A preliminary version of this paper appears in the *ACM Computer and Communications Security (CCS) Conference*, 2017. This is the full version.

Identity-Based Format-Preserving Encryption

Mihir Bellare[1] Viet Tung Hoang[2]

September 2017

Abstract

We introduce identity-based format-preserving encryption (IB-FPE) as a way to localize and limit the damage to format-preserving encryption (FPE) from key exposure. We give definitions, relations between them, generic attacks and two transforms of FPE schemes to IB-FPE schemes. As a special case, we introduce and cover identity-based tweakable blockciphers. We apply all this to analyze DFF, an FPE scheme proposed to NIST for standardization.

[1] Department of Computer Science & Engineering, University of California San Diego, La Jolla, California 92093, USA. Email: `mihir@eng.ucsd.edu`. URL: `http://cseweb.ucsd.edu/~mihir/`. Supported in part by NSF grants CNS-1526801 and CNS-1717640, ERC Project ERCC FP7/615074 and a gift from Microsoft corporation.

[2] Department of Computer Science, Florida State University, Tallahassee, Florida 32304, USA. Email: `tvhoang@cs.fsu.edu`. URL: `http://www.cs.fsu.edu/~tvhoang/`. Supported in part by NSF grant CICI-1738912 and First Year Assistant Professor Award of Florida State University.

Contents

1 Introduction 3

2 Preliminaries 8

3 FPE and IB-FPE 9

4 Attacks 18

5 The PRF construction 20

6 The Dbl construction 24

7 Pre-masking-based IB-FPE 31

8 Security Analysis of DFF 36

References **38**

A Proof of Lemma 7.2 41

B The FF2 and DFF IB-FPE schemes 46

1 Introduction

Schemes for format-preserving encryption (FPE) have been standardized [20] and are in widespread use for the encryption of credit-card numbers. Towards limiting the damage from key exposure, we introduce identity-based FPE (IB-FPE). We provide a provable-security treatment involving definitions, attacks and two design paradigms. We apply this to analyze DFF [39], an FPE scheme proposed to NIST for standardization.

FPE. Format-preserving encryption (FPE) originates with [9, 12]. An FPE scheme F specifies a deterministic encryption function $\mathsf{F.E}: \{0,1\}^{\mathsf{F.kl}} \times \mathsf{F.TS} \times \mathsf{F.Dom} \rightarrow \mathsf{F.Dom}$ that takes a $\mathsf{F.kl}$-bit key J, a tweak T and a message X to return a ciphertext $Y = \mathsf{F.E}(J, T, X)$. There is a corresponding decryption function $\mathsf{F.D}: \{0,1\}^{\mathsf{F.kl}} \times \mathsf{F.TS} \times \mathsf{F.Dom} \rightarrow \mathsf{F.Dom}$ such that the maps $\mathsf{F.E}(J, T, \cdot), \mathsf{F.D}(J, T, \cdot)$ are permutations over $\mathsf{F.Dom}$ that are inverses of each other. What makes FPE special is that the domain $\mathsf{F.Dom}$ can be arbitrary and in particular very small. Some examples are $\mathsf{F.Dom} = \{0,1\}^8$ —encrypt a byte so that the ciphertext is also a byte— $\mathsf{F.Dom} = \mathbb{Z}_{10}^4$ —encrypt a 4 digit PIN so that the ciphertext is also four decimal digits— $\mathsf{F.Dom} = \mathbb{Z}_{10}^{16}$ —encrypt a 16-digit credit-card number so that the result is also a 16-digit credit-card number. FPE is motivated by legacy constraints which in many systems mandate that the ciphertext replace the plaintext, and must thus have the same "format" as the plaintext. Tweakable blockciphers [28] are the special case where $\mathsf{F.Dom} = \{0,1\}^{\mathsf{F.bl}}$ for some integer $\mathsf{F.bl}$ called the block length.

The canonical metric of security for an FPE scheme F is prp security [9, 27]. The game picks a challenge bit b and random key $J \in \{0,1\}^{\mathsf{F.kl}}$. For each tweak T it also lets $\Pi(T, \cdot)$ be a random permutation over $\mathsf{F.Dom}$. The adversary $\mathcal{A}$ can ask for encryption under a tweak T and message X of its choice, being returned $\mathsf{F}(J, T, X)$ if $b = 1$ or $\Pi(T, X)$ if $b = 0$, and similarly for decryption.

