# Note on the RKA security of Continuously Non-Malleable Key-Derivation Function from PKC 2015

Eiichiro Fujisaki and Keita Xagawa

NTT Secure Platform Laboratories
{fujisaki.eiichiro,xagawa.keita}@lab.ntt.co.jp

**Abstract.** Qin, Liu, Yuen, Deng, and Chen (PKC 2015) gave a new security notion of key-derivation function (KDF), continuous non-malleability with respect to $\Phi$-related-key attacks ($\Phi$-CNM), and its application to RKA-secure public-key cryptographic primitives. They constructed a KDF from cryptographic primitives and showed that the obtained KDF is $\Phi_{\mathsf{hoe\&iocr}}$-CNM, where $\Phi_{\mathsf{hoe\&iocr}}$ contains the identity function, the constant functions, and functions that have high output-entropy (HOE) and input-output collision-resistance (IOCR) simultaneously.

This short note disproves the security of their KDF by giving $\Phi_{\mathsf{hoe\&iocr}}$-RKAs by exploiting the components of their KDF. We note that their proof is still correct for $\Phi$-CNM for a subset of $\Phi_{\mathsf{hoe\&iocr}}$; for example the KDF satisfies $\Phi_{\mathsf{poly}(d)}$-CNM, in which an adversary can tamper with a secret by using polynomials of degree at most $d$.

## 1 Introduction

Security against related-key attacks (RKA-security) is one of the advanced security notions. It captures the security of cryptographic primitives and protocols in situations in which keys are correlated with each other or are tampered with by an adversary [Bih94, Knu93, BK03, BCM11]. Roughly speaking, we say that a primitive is RKA-secure if no efficient adversary cannot win the original security game even if the adversary is additionally allowed to access a system $\langle G, s \rangle$ by using input $x$ and a function $\phi \in \Phi$ and obtain $G(\phi(s), x)$ with certain restrictions. The function $\phi$ is called as a related-key derivation (RKD) function. Apparently, the broader $\Phi$ we allow, the stronger the security notion is. Unfortunately, we cannot achieve RKA security with respect to a set of any efficient RKD functions; for example, we should avoid bit-fixing functions $\phi_i$ that fix the $i$-th bit of input to 0 because they allow us to recover a secret by mounting a "testing-for-malfunctioning" attack [GLM$^+$04].

Therefore, the researchers broadened the class $\Phi$ of RKD functions while avoiding impossible RKD functions: addition or multiplication [BC10, AHI11, BCM11, Wee12], Affine functions [BPT12, JLLM13, JLLM14], and a subset of polynomials [GOR11, LMR14, ABPP14]. On public-key primitives, Bellare, Paterson, and Thomson [BPT12] proposed RKA-secure primitives with respect to $\Phi_{\mathsf{poly}(d)}[\mathbb{F}_q]$, a set of polynomials over a field $\mathbb{F}_q$ of degree at most $d$, based on the $d$-extended DBDH assumption. The important problem is how we can show $\Phi$-RKA-security for beyond "algebraic" RKD functions. For example, bit-flipping functions are not included in $\Phi_{\mathsf{poly}(d)}[\mathbb{F}_q]$ when $q$ is an odd prime.

Recently, several papers have addressed the problem and proposed cryptographic primitives that are $\Phi$-RKA secure beyond algebraic RKD functions: Jafargholi and Wichs [JW15],[1] Qin, Liu, Yuen, Deng, and Chen [QLY$^+$15], Fujisaki and Xagawa [FX15], and Abdalla, Benhamouda, and Passelègue [ABP15]. We here focus on Qin et al.'s result.

*Review of Qin et al.'s result.* Qin et al. [QLY$^+$15] considered $\Phi_{\mathsf{hoe\&iocr}}$-RKA security, which is wider than $\Phi_{\mathsf{poly}(d)}$ when we consider a field as a secret space. Roughly speaking, we say that $\phi$ satisfies high output-entropy (HOE) if $\phi(s)$ has sufficiently high min-entropy when $s$ is chosen uniformly at random and $\phi$ satisfies input-output collision resistance (IOCR) if $\phi$'s fixed points are sufficiently small. $\Phi_{\mathsf{hoe\&iocr}}$

[1] Employing non-persistent continuous non-malleable codes without self-destruction often provides RKA security, depending on the definition of RKA security. See [JW15, Section 1.4].

consists of the identity function $\mathfrak{id}$, the constant functions, and RKD functions satisfying both HOE and IOCR.

They define the *continuous non-malleability* (*CNM*) of the key-derivation function (KDF) by extending the non-malleability of KDFs [FMVW14]. Their definition of KDF is the public-key version, in which a sampling algorithm $\mathsf{Sample}$ outputs a secret $s \in \mathcal{S}$ and a public information $\pi \in \Pi$ and a key-derivation algorithm $\mathsf{Derive}$ on inputs $\pi$ and $s$ outputs a derived key $r$. They constructed their KDF scheme from one-time lossy filter [QL13], pair-wise independent hash functions, and one-time signature. They showed that their KDF satisfies $\Phi_{\mathsf{hoe\&iocr}}$-CNM; roughly speaking, any efficient adversary cannot distinguish a real derived key from a random derived key even if it can access the derivation oracle *many times* with queries $\phi \in \Phi_{\mathsf{hoe\&iocr}}$ and $\pi$, which returns the special symbol $\mathsf{same}^*$ if $(\phi(s^*), \pi) = (s^*, \pi^*)$ and $\mathsf{Derive}(\pi, \phi(s))$ otherwise (see formal definition in Section 2.2). Their CNM-KDFs can be instantiated from the standard assumptions, say, the DDH assumption, the DCR assumption, and the subgroup-decision assumptions, since a one-time lossy filter can be instantiated from the standard assumptions (DDH, DCR, and SD) [QL13, QL14].

They also propose a conversion from an ordinal public-key cryptographic primitive to a $(\Phi_{\mathsf{hoe\&iocr}}[\mathcal{S}] \times \Phi_{\mathsf{all}}[\Pi])$-RKA-secure one by using the $\Phi_{\mathsf{hoe\&iocr}}[\mathcal{S}]$-CNM KDF, where $\Phi_{\mathsf{all}}[\Pi]$ is a set of all efficient functions over $\Pi$ (see Section 3.4).

### 1.1 Our Contribution

We revisit the $\Phi_{\mathsf{hoe\&iocr}}$-CNM security of KDFs and find a flaw in the proof of Qin et al. [QLY+15] by giving a concrete $\Phi_{\mathsf{hoe\&iocr}}$-RKA. Our attack is very simple; let $\phi$ map $s$ to itself if $\mathsf{Derive}(\pi^*, s) = r^*$ and $s + 1$ otherwise. Given $\pi^*$ and $r^*$, which is a real derived-key $\mathsf{Derive}(\pi^*, s^*)$ or a random derived-key, we query $\phi$ and $\pi^*$ to the oracle; if $r^*$ is a real derived-key, we have $\phi(s^*) = s^*$ and receive the special symbol $\mathsf{same}^*$; otherwise, that is, if $r^*$ is a random derived-key, we have $\phi(s^*) \neq s^*$ with a sufficiently high probability and rarely receive $\mathsf{same}^*$. Therefore, we can distinguish the real derived-key from the random one. The remaining task is to show that $\phi$ is in $\Phi_{\mathsf{hoe\&iocr}}$. This is easy because $\pi^*$ includes the image of $s^*$ under an injective map, one-time lossy filter. The details are in Section 3. We additionally give a key-recovery attack by modifying this $\phi$ slightly.

