ has range R_f of size $|R_f| \leq 2^m$. Let $\mathcal{G} = \{g : \{0,1\}^n \to \{0,1\}^n\}$ be a family of functions. Let $\mathcal{C}$ be a family of distributions over $\{0,1\}^n$ such that every $X \in \mathcal{C}$ has min-entropy $H_\infty(X) \geq k$. If*

$$\begin{cases} t \geq 2(2m + \log|\mathcal{F}| + \log|\mathcal{G}| + \log|\mathcal{C}| + \log(1/\delta) + 3), \\ m \leq (k - 3\log(1/\varepsilon) - \log t - 5)/2, \end{cases}$$

then with probability at least $(1-\delta)$ over $h \leftarrow \mathcal{H}$, it holds that for every $f \in \mathcal{F}$, $g \in \mathcal{G}$, and $X \in \mathcal{C}$,

$$\Delta((f(g(X)), f(h(X))), (f(g(X), f(U_n))) \leq \varepsilon.$$

Proof. (Sketch) The lemma follows in essentially identical way to the proof of Lemma 5.3, where we first apply Lemma A.2 to show that X has sufficient entropy conditioning on the leakage $f(g(X))$, and then applying Lemma B.2 together with union bounds over $\mathcal{F}$, $\mathcal{G}$, $\mathcal{C}$, and the leakage. □

Finally, the following rather standard lemma about pairwise independent hash function is useful for us.

Lemma B.5. *Let $n = 3m \in \mathbb{N}$ and $\mathcal{H} = \{h : \{0,1\}^m \to \{0,1\}^n\}$ be a pairwise independent hash function family. With probability at least $1 - 2^{-\Omega(n)}$ over $h \leftarrow \mathcal{H}$, h is injective.*

C Proof of Theorem 5.13

Assume for contradiction there exists a polynomial q and probabilistic polynomial time adversary A such that A breaks the s-bounded q-circular RDM security of Π. We define hybrids exactly as in Theorem 5.6: for $1 \leq j \leq q(k)+1$, $\mathrm{H}^{\Pi}_{j}(A,k)$ is as follows

$$
\begin{array}{rl}
\mathrm{H}^{\Pi}_{j}(A,k) := & (pk, sk) \leftarrow \mathrm{Gen}(1^k) \\
& f^1, f^2, \ldots, f^{q(k)}, \mathsf{state} \leftarrow A_1(1^k, pk) \\
& r^1, r^2 \ldots r^{q(k)} \leftarrow \{0,1\}^{q(k)|r|} \\
& \text{for } i = 1, 2 \ldots, q(k) \\
& \quad \text{if } i < j \text{ then } c^i \leftarrow \mathrm{Enc}_{pk}(f^i(r^{i-1}, c^{i-1}); r^i) \\
& \quad \text{else } c^i \leftarrow \mathrm{Enc}_{pk}(0; r^i) \\
& o \leftarrow A_2(\bar{c}, \mathsf{state}) \\
& \text{Output } o
\end{array}
$$

where $r^0 = r^{q(k)}$, c^0 is the empty string and $\bar{c} = c^1, c^2, \ldots c^{q(k)}$. As before, $\mathrm{H}^{\Pi}_{1}(A,k) = \mathrm{BCIR}^{\Pi}_{1}(A,k)$ and $\mathrm{H}^{\Pi}_{q(k)+1}(A,k) = \mathrm{BCIR}^{\Pi}_{0}(A,k)$. Note that by 5.11, Π is s-bounded RDM secure and hence CPA secure. The indistinguishability of $\mathrm{H}^{\Pi}_{j}(A,k)$ and $\mathrm{H}^{\Pi}_{j+1}(A,k)$ for $j = 1, 2, \ldots q-1$ follows in a straightforward manner from the CPA security of Π, just as it does in Theorem 5.6. Indeed, $\mathrm{H}^{\Pi}_{j}(A,k)$ and $\mathrm{H}^{\Pi}_{j+1}(A,k)$ for $j = 1, 2, \ldots q-1$ differ in only c^j which in $\mathrm{H}^{\Pi}_{j}(A,k)$ is an encryption of 0 and in $\mathrm{H}^{\Pi}_{j+1}(A,k)$ is an encryption of $f^j(r^{j-1}, c^{j-1})$. Furthermore both experiments can be simulated exactly without $f^{j+1}(r^j, c^j)$, where r^j is the randomness used for the above encryptions. Hence, if $\mathrm{H}^{\Pi}_{j}(A,k)$ and $\mathrm{H}^{\Pi}_{j+1}(A,k)$ are distinguishable then an efficient CPA adversary can be constructed that breaks the CPA security of Π.