FPE is not easy to build. Today the most practical approach is Feistel with strong —AES-based— round functions and a number of rounds $r \geq 8$. NIST SP 800-38G [20] standardizes two such schemes, $\mathsf{FF1}$ ($r = 10$) and $\mathsf{FF3}$ ($r = 8$). Recent attacks [6, 18] suggest that it would be good to increase the number of rounds when the inputs are very short, but this is largely orthogonal to our work.

Corporations offering FPE-based products include HPE Voltage, Verifone, Protegrity, Ingenico, Thales/Vormetric and Gemalto. Tens of millions of credit-cards have been encrypted with these products.

IB-FPE. We define an identity-based FPE (IB-FPE) scheme as a pair $(\mathsf{F}, \mathsf{KDF})$ consisting of a (base) FPE scheme F and an associated *key-derivation function* KDF. The latter takes a master key K and identity I to (deterministically) return a key $J = \mathsf{KDF}(K, I) \in \{0,1\}^{\mathsf{F.kl}}$ for I to use with F.

In the traditional usage of an FPE scheme F, an organization would have a single key K for F stored at many different devices (for example, point-of-sale terminals) that each encrypts directly under K. But each device is at some risk of compromise due to physical, insider or side-channel attacks. Compromise of even one device (which could be quite likely) then has the global consequence of exposure of K. IB-FPE allows us to localize, and thus limit, the damage from key exposure. With IB-FPE, we can associate an identity I to a device and delegate to it the derived key $J_I = \mathsf{KDF}(K, I)$, allowing the device to (effectively) encrypt under K without actually having K. (The master key K would be stored in a secure location, for example in secure hardware.) Compromise of device I would now have only local consequences, encryptions under J_I being compromised but (for an IB-FPE scheme meeting the definitions we will give) encryption under other identities remaining secure.

Another benefit of IB-FPE is to increase the lifetime of the key K. In practice it is recommended to limit the number of encryptions under a particular key, changing (rotating) the key periodically. With u identities each performing q encryptions, direct encryption with a traditional FPE scheme would result in uq encryptions under the base key. With IB-FPE, we have u key derivations under the master key and only q encryptions under each of u different derived keys. This structure can significantly increase the number uq of encryptions that can be safely performed [1, 37].

IB-FPE SECURITY. Let $(\mathsf{F}, \mathsf{KDF})$ be an IB-FPE scheme. We give a prp style definition of security called ib-prp. We also give two key-recovery security definitions called ib-kr-ai and ib-kr-ti. We show relations between them, summarized in Fig. 4.

While natural, these definitions are strong, in particular allowing selective opening attacks [19, 7, 5, 23] that make them hard to provably achieve. We also define non-adaptive versions, which continue to relate to each other as per Fig. 4, and which our schemes are shown to achieve.

THE DEFINITIONS. The ib-prp game picks a random challenge bit b and random master key K, and associates key $J_I = \mathsf{KDF}(K, I) \in \{0,1\}^{\mathsf{F.kl}}$ to identity I. The adversary gets oracle $\textsc{Enc}$ taking identity I, tweak T and message X, and oracle $\textsc{Dec}$ taking I, T and ciphertext Y. Initially, they respond with $\mathsf{F.E}(J_I, T, X)$ and $\mathsf{F.D}(J_I, T, Y)$, respectively. At any point, the adversary can either expose the key of I, querying $\textsc{Exp}(I)$ to get J_I, or switch I to challenge mode by querying $\textsc{Ch}(I)$, restricted, of course to not being able to do both for the same I. If I is switched to challenge mode, oracles $\textsc{Enc}, \textsc{Dec}$ change in the $b = 0$ case, with $\textsc{Enc}(I, T, \cdot)$ and $\textsc{Dec}(I, T, \cdot)$ now becoming permutations that are random but consistent with prior replies.