We note that their proof is correct with respect to $\Phi \subseteq \Phi_{\mathsf{hoe\&iocr}}$ if any function $\phi \in \Phi$ does not change its action when we change the games in the proof. For example, their proof is correct with respect to $\Phi_{\mathsf{poly}(d)}$-RKA security, in which an adversary can tamper with a secret by polynomials of degree at most $d$. Defining appropriate subclass $\Phi \subset \Phi_{\mathsf{hoe\&iocr}}$, we still recover the $\Phi$-RKA security of their KDF scheme against RKD functions "beyond polynomials." Unfortunately, we do not know how we can define a wide class of RKD functions whose actions do not change when we change the games in the proof of Qin et al. We leave such a definition as an open problem. See Section 4 for discussion.

## 2 Preliminaries

We here briefly recall pairwise independent hash functions, RKD functions, one-time lossy filter, and key-derivation function.

Let $\kappa$ be the security parameter. We use the standard $O$-notations, $O$, $\Omega$, $\Theta$, $o$, and $\omega$. For a positive real $x$, $\lg(x) := \log_2(x)$ denotes the logarithm of $x$ with base 2.

*Pairwise independent hash functions.*

**Definition 2.1.** *A family of functions $\mathcal{H}_\kappa = \{h \mid h : \mathcal{S} \to \mathcal{R}\}$ is said to be* a family of pairwise independent hash functions *if, for all distinct pair $s_1 \neq s_2 \in \mathcal{S}$ and all $a_1, a_2 \in \mathcal{R}$,*

$$\Pr_{h \leftarrow \mathcal{H}_\kappa}[h(s_1) = a_1 \wedge h(s_2) = a_2] = (1/\#\mathcal{R})^2.$$

*RKD functions.* We let $\mathcal{S}$ denote the key space, which is defined as the public parameters of the specific scheme. We regard an RKD function $\phi \in \Phi$ simply as an *efficiently computable* function from $\mathcal{S}$ to $\mathcal{S}$, depending on the specification of the cryptographic protocol and the public parameters. $\Phi_{\mathsf{all}}[\mathcal{S}]$ denotes the set of all efficiently computable functions on $\mathcal{S}$. We often write $\Phi[\mathcal{S}]$ and $\Phi \subseteq \Phi_{\mathsf{all}}[\mathcal{S}]$ when we want to stress the secret space $\mathcal{S}$.

*Classes of RKD functions.* We denote an identity function on $\mathcal{S}$ by $\mathfrak{id}$. We write $\Phi_{\mathsf{const}}[\mathcal{S}]$ to denote all constant functions, $\phi_a : s \mapsto a$. If $\mathcal{S}$ is a finite ring, $\Phi_{\mathsf{poly}(d)}[\mathcal{S}]$ denotes all $\mathcal{S}$-coefficient polynomials of degree at most $d$, that is, $\{f : s \mapsto f(s) \mid f(x) \in \mathcal{S}[x], \deg(f) \leq d\}$.

We recall the RKD function class which is the target of [QLY+15].

**Definition 2.2** ( **[QLY+15, Definition 1]).** *Let $\mathcal{S}$ be a set with size $2^{\omega(\lg(\kappa))}$. The RKD function class $\Phi_{\mathsf{hoe\&iocr}}$ is called as* high output-entropy and input-output collision resistance (HOE&IOCR) *if it satisfies*

- *(HOE:)* $\Phi_{\mathsf{hoe\&iocr}} \setminus \Phi_{\mathsf{const}} \subseteq \left\{\phi \in \Phi_{\mathsf{all}}[\mathcal{S}] : \max_{y \in \mathcal{S}} \Pr_{x \leftarrow \mathcal{S}}[\phi(x) = y] = \mathsf{negl}(\kappa)\right\}$,
- *(IOCR:)* $\Phi_{\mathsf{hoe\&iocr}} \setminus \{\mathfrak{id}\} \subseteq \left\{\phi \in \Phi_{\mathsf{all}}[\mathcal{S}] : \Pr_{x \leftarrow \mathcal{S}}[\phi(x) = x] \leq \mathsf{negl}(\kappa)\right\}$.

HOE implies that all functions except constant functions in the class have high output-entropy; IOCR implies that all functions except the identical function in the class have few fixed points.

*Remark 2.1.* We say that an RKD function $\phi : \mathcal{S} \to \mathcal{S}$ is HOE if it satisfies $\max_{y \in \mathcal{S}} \Pr_{x \leftarrow \mathcal{S}}[\phi(x) = y] = \mathsf{negl}(\kappa)$. We also say that an RKD function $\phi : \mathcal{S} \to \mathcal{S}$ is IOCR if it satisfies $\Pr_{x \leftarrow \mathcal{S}}[\phi(x) = x] \leq \mathsf{negl}(\kappa)$.

## 2.1 One-Time Lossy Filter (OT-LF)

A one-time lossy filter is introduced by Qin and Liu [QL13] as a weakened primitive of a lossy algebraic filter [Hof13] and lossy trapdoor functions [PW08].

*Syntax:* A OT-LF scheme $\mathsf{LF}$ consists of three algorithms:

- The setup algorithm $\mathsf{Setup}$ that, on input $1^\kappa$, outputs an evaluation key $ek$ and a trapdoor $td$. The evaluation key defines a tag space $\mathcal{T} = \{0,1\}^* \times \mathcal{T}_c$ that contains two disjoint subsets, that of lossy tags $\mathcal{T}_{\mathrm{loss}}$ and that of injective tags $\mathcal{T}_{\mathrm{inj}}$.
- The evaluation algorithm $\mathsf{Eval}$ that, on input $ek$, a tag $t \in \mathcal{T}$, and a preimage $s \in \mathcal{S}$, outputs a image $y \in \mathcal{Z}$. We denote $y = \mathsf{LF}_{ek,t}(s)$.
- The lossy-tag generation algorithm $\mathsf{LTGen}$ that, on input $td$ and $t_a \in \{0,1\}^*$, outputs $t_c \in \mathcal{T}_c$ such that $t = (t_a, t_c) \in \mathcal{T}_{\mathrm{loss}}$.

*Security:*

**Definition 2.3** ( **[QLY+15, Section 4]).** *We say a OT-LF scheme* $\mathsf{LF} = (\mathsf{Setup}, \mathsf{Eval}, \mathsf{LTGen})$ *is* $(\mathcal{S}, \ell)$*-one-time lossy if it satisfies the following three properties:*

- *Lossiness: If $t$ is injective, that is, $t \in \mathcal{T}_{\mathrm{inj}}$, $\mathsf{LF}_{ek,t}(\cdot)$ is injective. If $t$ is lossy, that is, $t \in \mathcal{T}_{\mathrm{loss}}$, $\#\mathsf{LF}_{ek,t}(\mathcal{S}) \leq 2^\ell$, that is, the number of possible values is at most $2^\ell$.*
- *Indistinguishability of tags: Any PPT adversary cannot distinguish a lossy tag from a random tag. Formally, we require that, for any PPT adversary, its advantage defined by*

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathsf{A},\mathsf{LF}}(\kappa) = \left| \begin{array}{c} \Pr\left[(ek, td) \leftarrow \mathsf{Setup}(1^\kappa); (t_a, st) \leftarrow \mathsf{A}(ek); t_c \leftarrow \mathsf{LTGen}(td, t_a) : \mathsf{A}(t_c, st) = 1\right] \\ - \Pr\left[(ek, td) \leftarrow \mathsf{Setup}(1^\kappa); (t_a, st) \leftarrow \mathsf{A}(ek); t_c \leftarrow \mathcal{T}_c : \mathsf{A}(t_c, st) = 1\right] \end{array} \right|$$