It remains to show that $\mathrm{H}^{\Pi}_{q}(A,k)$ and $\mathrm{H}^{\Pi}_{q+1}(A,k)$ are indistinguishable. Unlike the previous cases, this does not follow in the same straightforward manner from the CPA security of Π because here the reduction needs $f^1(r^q)$ to simulate the experiment, where r^q is the randomness used to encrypt the challenge ciphertext in the CPA security game. We will use Lemma 5.12 to show that both $\mathrm{H}^{\Pi}_{q}(A,k)$ and $\mathrm{H}^{\Pi}_{q+1}(A,k)$ are indistinguishable from the same distribution: $\mathrm{H}^{*\Pi}(A,k)$ where

$$
\begin{array}{rl}
\mathrm{H}^{*\Pi}(A,k) := & (pk, sk) \leftarrow \mathrm{Gen}(1^k) \\
& f^1, f^2, \ldots, f^{q(k)}, \mathsf{state} \leftarrow A_1(1^k, pk) \\
& r^1, r^2 \ldots r^{q(k)} \leftarrow \{0,1\}^{q(k)|r|} \\
& \text{for } i = 1, 2 \ldots, q(k)-1 \\
& \quad c^i \leftarrow \mathrm{Enc}_{pk}(f^i(r^{i-1}, c^{i-1}); r^i) \\
& c^{q(k)} \leftarrow Enc_{pk}(U_n) \\
& o \leftarrow A_2(\bar{c}, \mathsf{state}) \\
& \text{Output } o
\end{array}
$$

where $r^0 = r^{q(k)}$, c^0 is the empty string and $\bar{c} = c^1, c^2, \ldots c^{q(k)}$. First we show $\mathrm{H}^{\Pi}_{q+1}(A,k)$ is indistinguishable from $\mathrm{H}^{*\Pi}(A,k)$. The view of A in both experiments can be described by $(pk, h, \bar{c}, \sigma)$, where pk is the key of the lossy trapdoor function and h is the t-wise independent hash function used by Π and σ is the internal randomness of A. Applying Lemma 5.12 with $X = (f^q(r^{q-1}, c^{q-1}), r^q)$, h as the same t-wise independent hash function used by Π, $f = F_{pk}$ and g as a randomized function such that $g(X) = h(f^1(r^q)||r^1||o^{2(l+R)})$ we get that with $1 - 2^{-u}$ probability over h, for any pk, any size s functions f^q and f^1 and any r^1, $(c^1, F_{pk}(h(f^q(r^{q-1}, c^{q-1}), r^q)))$ and $(c^1, F_{pk}(h(U_n)))$ are 2^{-u} close, where $c^1 = F_{pk}(h(f^1(r^q), r^1))$. Hence (pk, h, c^1, c^q) in both experiments are statistically close. The rest of A's view is derived from (pk, h, c^1, c^q) in both experiments in the same manner.

It follows that $\mathrm{H}^{\Pi}_{q+1}(A,k)$ and $\mathrm{H}^{*\Pi}(A,k)$ are statistically close. Similarly applying Lemma 5.12 with $X = (0, r^q)$ gives us that $\mathrm{H}^{\Pi}_{q}(A,k)$ and $\mathrm{H}^{*\Pi}(A,k)$ are statistically close.