Theoretical work has traditionally formalized only strong goals that represent the most desirable targets for security proofs, ib-prp in our case. But we also formalize weaker key-recovery security goals (ib-kr-ai and ib-kr-ti). Oracles in the games are like in the $b = 1$ case of ib-prp. The adversary returns a key J' and identity I'. In the ib-kr-ti (target identity) case, it wins if $J' = J_{I'}$ is the key for the identity it names, while in the ib-kr-ai (any identity) case, I' is ignored and it wins if $J' = J_I$ for any un-exposed challenge identity I. The motivation is that (1) We are interested not just in security proofs but in attacks, for which we want to make claims that are strong (violating ib-kr-ai or ib-kr-ti is much more damaging than violating ib-prp) as well as precise (which requires that key-recovery advantages be formalized), and (2) We might be able to prove better security (in terms of bounds on adversary advantage) for ib-kr-ai or ib-kr-ti than for ib-prp.

So far adversaries are adaptive in the sense that they can query $\textsc{Enc}, \textsc{Dec}$ with I before deciding to expose I. We say that an adversary (whether ib-prp, ib-kr-ai or ib-kr-ti) is non-adaptive if its exposure decision for I does not depend on seeing encryptions or decryptions under I: if it queries $\textsc{Exp}(I)$, it has not previously queried $\textsc{Enc}(I, \cdot, \cdot)$ or $\textsc{Dec}(I, \cdot, \cdot)$.

Security in the face of exposure queries captures the above-mentioned application goal that the damage from compromise is local rather than global. (Encryption for an identity is secure even if the keys of other identities are known to the attacker.) Exposure is thus a central element of the framework, and is a powerful adversary capability even in the non-adaptive case. The definition adapts the classical one for IBE [13], differences being that our setting is symmetric (there is no public master key), encryption is deterministic, the goal is prp style security (rather than semantic security) and there are multiple challenge identities, not just one. In the adaptive case, the combination of these elements allows a selective opening attack [19, 7, 5, 23]. We stress that non-adaptive security, even if weaker than adaptive, is hardly a weak notion, and seems more than adequate for practice.

RELATIONS. It is clear that ib-kr-ai security (tightly) implies ib-kr-ti security. (If you can find the key for an identity you name, you can find a key for some identity.) Proposition 3.2 says that, conversely, ib-kr-ti *tightly* implies ib-kr-ai, because, given a candidate key, one can (under some conditions) test to see which identity it matches. We would expect that ib-prp implies ib-kr-ti (and thus, by the above, ib-kr-ai), and while Theorem 3.3, at the highest level, validates this, the truth it shows is more delicate. The difficulty is that in FPE the domain size can be small, and the reduction is parameterized to adjust. The relations, summarized in Fig. 4, hold in both the adaptive and non-adaptive cases.

Attacks. We give attacks on the security of *any* IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$, showing inherent limitations in achievable security. The attacks are strong (they violate non-adaptive ib-kr-ai, not just ib-prp) and rigorously analyzed (Theorems 4.1 and 4.2 provide and prove precise lower bounds on adversary advantage). Their implication is that for $(\mathsf{F}, \mathsf{KDF})$ to have k-bits of (even non-adaptive, ib-kr-ai) security, FPE scheme F must have $2k$-bit keys, regardless of the length of the master key and the choice of KDF. We call this the *double-key condition*.

The challenge with the attacks is to cover *all* IB-FPE schemes $(\mathsf{F}, \mathsf{KDF})$. We give two attacks, calling the first the *matching attack* and the second, which generalizes DP [21], the *exhaustive search attack*. Depending on the value of a quantity we define, called the *diversity* of the key-distribution function KDF, we are able to show that one or the other attack always has constant non-adaptive ib-kr-ai advantage with effort around $2^{\mathsf{F.kl}/2}$.

Building IB-FPE schemes. We now turn to constructing IB-FPE schemes that do as well as possible subject to the limitations uncovered by our attacks. Given that FPE schemes F (satisfying standard prp security) are hard to build, we want to leverage existing constructions of them. Accordingly, our approach is modular: taking as given a (base) FPE scheme F, we design key-derivation functions KDF for it and prove non-adaptive ib-prp security of $(\mathsf{F}, \mathsf{KDF})$ assuming the prp security of F and also possibly assuming something about KDF. We aim to make the master key of KDF as short as we can and to make KDF as efficient as we can. We also aim for instantiations of our key-derivation functions that use only a blockcipher, and moreover one that (like AES) has the same key and block length. (This is because practical FPE schemes already use such blockciphers, as Feistel round functions.) Below we first give a natural, standard-model key-derivation construction **PRF**. Then, to improve efficiency and get an analysis of DFF, we give and analyze an ideal-cipher model construction **Dbl**.