*is negligible in $\kappa$.*

- *Evasiveness: Any PPT adversary cannot produce a non-injective tag even if it sees a lossy tag. Formally, we require that, for any PPT adversary, its advantage defined by*

$$\mathsf{Adv}^{\mathsf{eva}}_{\mathsf{A},\mathsf{LF}}(\kappa) = \Pr\left[\begin{array}{l}(ek, td) \leftarrow \mathsf{Setup}(1^\kappa); (t'_a, st) \leftarrow \mathsf{A}(ek); t'_c \leftarrow \mathsf{LTGen}(td, t'_a); (t_a, t_c) \leftarrow \mathsf{A}(t'_c, st) : \\ (t'_a, t'_c) \neq (t_a, t_c) \wedge (t'_a, t'_c) \in \mathcal{T} \setminus \mathcal{T}_{\mathrm{inj}}\end{array}\right]$$

  *is negligible in $\kappa$.*

We note a simple lemma that if we choose $t_c$ uniformly at random, then the tag $(t_a, t_c)$ will be injective with high probability.

**Lemma 2.1.** *Suppose that* $\mathsf{LF} = (\mathsf{Setup}, \mathsf{Eval}, \mathsf{LTGen})$ *is $(\mathcal{S}, \ell)$-one-time lossy. Then, for any $t_a$,*

$$\Pr[(ek, td) \leftarrow \mathsf{Setup}(1^\kappa); t_c \leftarrow \mathcal{T}_c : (t_a, t_c) \in \mathcal{T}_{\mathrm{inj}}] \geq 1 - 1/\kappa.$$

*Proof.* If not, there exists $t_a^*$ such that randomly chosen $t_c$ results in $(t_a, t_c) \notin \mathcal{T}_{\mathrm{inj}}$ with probability at least $1/\kappa$. This breaks evasiveness of $\mathsf{LF}$. □

## 2.2 Key-Derivation Function (KDF)

The non-malleability of KDF is introduced by Faust et al. [FMVW14]. Qin et al. slightly modified the syntax of the KDF and gave the security notion of the *continuous non-malleability* of KDF [QLY+15]. We here adopt the syntax and the security notion of Qin et al. because we will analyze their scheme.

*Syntax:* A key derivation function (KDF) scheme consists of three algorithms: the setup algorithm $\mathsf{Setup}$ that, on input $1^\kappa$, outputs public parameter $\textsc{pp}$, the sampling algorithm $\mathsf{Sample}$ that, on input $\textsc{pp}$, outputs a secret string $s \in \mathcal{S}$ and a public information $\pi \in \Pi$, and the derivation algorithm $\mathsf{Derive}$ that, on inputs $\textsc{pp}$, $\pi$, and $s$, outputs a derived key $r \in \mathcal{R}$ or the special symbol $\perp$.

*Remark 2.2.* In the definition of Faust et al. [FMVW14], the sampling algorithm outputs $s$ only and the derivation algorithm outputs $r$ on inputs $\textsc{pp}$ and $s$.

*Security:*

**Definition 2.4 (Continuous non-malleability [QLY+15, Section 4]).** *For KDF scheme* $\mathsf{KDF} = (\mathsf{Setup}, \mathsf{Sample}, \mathsf{Derive})$*, a class of RKD functions $\Phi$, and a bit $b \in \{0, 1\}$, we define experiment* $\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, b)$ *between adversary* $\mathsf{A}$ *and a challenger as follows.*

```
$\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, b)$:
  $\textsc{pp} \leftarrow \mathsf{Setup}(1^\kappa)$, $(s^*, \pi^*) \leftarrow \mathsf{Sample}(\textsc{pp})$
  $r_0 \leftarrow \mathsf{Derive}(\textsc{pp}, \pi, s)$, $r_1 \leftarrow \mathcal{R}$,
  $r^* \leftarrow r_b$
  $b' \leftarrow \mathsf{A}^{\textsc{RK-Derive}(\cdot,\cdot)}(\textsc{pp}, \pi^*, r^*)$
  Return $b'$
```

```
$\textsc{RK-Derive}(\phi, \pi)$:
  If $\phi \notin \Phi$, then return $\perp$
  If $(\phi(s^*), \pi) = (s^*, \pi^*)$,
    then return a special symbol $\mathsf{same}^*$
  Else, return $r \leftarrow \mathsf{Derive}(\textsc{pp}, \phi(s^*), \pi)$
```

*We define the advantage of* $\mathsf{A}$ *as*

$$\mathsf{Adv}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa) = \left|\Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, 0) = 1\right] - \Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, 1) = 1\right]\right|.$$

*We say that* KDF *is* continuously non-malleable with respect to $\Phi$ *($\Phi$-CNM in short) if, for any PPT adversary* $\mathsf{A}$*, its advantage is negligible in $\kappa$.*

# 3 Related-Key Attacks against CNM-KDFs

We first show an impossibility result for $\Phi_{\mathsf{all}}[\mathcal{S}]$-CNM KDF as a warm up, which is very similar to the attacks to show impossibility results in [GLM+04, DFMV13]. We then exemplify a concrete $\Phi_{\mathsf{hoe\&iocr}}$-RKA against Qin et al.'s CNM-KDF by using the RKD functions used to show the general impossibility.

### 3.1 Impossibility of $\Phi_{\mathsf{all}}[\mathcal{S}]$-CNM KDF.

**Theorem 3.1.** *There is no* $\Phi_{\mathsf{all}}[\mathcal{S}]$*-CNM KDF.*

*Proof.* Suppose that $\mathcal{S}$ is an additive group with binary operation + (we can easily remove this assumption). Let us fix a non-zero element in $\mathcal{S}$ and denote it by $e$. We define $\phi^* : \mathcal{S} \to \mathcal{S}$ by

$$\phi^*(s) = \begin{cases} s & \text{if } \mathsf{Derive}(\textsc{pp}, \pi^*, s) = r^*, \\ s + e & \text{otherwise.} \end{cases}$$

We show that $\phi^*$ violates the $\Phi_{\mathsf{all}}[\mathcal{S}]$-CNM security of KDFs.

Consider the adversary A:

1. A receives $\textsc{pp}, \pi^*$, and $r^*$ from its challenger, where $(s^*, \pi^*) \leftarrow \mathsf{Sample}(\textsc{pp}), r_0 \leftarrow \mathsf{Derive}(\textsc{pp}, \pi^*, s^*)$, and $r_1 \leftarrow \mathcal{R}$, and $r^* \leftarrow r_b$.
2. A queries $(\phi^*, \pi^*)$ and receives $\xi$.
3. A returns $b' = 0$ if $\xi = \mathsf{same}^*$; otherwise, A returns $b' = 1$.

We note that $\textsc{pp}$, $\pi^*$, and $r^*$ are hardwired to $\phi^*$, and thus, $\phi^*$ is efficiently computable. Notice that the oracle RK-Derive on query $(\phi^*, \pi^*)$ returns $\mathsf{same}^*$ if $\mathsf{Derive}(\textsc{pp}, \pi, s^*) = r^*$ and returns $r' \leftarrow \mathsf{Derive}(\textsc{pp}, \pi^*, s^* + e)$ if $\mathsf{Derive}(\textsc{pp}, \pi, s^*) \neq r^*$.

If $r^* = r_0$, the adversary always outputs 0 because $\mathsf{Derive}(\textsc{pp}, \pi^*, s^*) = r_0$ holds and the oracle returns $\xi = \mathsf{same}^*$.

Otherwise, that is, if $r^* = r_1$, the adversary outputs 1 unless $\mathsf{Derive}(\textsc{pp}, \pi^*, s^*) = r_1$. Let Bad be the event that $\mathsf{Derive}(\textsc{pp}, \pi, s) = r_1$ holds, where $r_1 \leftarrow \mathcal{R}$; $\textsc{pp} \leftarrow \mathsf{Setup}_{\mathsf{KDF}}(1^\kappa)$; $(\pi, s) \leftarrow \mathsf{Sample}(\textsc{pp})$. Since $r_1$ is chosen uniformly at random, we have $\Pr[\mathsf{Bad}] \leq 1/\#\mathcal{R} \leq 1/2$.

Summarizing the above, the adversary's advantage is

$$\begin{aligned} \mathsf{Adv}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi_{\mathsf{all}}[\mathcal{S}]}(\kappa) &= \left|\Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, 0) = 1\right] - \Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, 1) = 1\right]\right| \\ &= |0 - (1 - \Pr[\mathsf{Bad}])| \geq 1 - 1/2 \geq 1/2, \end{aligned}$$

which is apparently constant. Therefore, any KDF cannot satisfy $\Phi_{\mathsf{all}}[\mathcal{S}]$-CNM. □

*Remark 3.1.* We note that the RKD function $\phi^*$ satisfies the HOE condition because any $s \in \mathcal{S}$ has at most two preimages $s$ and $s - e$. In addition, if the number of $s$ that satisfies $\mathsf{Derive}(\textsc{pp}, \pi^*, s) = r^*$ is sufficiently small, $\phi^*$ also satisfies the IOCR condition. In general, we cannot say the function is IOCR. For example, consider the case that $\#\mathcal{R} = 2$.

### 3.2 Related-Key Attacks against Qin et al.'s CNM-KDF.

Let us review how Qin et al. constructed their KDF scheme from one-time lossy filter (Definition 2.1), one-time signature, and pairwise-independent hash functions.

**Definition 3.1** ($\mathsf{KDF}_{\mathsf{QLY+15}}$ **[QLY⁺15]).** *Let* $\mathsf{LF} = (\mathsf{Gen}_{\mathsf{LF}}, \mathsf{Eval}, \mathsf{LTag})$ *be a one-time lossy filter scheme whose domain is* $\mathcal{S}$*, range is* $\mathcal{Z}$*, residual leakage is* $\ell$*, and tag space is* $\mathcal{T} = \{0,1\}^* \times \mathcal{T}_c$*. Let* $\mathcal{H} = \{\mathcal{H}_\kappa\}$ *be a family of pairwise independent hash functions whose domain is* $\mathcal{S}$ *and range is* $\mathcal{R} = \{0,1\}^m$*. Let* $\mathsf{OTS} = (\mathsf{Setup}_{\mathsf{OTS}}, \mathsf{Gen}_{\mathsf{OTS}}, \mathsf{Sign}_{\mathsf{OTS}}, \mathsf{Vrfy}_{\mathsf{OTS}})$ *be a one-time signature scheme whose signature space is denoted by* $\Sigma$*.*

- $\mathsf{Setup}_{\mathsf{KDF}}(1^\kappa)$*:* $(ek, td) \leftarrow \mathsf{Gen}_{\mathsf{LF}}(1^\kappa)$*;* $\rho \leftarrow \mathsf{Setup}_{\mathsf{OTS}}(1^\kappa)$*;* $h \leftarrow \mathcal{H}_\kappa$*; Output* $\textsc{pp} = (ek, \rho, h)$*.*

- $\mathsf{Sample}(\textsc{pp})$: $(ovk, osk) \leftarrow \mathsf{Gen}_{\mathsf{OTS}}(\rho)$; $s \leftarrow \mathcal{S}$; $t_c \leftarrow \mathcal{T}_c$; *compute*
$$y \leftarrow \mathsf{Eval}_{ek,(ovk,t_c)}(s) \text{ and } \sigma \leftarrow \mathsf{Sign}_{\mathsf{OTS}}(osk, (t_c, y)).$$
*Output* $s$ *and* $\pi = (t = (ovk, t_c), y, \sigma)$.
- $\mathsf{Derive}(\textsc{pp}, \pi, s)$: *If* $\mathsf{Eval}_{ek,(ovk,t_c)}(s) \neq y$ *or* $\mathsf{Vrfy}_{\mathsf{OTS}}(ovk, (t_c, y), \sigma) \neq 1$, *then return* $\perp$; *otherwise, output* $r \leftarrow h(s)$.

Qin et al. showed that the obtained KDF scheme is secure against $\Phi_{\mathsf{poly}(d)}[\mathcal{S}]$-RKAs:

**Theorem 3.2 ( [QLY+15, Theorem 1]).** *Let* $d = \mathrm{poly}(\kappa)$ *be an integer.* $\mathsf{KDF}_{\mathsf{QLY+15}}$ *is* $\Phi_{\mathsf{poly}(d)}[\mathcal{S}]$*-CNM KDF if* $\mathsf{LF}$ *is* $(\mathcal{S}, \ell)$*-lossy,* $\mathsf{OTS}$ *is strongly EUF-CMA secure, and* $\lg \#\mathcal{S} \geq \ell + m + \omega(\lg \kappa)$.

Following the recommendation of reviewers of PKC 2015, Qin et al. insisted that the following theorem is correct.

**Theorem 3.3 (Incorret generalization of [QLY+15, Theorem 1]).** $\mathsf{KDF}_{\mathsf{QLY+15}}$ *is* $\Phi_{\mathsf{hoe\&iocr}}[\mathcal{S}]$*-CNM KDF if* $\mathsf{LF}$ *is* $(\mathcal{S}, \ell)$*-lossy,* $\mathsf{OTS}$ *is strongly EUF-CMA secure, and* $\lg \#\mathcal{S} \geq \ell + m + \omega(\lg \kappa)$.

In the following, we disprove this generalization.

**Distinguishing attack.** We propose a related-key attack against the CNM-KDF, $\mathsf{KDF}_{\mathsf{QLY+15}}$. Our attack exploits the fact that the function $\mathsf{Eval}_{ek,(ovk^*,t_c^*)}$ is injective if $(ovk^*, t_c^*)$ is injective.

**Theorem 3.4.** $\mathsf{KDF}_{\mathsf{QLY+15}}$ *is not* $\Phi_{\mathsf{hoe\&iocr}}$*-CNM KDF.*

*Proof.* Let us apply our candidate $\phi^*$ to the CNM-KDF. We define

$$\phi^*(s) = \begin{cases} s & \text{if } \mathsf{Eval}_{ek,(ovk^*,t_c^*)}(s) = y^*, \mathsf{Vrfy}_{\mathsf{OTS}}(ovk, (t_c^*, y^*), \sigma^*) = 1, \text{ and } r^* = h(s), \\ s + e & \text{otherwise,} \end{cases}$$

which is efficiently computable function from $\mathcal{S}$ to $\mathcal{S}$. We verify that $\phi^*$ is HOE and IOCR with overwhelming probability by the following claim.

*Claim.* If the tag $(ovk^*, t_c^*)$ is injective, $\phi^* \in \Phi_{\mathsf{hoe\&iocr}}$.