The **PRF** construction. We show in Section 5 that PRFs make good key-derivation functions: If $\mathsf{KDF}\colon \mathsf{KDF.MKS} \times \mathsf{KDF.IS} \to \{0,1\}^{\mathsf{F.kl}}$ is a PRF and base FPE scheme F is prp secure then IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ is non-adaptive ib-prp secure. We call this the **PRF** construction of an IB-FPE scheme. Assuming $\mathsf{F.kl} = 2k$, the concrete reduction, as given by Theorem 5.1, implies that if KDF has k-bits of prf security and F has $2k$-bits of prp security then $(\mathsf{F}, \mathsf{KDF})$ has k bits of non-adaptive ib-prp security. Our attacks discussed above imply that the reduction is optimal.

For an instantiation we would like to base KDF solely on AES and achieve full 128-bit security with the master key being a (128-bit) AES key. Abstractly, assuming given a base FPE scheme F that has $2k$ bits of prp security with $\mathsf{F.kl} = 2k$, this means that we want to build $\mathsf{KDF}\colon \{0,1\}^k \times \mathsf{KDF.IS} \to \{0,1\}^{2k}$, with k bits of prf security, solely from a blockcipher $E\colon \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ having k bits of prp-cpa security. This is a challenging goal, but we can reach it via DHT's new analysis [15] of the XOR prp-to-prf transform of BKR [8]. Our key-derivation function, shown in Fig. 9, has a computational cost of four invocations of the blockcipher E.

In summary, the **PRF** construction instantiated as above is an efficient way to generically turn an FPE scheme into an IB-FPE scheme with optimal security, a standard-model proof and a reasonable key-derivation cost of four blockcipher invocations. There are two motivations for the alternative key-derivation method that follows: (1) Our results about it will eventually yield an analysis of the DFF scheme proposed to NIST for standardization, and (2) It uses only two blockcipher invocations.

The **Dbl** construction. Letting F be the given prp-secure FPE scheme with $\mathsf{F.kl} = 2k$, our **Dbl** ("Double") construction of an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ lets $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ be a blockcipher and then defines key-derivation function $\mathsf{KDF} : \{0,1\}^k \times \mathsf{KDF.IS} \to \{0,1\}^{2k}$ by

$$\mathsf{KDF}(K, I) = E(K, \mathsf{M}_0(I)) \,\|\, E(K, \mathsf{M}_1(I)) \,, \tag{1}$$

where $\mathsf{M}_0, \mathsf{M}_1 : \mathsf{KDF.IS} \to \{0,1\}^k$ are injective functions with disjoint ranges. We refer to M as an embedding scheme, and it parameterizes the construction. Theorem 6.1 implies that $(\mathsf{F}, \mathsf{KDF})$ has k-bits of ib-prp security assuming F has $2k$-bits of prp security and E is an ideal cipher. The double-key condition emanating from our attacks says that the analysis of Theorem 6.1 is optimal. Next we discuss some technical elements of the result.

One might have hoped to establish prf security of KDF in the ideal-cipher model and then apply our result about **PRF**, but, even in the ideal-cipher model, the key-derivation function KDF of Eq. (1) has only $k/2$ bits of prf security. Instead we give a direct analysis.

In practice we expect that $E = \mathsf{AES}$ will be used, not only by KDF, but also by F. To model this, we allow F to have oracle access to the *same* ideal cipher E that is used by KDF. This common use of the ideal primitive precludes a modular proof and makes the analysis more challenging. Given an ib-prp adversary $\mathcal{A}$ against F under KDF, the reduction aims to build a prp adversary $\overline{\mathcal{A}}$ and bound ϵ, the ib-prp advantage of $\mathcal{A}$ against F, KDF, as a function of $\overline{\epsilon}$, the prp advantage of $\overline{\mathcal{A}}$ against F. The natural approach is a hybrid argument. The difficulty is that, due to the structure of KDF, keys of different users are not statistically independent. If u is the number of users invoked by $\mathcal{A}$, the straightforward hybrid argument would incur a loss of $O(u/2^k)$ per hybrid step, resulting in a bound of the form $\epsilon \le u\overline{\epsilon} + \delta$ where $\delta = O(u^2/2^k)$. This would imply only $k/2$ bits of security for F under KDF, well short of what we want and believe to be true. Theorem 6.1 gives a different proof that includes a more sophisticated hybrid argument to obtain $\delta = O(u/2^k)$, which implies k-bit ib-prp security for $(\mathsf{F}, \mathsf{KDF})$, as desired.