*Proof (Proof of claim).* $\mathsf{Eval}_{ek,(ovk^*,t_c^*)}$ is injective since the tag $(ovk^*, t_c^*)$ is injective by the hypothesis. $s^*$ is a unique preimage of $y^*$ under the function $\mathsf{Eval}_{ek,(ovk^*,t_c^*)}$ and other elements cannot satisfy $\mathsf{Eval}_{ek,(ovk^*,t_c^*)}(s) = y^*$. Hence, the number of fixed points of $\phi^*$ is at most 1. This implies the IOCR property of $\phi^*$.
In addition, the function $\phi^*$ maps any $s \in \mathcal{S} \setminus \{s^*\}$ to $s + e$. Therefore, the output of $\phi^*$ is almost uniformly at random if $s$ is chosen randomly. This shows the HOE property of $\phi^*$. □

Now, let us consider an adversary $\mathsf{A}$ that distinguishes a real game with a random game as follows.

1. $\mathsf{A}$ receives $\textsc{pp}$, $\pi^*$, and $r^*$ from its challenger, where $(s^*, \pi^*) \leftarrow \mathsf{Sample}(\textsc{pp})$, $r_0 \leftarrow \mathsf{Derive}(\textsc{pp}, \pi^*, s^*)$, and $r_1 \leftarrow \mathcal{R}$, and $r^* \leftarrow r_b$.
2. $\mathsf{A}$ queries $(\phi^*, \pi^*)$ and receives $\xi$.
3. If $\mathsf{A}$ receives $\xi = \perp$, then it outputs $\perp$.
4. $\mathsf{A}$ returns $b' = 0$ if $\xi = \mathsf{same}^*$; otherwise, $\mathsf{A}$ returns $b' = 1$.

Suppose that $(ovk^*, t_c^*)$ is injective, which happens with probability at least $1 - 1/\kappa$ by Lemma 2.1. Let us estimate the probabilities that $\phi^*(s^*) = s^*$ in both games. In a real game, we always have $\phi^*(s^*) = s^*$. In a random game where $r^* \leftarrow \mathcal{R}$, $h(s^*) = r^*$ happens with probability $1/\#\mathcal{R}$, because $h$ is chosen from a family of pairwise independent hash functions. Therefore, we have $\phi^*(s^*) = s^*$ with a probability of at most $1/\#\mathcal{R} + \mathsf{negl}(\kappa)$.

We have

$$\begin{aligned}\mathsf{Adv}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi_{\mathsf{hoe\&iocr}}}(\kappa) &= \left|\begin{matrix}\Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi_{\mathsf{hoe\&iocr}}}(\kappa, 0) = 1\right] \\ -\Pr\left[\mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi_{\mathsf{hoe\&iocr}}}(\kappa, 1) = 1\right]\end{matrix}\right| \\ &= (1 - 1/\kappa)|0 - (1 - 1/\#\mathcal{R} - \mathsf{negl}(\kappa))| \\ &\geq (1 - 1/\kappa)(1 - 1/\#\mathcal{R} - \mathsf{negl}(\kappa)) \geq 1/4.\end{aligned}$$

which is apparently constant. □

**Key-recovery attack.** The previous attack exemplifies that we can distinguish a real derived key $r_0$ from a random derived key $r_1$ by utilizing $\phi^* \in \Phi_{\mathsf{hoe\&iocr}}$. We additionally give a stronger attack; by using variants of $\phi^*$, we can retrieve $s^*$ by using the oracle RK-Derive when we are given $(\textsc{pp}, \pi^*, r^*)$, where $r^* \leftarrow \mathsf{Derive}(\textsc{pp}, \pi^*, s^*)$,

Suppose that $\mathcal{S} \subseteq \{0, 1\}^n$. We denote the $i$-th bits of $s$ and $s^*$ for $i = 1, \dots, n$ by $s_i$ and $s_i^*$, respectively. Define, for $i = 1, \dots, n$ and $b \in \{0, 1\}$,

$$\phi^*_{i,b}(s) = \begin{cases} s & \text{if } \mathsf{Eval}_{ek,(ovk^*,t_c^*)}(s) = y^*, \mathsf{Vrfy}_{\mathsf{OTS}}(ovk, (t_c^*, y^*), \sigma^*) = 1, \\ & r^* = h(s), \text{ and } s_i = b, \\ s + e & \text{otherwise.} \end{cases}$$

*Claim.* $\phi^*_{i,b} \in \Phi_{\mathsf{hoe\&iocr}}$ if $(ovk^*, t_c^*)$ is injective.

*Proof.* The proof is very similar to the claim on $\phi^*$, so we omit it.

*Claim.* Suppose that $(ovk^*, t_c^*)$ is injective. We have that $s_i^* = b$ if and only if $\phi_{i,b}(s^*) = s^*$.

*Proof.* The proof is again very similar to the claim on $\phi^*$, so we omit it.

Let us consider an adversary A defined as follows.

1. A receives $\textsc{pp}$, $\pi^*$, and $r^*$ from its challenger, where $(s^*, \pi^*) \leftarrow \mathsf{Sample}(\textsc{pp})$ and $r^* \leftarrow \mathsf{Derive}(\textsc{pp}, \pi^*, s^*)$.
2. For $i = 1, \dots, n$,
   (a) A queries $(\phi^*_{i,0}, \pi^*)$ and $(\phi^*_{i,1}, \pi^*)$ to the oracle RK-Derive and receives $\xi_0$ and $\xi_1$, respectively.
   (b) A sets $s_i = b$ if $\xi_b = \mathsf{same}^*$.
3. A outputs $s = s_1 \dots s_N$.

A obtains $s^*$ if $(ovk^*, t_c^*)$ is injective, which happens with probability at least $1 - 1/\kappa$.

### 3.3 Identifying a Pitfall

Let us review the proofs in [QLY+15] and discuss where a pitfall exists. If two games, Game and Game′, are computationally indistinguishable, we write Game $\approx_c$ Game′.

| Games | Tag | Key Derivation Rules |
|---|---|---|
| $\mathsf{Game}_0(b)$: | Injective | **R0**: If $(\phi(s^*), \pi) = (s^*, \pi^*)$, return $\mathsf{same}^*$;<br>else if $\mathsf{Eval}_{ek,(ovk,t_c)}(\phi(s^*)) = y$ and $\mathsf{Vrfy}_{\mathsf{OTS}}(ovk, (t_c, y), \sigma) = 1$,<br>return $\mathsf{Derive}(\textsc{pp}, \phi(s^*), \pi)$;<br>else return $\perp$ |
| $\mathsf{Game}_1(b)$: | Injective | **R1**: If $(\phi, \pi) = (\mathfrak{id}, \pi^*)$, return $\mathsf{same}^*$;<br>else if $\phi = \phi_c \in \Phi_{\mathsf{const}}$, $\pi = \pi^*$, and $y = \mathsf{Eval}_{ek,(ovk,t_c)}(c)$,<br>return $\mathsf{same}^*$;<br>else if $\phi = \phi_c \in \Phi_{\mathsf{const}}$ and $\pi \neq \pi^*$, return $\mathsf{Derive}(\textsc{pp}, \pi, c)$;<br>else if $\phi = \phi_c \in \Phi_{\mathsf{const}}$ and $y \neq \mathsf{Eval}_{ek,(ovk,t_c)}(c)$,<br>return $\mathsf{Derive}(\textsc{pp}, \pi, c)$;<br>**R0** |
| $\mathsf{Game}_2(b)$: | Injective | **R1**<br>**R2**: If $ovk = ovk^*$, but $((t_c, y), \sigma) \neq ((t_c^*, y^*), \sigma^*)$, return $\perp$<br>**R0** |
| $\mathsf{Game}_3(b)$: | Injective | **R1**<br>**R2**<br>**R3**: If $\pi = \pi^*$, but $\phi(s^*) \neq s^*$, return $\perp$<br>**R0** |
| $\mathsf{Game}_4(b)$: | Lossy | **R1**<br>**R2**<br>**R3**<br>**R0** |
| $\mathsf{Game}_5(b)$: | Lossy | **R1**<br>**R2**<br>**R3'**: Return $\perp$<br>**R0'**: Return $\perp$ |

**Table 1.** The games defined in [QLY+15, Section 4]. If $b = 0$, then $r^* = r_0$. Otherwise, $r^*$ is chosen uniformly at random.

Table 1 summarizes the games in their proofs, where $\mathsf{Game}_0(b) = \mathsf{Expt}^{\mathsf{cnm}}_{\mathsf{A},\mathsf{KDF},\Phi}(\kappa, b)$. To show $\mathsf{Game}_0(0) \approx_c \mathsf{Game}_0(1)$, they show $\mathsf{Game}_i(b) \approx_c \mathsf{Game}_{i+1}(b)$ for $i = 0, \ldots, 4$ and $\mathsf{Game}_5(0) \approx_c \mathsf{Game}_5(1)$.

We can verify that the proofs that show $\mathsf{Game}_0(b) \approx_c \mathsf{Game}_1(b) \approx_c \mathsf{Game}_2(b) \approx_c \mathsf{Game}_3(b)$ are correct. We also verify that the proofs that show $\mathsf{Game}_4(b) \approx_c \mathsf{Game}_5(b)$ and $\mathsf{Game}_5(0) \approx_c \mathsf{Game}_5(1)$ are correct.

The pitfall exists in the lemma showing $\mathsf{Game}_3(b) \approx_c \mathsf{Game}_4(b)$. We found that $\phi^*$ becomes an invalid RKD function when we go from $\mathsf{Game}_3(b)$ to $\mathsf{Game}_4(b)$. In $\mathsf{Game}_i(b)$ for $i = 0, 1, 2, 3$, the tag in $\pi^*$ is injective (with overwhelming probability). Therefore, $\phi^* \in \Phi_{\mathsf{hoe\&iocr}}$ as in $\mathsf{Game}_0(b)$. However, the tag in $\pi^*$ becomes lossy in $\mathsf{Game}_4(b)$. This makes $\phi^*$ non-IOCR; a number of $s$ that satisfies $\mathsf{Eval}_{ek,(ovk^*,t_c^*)}(s) = y^*$ and $h(s) = r^*$ exceeds the threshold of IOCR.

*Remark 3.2.* Originally, Qin et al. considered $\Phi_{\mathsf{poly}(d)}$-RKA security and generalized their proof to $\Phi_{\mathsf{hoe\&iocr}}$-RKA security by the recommendation of a reviewer [QLY+15, Acknowledgment]. We note that Qin et al.'s proof is correct for $\Phi_{\mathsf{poly}(d)}$-RKA security because any $\phi \in \Phi_{\mathsf{poly}(d)}$ is valid in any game. Moreover, their proof is correct for $\Phi_{\mathsf{fixed}} \subseteq \Phi_{\mathsf{hoe\&iocr}}$ if the function class is fixed a priori and recognizable and does not change in all games. See Section 4 for discussion.

### 3.4 Attack against the Embedded CNM-KDF

Qin et al. converted a public-key cryptographic system into a $\Phi$-RKA secure public-key cryptographic system with a $\Phi$-CNM-KDF. Consider an identity-based encryption (IBE) scheme as an example. The conversion is summarized as follows.

**Definition 3.2** ($\mathsf{IBE}_{\mathsf{QLY+15}}$)**.** *Let* $\mathsf{IBE} = (\mathsf{Setup}, \mathsf{Gen}, \mathsf{Extract}, \mathsf{Enc}, \mathsf{Dec})$ *be an IBE scheme. Let* $\mathcal{R}_{\mathsf{Gen}}$ *be a randomness space of* $\mathsf{Gen}$*. Let* $\mathsf{KDF} = (\mathsf{Setup}_{\mathsf{KDF}}, \mathsf{Sample}, \mathsf{Derive})$ *be a KDF scheme whose derived key is in* $\mathcal{R}_{\mathsf{Gen}}$*. A new IBE scheme is defined as follows.*

$\overline{\mathsf{Setup}}(1^\kappa)$**:** $\textsc{pp}_{\mathsf{KDF}} \leftarrow \mathsf{Setup}_{\mathsf{KDF}}(1^\kappa)$*;* $\textsc{pp}_{\mathsf{IBE}} \leftarrow \mathsf{Setup}(1^\kappa)$*; Output* $pp = (\textsc{pp}_{\mathsf{KDF}}, \textsc{pp}_{\mathsf{IBE}})$.

$\overline{\mathsf{Gen}}(pp)$**:** $(s, \pi) \leftarrow \mathsf{Sample}(\textsc{pp}_{\mathsf{KDF}})$*;* $r \leftarrow \mathsf{Derive}(\textsc{pp}_{\mathsf{KDF}}, \pi, s)$*;* $(\textsc{mpk}, \textsc{msk}) \leftarrow \mathsf{Gen}(\textsc{pp}_{\mathsf{IBE}}; r)$*; Output* $mpk = (\textsc{mpk}, \pi)$ *and* $msk = (s, \pi)$.

$\overline{\mathsf{Extract}}(pp, msk, id)$**:** $r \leftarrow \mathsf{Derive}(\textsc{pp}_{\mathsf{KDF}}, \pi, s)$*; If* $r = \perp$*, then output* $\perp$*; otherwise,* $(\textsc{mpk}, \textsc{msk}) \leftarrow \mathsf{Gen}(\textsc{pp}_{\mathsf{IBE}}; r)$*;* $dk_{id} \leftarrow \mathsf{Extract}(\textsc{pp}_{\mathsf{IBE}}, \textsc{msk}, id)$*. Output* $dk_{id}$.

$\overline{\mathsf{Enc}}(pp, mpk, id, m)$**:** *Output* $ct \leftarrow \mathsf{Enc}(\textsc{pp}_{\mathsf{IBE}}, \textsc{mpk}, id, m)$.

$\overline{\mathsf{Dec}}(pp, mpk, dk_{id}, ct)$**:** *Output* $m/\perp \leftarrow \mathsf{Dec}(\textsc{pp}_{\mathsf{IBE}}, \textsc{mpk}, dk_{id}, ct)$.

Qin et al. show the following theorem [QLY+15, Theorem 2].

**Theorem 3.5.** *The scheme* $\mathsf{IBE}_{\mathsf{QLY+15}}$ *is* $\Phi_{\mathsf{hoe\&iocr}}$*-IND-fID-CP-RKA secure if the basic IBE scheme is IND-fID-CPA secure and* $\mathsf{KDF}$ *is* $\Phi_{\mathsf{hoe\&iocr}}$*-CNM.*

Let $\mathcal{S}$ and $\Pi$ be a set of the secret $s$ and public information $\pi$ of KDF, respectively. Let $\mathcal{S}' = \mathcal{S} \times \Pi$ be a space of the master secret key of $\mathsf{IBE}_{\mathsf{QLY+15}}$. Strictly speaking, their theorem is restated as follows.