IB-FPE from Pre-masking FPE. **Dbl** builds an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ assuming as given the base FPE scheme $\mathsf{F} : \{0,1\}^{2k} \times \mathsf{F.TS} \times \mathsf{F.Dom} \to \mathsf{F.Dom}$. We now ask if the assumption can be dropped. That is, we want to build a practical F from our blockcipher $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ so that, with KDF as in Eq. (1), IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ can be shown to have k bits of ib-prp security assuming nothing more than ideality of E. The difficulty is that practical FPE schemes F are mostly Feistel-based, and Feistel (as we explain further in Section 7) notoriously lacks tight analyses showing prp security for small domains and number of rounds. However we show that the goal can be reached if we target key-recovery security rather than prp security.

Our results are quite general. We define a class of FPE schemes that we call pre-masking. This class includes Feistel-based schemes. The schemes use a blockcipher $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ but have $2k$-bit keys. Encryption and decryption do not have direct access to the key but can call an oracle that uses the key in conjunction with the blockcipher in a restricted way. (See Section 7 for the full definition.) Now, take any F in this class and adjoin the key-derivation function KDF of **Dbl** as per Eq. (1) to get IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$. Then Theorem 7.3 establishes that $(\mathsf{F}, \mathsf{KDF})$ has k bits of ib-kr-ti security.

Security of DFF. Two FPE schemes proposed to NIST for standardization, namely FF2 [38] — not standardized due to the attack of [21]— and DFF [39] —still under consideration— derive a subkey from the tweak and then encrypt under the subkey with an un-tweaked cipher. The authors highlight this method as providing a feature they call delegation, where knowledge of the subkey for one tweak would not impact security of encryption under another tweak. Our IB-FPE framework allows us to formalize this claim and evaluate the security, relative to it, of DFF.

We view $\mathsf{FF2} = (\mathsf{F}_{\mathsf{ff2}}, \mathsf{KDF}_{\mathsf{ff2}})$ and $\mathsf{DFF} = (\mathsf{FF}_{\mathsf{dff}}, \mathsf{KDF}_{\mathsf{dff}})$ as IB-FPE schemes with identity space the tweak space of the original scheme, and a tweak space that is trivial, consisting, say, of just the empty string. In $\mathsf{FF2}$, the master key K is 128 bits and the key delegated to I is $J_I = \mathsf{KDF}_{\mathsf{ff2}}(K, I) = \mathsf{AES}(K, I)$. Since $\mathsf{F}_{\mathsf{ff2}}$ (accordingly) has 128-bit keys, our attacks from Section 4 say that $\mathsf{FF2}$ has at most 64 bits of ib-kr-ti security, explaining the Dworkin and Perlner (DP) attack [21]. Arguing that a scheme with a 128-bit (master) key should provide 128-bits of security, NIST rejected $\mathsf{FF2}$. Seeking 128 bits of security, DFF continues to have a 128-bit master key, but derived keys (and thus keys for $\mathsf{FF}_{\mathsf{dff}}$) are 256 bits long. Our attacks indicate that the security is at most 128 bits. The question relevant to standardization is whether it actually is 128, or significantly less.

Let the domain be $\mathbb{Z}_{\mathsf{rdx}}^n$, the set of length n strings over alphabet $\mathbb{Z}_{\mathsf{rdx}}$ ($2 \leq \mathsf{rdx}, n < 2^8$), and regard rdx, n as fixed. Let $k = 128$. Then the key-derivation function $\mathsf{KDF}_{\mathsf{dff}}$ of $\mathsf{DFF} = (\mathsf{FF}_{\mathsf{dff}}, \mathsf{KDF}_{\mathsf{dff}})$ can be viewed as obtained by applying our **Dbl** transform with $E = \mathsf{AES}$ and embedding scheme M defined by $\mathsf{M}_0(I) = [\mathsf{rdx}]^1 \| [|I|]^1 \| [n]^1 \| [I]^{13}$ and $\mathsf{M}_1(I) = [0]^3 \| [I]^{13}$, where $[x]^\ell$ denotes the encoding of x as an ℓ-byte string. Then (1) our results from Section 6 say that DFF has about 128 bits of ib-prp security if E is ideal and FPE scheme $\mathsf{FF}_{\mathsf{dff}}$ is assumed to have about 256 bits of prp security, and (2) Observing that $\mathsf{FF}_{\mathsf{dff}}$ is an E-based pre-masking FPE scheme, our results from Section 7 say that DFF has about 128 bits of ib-kr-ti security assuming only that E is ideal.

There is, however, a caveat that our analysis uncovers. For our results to apply, the functions $\mathsf{M}_0, \mathsf{M}_1$ defined above must be injective. This is not, strictly speaking, true for DFF, because the identity space is the set of all binary strings of length at most 13 bytes, and so, for example, $\mathsf{M}_1(001) = \mathsf{M}_1(01)$. It is true (our conditions on M are met) if we restrict identities further, for example to all have the same length, or so that no two represent, in binary, the same integer. For the general case we have neither a proof, nor an attack showing security to be significantly smaller than the desired 128 bits. In Section 8 we expand on this and also give the best attack we know for the general case. We would suggest that the embedding function used in DFF be changed to meet our conditions, so that our results would apply to validate security in the general case as well. For example, let identities be binary strings of at most 12 bytes, let $\mathsf{M}_0(I) = [0]^1 \| [\mathsf{rdx}]^1 \| [|I|]^1 \| [n]^1 \| [I]^{12}$ and $\mathsf{M}_1(I) = [1]^1 \| [\mathsf{rdx}]^1 \| [|I|]^1 \| [n]^1 \| [I]^{12}$.

We clarify that, as designed and presented in [38, 39], $\mathsf{FF2}$ and DFF are FPE schemes targeting key delegation based on tweaks, not IB-FPE schemes. To translate findings above back to the original context, read "tweak" for "identity".

Related work. Identity-based cryptography was suggested by Shamir [36]. Identity-based encryption (IBE) was formalized and achieved by BF [13].

BHT [6] give message-recovery attacks on Feistel-based FPE schemes F, including the $\mathsf{FF1}$ and $\mathsf{FF3}$ standards [20] and $\mathsf{FF}_{\mathsf{dff}}$, in the case that the domain is tiny. DV [18] give small-domain attacks on $\mathsf{FF3}$. $\mathsf{FF1}$ and $\mathsf{FF3}$ are not relevant for us. (Having 128 bit keys, they cannot, by our attacks, be base schemes for high-security IB-FPE.) For $\mathsf{F} = \mathsf{FF}_{\mathsf{dff}}$, the validity of Theorems 5.1 (for **PRF**) and 6.1 (for **Dbl**) is not affected, but to get the full possible k-bits of security for the IB-FPE scheme $(\mathsf{FF}_{\mathsf{dff}}, \mathsf{KDF})$ from these results, one would have to increase the number of rounds in $\mathsf{FF}_{\mathsf{dff}}$ for tiny inputs. The BHT attacks do not contradict our proof of ib-kr-ti security of DFF because they are message-recovery attacks and do not succeed in key recovery.