**Theorem 3.6.** *Suppose that the basic IBE scheme is IND-fID-CPA secure and* $\mathsf{KDF}$ *is* $\Phi_{\mathsf{hoe\&iocr}}[\mathcal{S}]$*-CNM. The scheme* $\mathsf{IBE}_{\mathsf{QLY+15}}$ *is IND-fID-CP-RKA secure with respect to* $\Phi_{\mathsf{hoe\&iocr}}[\mathcal{S}] \times \Phi_{\mathsf{const}}[\Pi]$.

**Key-recovery attack.** Let us consider a key-recovery related-key attack against this IBE scheme when we adopt $\mathsf{KDF}_{\mathsf{QLY+15}}$, which is not $\Phi_{\mathsf{hoe\&iocr}}$-CNM-KDF as we have shown. In contrast to the direct key-recovery attack against CNM-KDF, we cannot access $r^*$ directly. However, we can still employ a variant of $\phi^*_{i,b}$ defined as follows. For $i = 1, \ldots, n$ and $b \in \{0, 1\}$, we define

$$\phi^+_{i,b}(s) = \begin{cases} s & \text{if } \mathsf{Eval}_{ek,(ovk^*,t^*_c)}(s) = y^*, \mathsf{Vrfy}_{\mathsf{OTS}}(ovk, (t^*_c, y^*), \sigma^*) = 1, \\ & (\textsc{mpk}, \cdot) = \mathsf{Gen}_{\mathsf{IBE}}(\textsc{pp}_{\mathsf{IBE}}; h(s)), \text{ and } s_i = b, \\ s + e & \text{otherwise.} \end{cases}$$

In a similar argument, $\phi^+_{i,b}$ is HOE and IOCR if $(ovk^*, t^*_c)$ is injective. We have that $s^*_i = b$ if and only if $\phi^+_{i,b}(s^*) = s^*$.

Consider an adversary $\mathsf{A}$ that retrieves $s$ given $\textsc{pp}$, $\pi^*$, and $\textsc{mpk}^*$ by using the oracle RK-Extract:

1. $\mathsf{A}$ receives $\textsc{pp}_{\mathsf{KDF}}$, $\textsc{pp}_{\mathsf{IBE}}$, $\pi^*$, and $\textsc{mpk}^*$ from its challenger, where $(s^*, \pi^*) \leftarrow \mathsf{Sample}(\textsc{pp}_{\mathsf{KDF}})$, $r^* \leftarrow \mathsf{Derive}(\textsc{pp}_{\mathsf{KDF}}, \pi^*, s^*)$, and $(\textsc{mpk}^*, \textsc{msk}^*) \leftarrow \mathsf{Gen}(\textsc{pp}_{\mathsf{IBE}}; r^*)$.
2. $\mathsf{A}$ randomly chooses an identity $id^*$.
3. For $i = 1, \ldots, n$,
   (a) $\mathsf{A}$ queries $((\phi^+_{i,0}, \mathfrak{id}), id^*)$ and $((\phi^+_{i,1}, \mathfrak{id}), id^*)$ and receives $\xi_0$ and $\xi_1$, respectively.
   (b) If $\xi_0 = \perp$, then $\mathsf{A}$ sets $s_i = 1$. It sets $s_i = 0$, otherwise.
4. $\mathsf{A}$ outputs $s = s_1 \ldots s_N$.

With overwhelming probability, $\mathsf{A}$ obtains $s^*$ if the tag is injective. If $s^*_i = b$, $\xi_b$ is correct. Meanwhile, $\xi_{1-b}$ should be $\perp$ because the derivation of $r$ becomes $\perp$.

## 4 Discussion

### 4.1 Review of RKA security in the Ideal Cipher Model

We found similar attacks employing the same principle in the context of the RKA security of *the ideal cipher*.

Let us look back over the history of this security. Bellare and Kohno [BK03] showed that the ideal cipher is $\Phi$-RKA secure if $\Phi$ is output-unpredictable and collision-resistant[2]. Let $\mathcal{K}$ and $\mathcal{D}$ denote the key space and the domain of permutations. Let $E, G : \mathcal{K} \times \mathcal{D} \to \mathcal{D}$ be keyed permutations over $\mathcal{D}$. We say that the ideal cipher is $\Phi$-RKA secure if any efficient adversary cannot distinguish the two oracles $E_{\mathsf{rka}}$ and $G_{\mathsf{rka}}$, which, on queries $\phi \in \Phi$ and $x \in \mathcal{D}$, return $E(\phi(K), x)$ and $G(\phi(K), x)$, respectively, even if it can access $E(\cdot, \cdot)$ and $E^{-1}(\cdot, \cdot)$, where $K$, $E$, and $G$ are randomly chosen. Their theorem states that, for any output-unpredictable and collision-resistant $\Phi$, the ideal cipher is $\Phi$-RKA secure.

As Albrecht, Farshim, Paterson, and Watson discussed in [AFPW11], the $\Phi$-RKA security of the ideal cipher holds if the ideal cipher $E$ is chosen independently from $\Phi$. Otherwise, that is, *if* $\Phi$ *depends on* $E$, an attack due to Bernstein (see [AFPW11]) exemplifies an RKA by exploiting the $E$-dependent RKD function class. Consider a class $\Phi_E$ that consists of two RKD functions, $\mathfrak{id} : K \mapsto K$ and $\phi_E : K \mapsto E(K, 0)$. If $E$ is indistinguishable from random permutations, the class is output-unpredictable and collision-resistant. However, these functions give a distinguishing attack. Harris [Har11] also gave a key-recovery RKA by exploiting the $E$-dependent RKD functions. See the details of attacks and discussions in [AFPW11].

Albrecht, Farshim, Paterson, and Watson revisited the RKA security of the ideal cipher in which RKD functions can access $E$ and $E^{-1}$. They defined the oracle-independent properties of a class of RKD functions and showed that the ideal cipher is $\Phi$-RKA secure if $\Phi$ satisfies such properties.

### 4.2 Discussion on the Class of RKD Functions

Our distinguishing and key-recovery attacks against the CNM-KDF in Section 3 also exploit the RKD functions that strongly depend on the algorithms of the scheme. If the RKD functions never change their behavior in the hopping of the games in Section 3.3, the proof in [QLY$^+$15] is correct.

In the ideal cipher model, Albrecht et al. salvaged the RKA security by defining the oracle-independent properties of RKD functions. We note that we (and the challenger) can check if the RKD functions are independent from $E$ because $E$ is *oracle* and an $E$-depending RKD function should have $E$-gates or subroutines calling $E$ explicitly.

In the CNM-KDF case, it seems hard to check if the RKD functions are independent from the algorithms of the scheme because the algorithms are not *oracle*. The easiest way to patch the proof is restricting the RKD functions more explicitly. For example, the class of polynomials is checked easily. Another possible way is restricting the number of functions as in the definitions of (continuous) non-malleable codes. Essentially speaking, the constraint on number will exclude our attacks depending on public information.

## References

ABP15. Michel Abdalla, Fabrice Benhamouda, and Alain Passelègue. An algebraic framework for pseudorandom functions and applications to related-key security. In Rosario Gennaro and Matthew Robshaw, editors, *CRYPTO 2015, Part I*, volume 9215 of *LNCS*, pages 388–409. Springer, Heidelberg, 2015. See also https://eprint.iacr.org/2015/554.

[2] Roughly speaking, we say $\Phi$ is output-unpredictable if any adversary cannot predict $\phi(K)$ and $\Phi$ is collision-resistant if any adversary cannot output two distinct functions $\phi_1, \phi_2 \in \Phi$ satisfying $\phi_1(K) = \phi_2(K)$ (formally, these definitions are given in a concrete-security style).