Shuffle-based FPE schemes [22, 30, 34] are a possible choice in the role of F to obtain IB-FPE schemes via the **PRF** or **Dbl** constructions. For efficiency, however, schemes in practice, including $\mathsf{FF}_{\mathsf{dff}}$, have been Feistel based, so we have focused on the latter in considering instantiating F via

```
Game $\mathbf{G}^{\mathsf{prf}}_{\mathsf{GG}}(\mathcal{A})$
$b \leftarrow\!\!\$ \{0,1\}; K \leftarrow\!\!\$ \mathsf{GG.Keys}$
$b' \leftarrow\!\!\$ A^{\mathrm{Fn}}$ ; Return $(b' = b)$

$\mathrm{Fn}(X)$
If $\mathrm{T}[X] \neq \bot$ then return $\mathrm{T}[X]$
If $b = 0$ then $\mathrm{T}[X] \leftarrow\!\!\$ \mathsf{GG.Rng}$
Else $\mathrm{T}[X] \leftarrow \mathsf{GG}(K, X)$
Return $\mathrm{T}[X]$
```

```
Game $\mathbf{G}^{\mathsf{prp\text{-}cpa}}_{\mathsf{GG}}(\mathcal{A})$
$b \leftarrow\!\!\$ \{0,1\}; K \leftarrow\!\!\$ \mathsf{GG.Keys}$
$b' \leftarrow\!\!\$ A^{\mathrm{Fn}}$ ; Return $(b' = b)$

$\mathrm{Fn}(X)$
If $\mathrm{ET}[X] \neq \bot$ then return $\mathrm{ET}[X]$
If $b = 0$ then $Y \leftarrow\!\!\$ \{\, Y \in \mathsf{GG.Dom} \,:\, \mathrm{DT}[Y] = \bot \,\}$
Else $Y \leftarrow \mathsf{GG}(K, X)$
$\mathrm{ET}[X] \leftarrow Y$ ; $\mathrm{DT}[Y] \leftarrow X$ ; Return $Y$
```

Figure 1: **Games defining PRF security (left) and PRP-CPA security (right) of GG.**

pre-masking FPE schemes.

2 Preliminaries

Notation and conventions. We let ε denote the empty string. If y is a string then $|y|$ denotes its length and $y[i]$ denotes its i-th bit for $1 \le i \le |y|$, and for $1 \le i \le j \le |y|$, let $y[i:j] = y[i] \cdots y[j]$. If X is a finite set, we let $x \leftarrow\!\!\$ X$ denote picking an element of X uniformly at random and assigning it to x. Algorithms may be randomized unless otherwise indicated. Running time is worst case. If A is an algorithm, we let $y \leftarrow A(x_1, \ldots; r)$ denote running A with random coins r on inputs $x_1, \ldots$ and assigning the output to y. We let $y \leftarrow\!\!\$ A(x_1, \ldots)$ be the result of picking r at random and letting $y \leftarrow A(x_1, \ldots; r)$. We let $[A(x_1, \ldots)]$ denote the set of all possible outputs of A when invoked with inputs $x_1, \ldots$.

We use the code based game playing framework of [10]. By $\Pr[\mathrm{G} \Rightarrow y]$ we denote the event that the execution of game G results in the game returning y. We write $\Pr[\mathrm{G}]$ as an abbreviation of $\Pr[\mathrm{G} \Rightarrow \mathsf{true}]$. In code of games, unless otherwise indicated, sets are assume initialized to empty, booleans to false, integers to 0 and anything else to $\bot$. We adopt the convention that the running time of an adversary refers to the worst-case execution time of the game with the adversary, so that the time for the execution of oracles to compute replies to oracle queries is included. This means that usually in reductions, adversary running time is roughly maintained.

If D, R are sets then $\mathrm{Func}(\mathsf{D}, \mathsf{R})$ denotes the set of all functions from domain D to range R, and $\mathrm{Perm}(\mathsf{D})$ the set of all permutations on D.

PRFs and PRPs. Recall that the prf advantage of an adversary $\mathcal{A}$ against a family of functions $\mathsf{GG} : \mathsf{GG.Keys} \times \mathsf{GG.Dom} \to \mathsf{GG.Rng}$ is defined as $\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{GG}}(\mathcal{A}) = 2\Pr[\mathbf{G}^{\mathsf{prf}}_{\mathsf{GG}}(\mathcal{A})] - 1$, where game $\mathbf{G}^{\mathsf{prf}}_{\mathsf{GG}}(\mathcal{A})$ is shown in Fig. 1. Also the prp-cpa advantage of an adversary $\mathcal{A}$ against a family of permutations $\mathsf{GG} : \mathsf{GG.Keys} \times \mathsf{GG.Dom} \to \mathsf{GG.Dom}$ is defined as $\mathbf{Adv}^{\mathsf{prp\text{-}cpa}}_{\mathsf{GG}}(\mathcal{A}) = 2\Pr[\mathbf{G}^{\mathsf{prp\text{-}cpa}}_{\mathsf{GG}}(\mathcal{A})] - 1$, where game $\mathbf{G}^{\mathsf{prp\text{-}cpa}}_{\mathsf{GG}}(\mathcal{A})$ is shown in Fig. 1.