ABPP14. Michel Abdalla, Fabrice Benhamouda, Alain Passelègue, and Kenneth G. Paterson. Related-key security for pseudorandom functions beyond the linear barrier. In Juan A. Garay and Rosario Gennaro, editors, *CRYPTO 2014, Part I*, volume 8616 of *LNCS*, pages 77–94. Springer, Heidelberg, 2014. See also https://eprint.iacr.org/2014/488.

AFPW11. Martin R. Albrecht, Pooya Farshim, Kenneth G. Paterson, and Gaven J. Watson. On cipher-dependent related-key attacks in the ideal-cipher model. In Antoine Joux, editor, *FSE 2011*, volume 6733 of *LNCS*, pages 128–145. Springer, Heidelberg, 2011. See also https://eprint.iacr.org/2011/213.

AHI11. Benny Applebaum, Danny Harnik, and Yuval Ishai. Semantic security under related-key attacks and applications. In Bernard Chazelle, editor, *ICS 2011*, pages 45–60. Tsinghua University Press, 2011. The full version is available at https://eprint.iacr.org/2010/544.

BC10. Mihir Bellare and David Cash. Pseudorandom functions and permutations provably secure against related-key attacks. In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 666–684. Springer, Heidelberg, 2010. The full version is available at https://eprint.iacr.org/2010/397.

BCM11. Mihir Bellare, David Cash, and Rachel Miller. Cryptography secure against related-key attacks and tampering. In Dong Hoon Lee and Xiaoyun Wang, editors, *ASIACRYPT 2011*, volume 7073 of *LNCS*, pages 486–503. Springer, Heidelberg, 2011. The full version is available at https://eprint.iacr.org/2011/252.

BK03. Mihir Bellare and Tadayoshi Kohno. A theoretical treatment of related-key attacks: RKA-PRPs, RKA-PRFs, and applications. In Eli Biham, editor, *EUROCRYPT 2003*, volume 2656 of *LNCS*, pages 491–506. Springer, Heidelberg, 2003.

BPT12. Mihir Bellare, Kenneth G. Paterson, and Susan Thomson. RKA security beyond the linear barrier: IBE, encryption and signatures. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 331–348. Springer, Heidelberg, 2012. The full version is available at https://eprint.iacr.org/2012/514.

Bih94. Eli Biham. New types of cryptanalytic attacks using related keys. *Journal of Cryptology*, 7(4):229–246, 1994. A preliminary version appeared in *EUROCRYPT '93*, 1993.

DFMV13. Ivan Damgård, Sebastian Faust, Pratyay Mukherjee, and Daniele Venturi. Bounded tamper resilience: How to go beyond the algebraic barrier. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part II*, volume 8270 of *LNCS*, pages 140–160. Springer, Heidelberg, 2013. See also https://eprint.iacr.org/2013/677.

FMVW14. Sebastian Faust, Pratyay Mukherjee, Daniele Venturi, and Daniel Wichs. Efficient non-malleable codes and key-derivation for poly-size tampering circuits. In Phong Q. Nguyen and Elisabeth Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 111–128. Springer, Heidelberg, 2014. See also https://eprint.iacr.org/2013/702.

FX15. Eiichiro Fujisaki and Keita Xagawa. Efficient RKA-secure KEM and IBE schemes against invertible functions. In Kristin Lauter and Francisco Rodríguez-Henríquez, editors, *LATINCRYPT 2015*, volume 9230 of *LNCS*, pages 3–20. Springer, Heidelberg, 2015.

GLM$^+$04. Rosario Gennaro, Anna Lysyanskaya, Tal Malkin, Silvio Micali, and Tal Rabin. Algorithmic tamper-proof (ATP) security: Theoretical foundations for security against hardware tampering. In Moni Naor, editor, *TCC 2004*, volume 2951 of *LNCS*, pages 258–277. Springer, Heidelberg, 2004.

GOR11. Vipul Goyal, Adam O'Neill, and Vanishree Rao. Correlated-input secure hash functions. In Yuval Isahi, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 182–200. Springer, Heidelberg, 2011. The full version is available at https://eprint.iacr.org/2011/233.

Har11. David G. Harris. Critique of the related-key attack concept. *Designs, Codes and Cryptography*, 59(1-3):159–168, 2011.

Hof13. Dennis Hofheinz. Circular chosen-ciphertext security with compact ciphertexts. In Thomas Johansson and Phong Q. Nguyen, editors, *EUROCRYPT 2013*, volume 7881 of *LNCS*, pages 520–536. Springer, Heidelberg, 2013. See also https://eprint.iacr.org/2012/150.

JW15. Zahra Jafargholi and Daniel Wichs. Tamper detection and continuous non-malleable codes. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *TCC 2015, Part I*, volume 9014 of *LNCS*, pages 451–480. Springer, Heidelberg, 2015. See also https://eprint.iacr.org/2014/956.

JLLM14. Dingding Jia, Bao Li, Xianhui Lu, and Qixiang Mei. Related key secure PKE from hash proof systems. In Maki Yoshida and Koichi Mouri, editors, *IWSEC 2014*, volume 8639 of *LNCS*, pages 250–265. Springer, Heidelberg, 2014.

JLLM13. Dingding Jia, Xianhui Lu, Bao Li, and Qixiang Mei. RKA secure PKE based on the DDH and HR assumptions. In Willy Susilo and Reza Reyhanitabar, editors, *ProvSec 2013*, volume 8209 of *LNCS*, pages 271–287. Springer, Heidelberg, 2013.

Knu93. Lars Ramkilde Knudsen. Cryptanalysis of LOKI91. In Jennifer Seberry and Yuliang Zheng, editors, *AUSCRYPT '92*, volume 718 of *LNCS*, pages 196–208. Springer, Heidelberg, 1993.

LMR14. Kevin Lewi, Hart Montgomery, and Ananth Raghunathan. Improved constructions of PRFs secure against related-key attacks. In Ioana Boureanu, Philippe Owesarski, and Serge Vaudenay, editors, *ACNS 2014*, volume 8479 of *LNCS*, pages 44–61. Springer, Heidelberg, 2014.

PW08. Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In Cynthia Dwork, editor, *STOC 2008*, pages 187–196. ACM, 2008.

QL13. Baodong Qin and Shengli Liu. Leakage-resilient chosen-ciphertext secure public-key encryption from hash proof system and one-time lossy filter. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part II*, volume 8270 of *LNCS*, pages 381–400. Springer, Heidelberg, 2013. See also https://eprint.iacr.org/2013/654.

QL14. Baodong Qin and Shengli Liu. Leakage-flexible CCA-secure public-key encryption: Simple construction and free of pairing. In Hugo Krawczyk, editor, *PKC 2014*, volume 8383 of *LNCS*, pages 19–36. Springer, Heidelberg, 2014. See also https://eprint.iacr.org/2015/272.

QLY$^+$15. Baodong Qin, Shengli Liu, Tsz Hon Yuen, Robert H. Deng, and Kefei Chen. Continuous non-malleable key derivation and its application to related-key security. In Jonathan Katz, editor, *PKC 2015*, volume 9020 of *LNCS*, pages 557–578. Springer, Heidelberg, 2015. See also https://eprint.iacr.org/2015/003.

Wee12. Hoeteck Wee. Public key encryption against related key attacks. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 262–279. Springer, Heidelberg, 2012.