Ideal primitives. An *ideal primitive* is defined simply as a set of functions. An instance (meaning, a particular function) P will be picked at random in the games and provided as an oracle, to algorithms that need it and to the adversary. For example, the ideal primitive corresponding to a random oracle with domain D and range R is $\mathrm{Func}(\mathsf{D}, \mathsf{R})$. Ideal ciphers are a bit more work since one must give access to both the map and its inverse. If K, D are sets then $\mathbf{IC}(\mathsf{K}, \mathsf{D})$ is the set of all maps $P : \mathsf{K} \times \mathsf{D} \times \{+, -\} \to \mathsf{D}$ with the property that $P(K, \cdot, +), P(K, \cdot, -) \in \mathrm{Perm}(\mathsf{D})$ are inverses of each other for every $K \in \mathsf{K}$. If $P \leftarrow\!\!\$ \mathbf{IC}(\mathsf{K}, \mathsf{D})$, and then P is provided as an oracle, we are in the ideal cipher model where one has oracle access to both the cipher $P(\cdot, \cdot, +)$ and its inverse

$P(\cdot,\cdot,-)$. As an abbreviation, we let $\mathbf{IC}(k,n) = \mathbf{IC}(\{0,1\}^k, \{0,1\}^n)$, capturing ideal blockciphers with key length k and block length n.

A USEFUL INEQUALITY. In some proofs we'll use the following.

Lemma 2.1 *Let $p \geq 1$ be an integer and $a \geq 0$ a real number. Assume $ap \leq 1$. Then $(1-a)^p \leq 1 - pa/2$.*

Proof: Let $x = 2 - a(p-1)$. We assumed $ap \leq 1$ and $a \geq 0$. This implies that $x = 2 - ap + a \geq 2 - ap \geq 2 - 1 = 1$. By inclusion-exclusion we have

$$\begin{aligned} (1-a)^p &\leq 1 - pa + \binom{p}{2}a^2 \\ &= 1 - pa + \frac{p(p-1)}{2}a^2 = 1 - \frac{pa}{2}\left(2 - (p-1)a\right) \\ &= 1 - \frac{pa}{2}x \leq 1 - \frac{pa}{2} . \end{aligned}$$

The last inequality is because $x \geq 1$. ▌

3 FPE and IB-FPE

We give definitions and basic results, including relations between notions, for FPE and IB-FPE.

FPE SCHEMES. A *format-preserving encryption* (FPE) scheme F [9, 12] specifies a deterministic encryption algorithm $\mathsf{F.E} : \{0,1\}^{\mathsf{F.kl}} \times \mathsf{F.TS} \times \mathsf{F.Dom} \to \mathsf{F.Dom}$ together with a deterministic decryption algorithm $\mathsf{F.D} : \{0,1\}^{\mathsf{F.kl}} \times \mathsf{F.TS} \times \mathsf{F.Dom} \to \mathsf{F.Dom}$. Here $\{0,1\}^{\mathsf{F.kl}}$ is the keyspace, $\mathsf{F.Dom}$ is the domain and $\mathsf{F.TS}$ is the tweak space. For every key $J \in \{0,1\}^{\mathsf{F.kl}}$ and tweak $T \in \mathsf{T}$, the functions $\mathsf{F.E}(J,T,\cdot), \mathsf{F.D}(J,T,\cdot) \in \mathrm{Perm}(\mathsf{F.Dom})$ are permutations over $\mathsf{F.Dom}$ that are inverses of each other. We refer to $\mathsf{F.kl}$ as the key length. The scheme may have an associated ideal primitive $\mathsf{F.IP}$, in which case $\mathsf{F.E}, \mathsf{F.D}$ have oracle access to a function $P \in \mathsf{F.IP}$. Tweakable blockciphers [28] are a special case: FPE scheme F is a *tweakable blockcipher* if $\mathsf{F.Dom} = \{0,1\}^{\mathsf{F.bl}}$ for an integer $\mathsf{F.bl}$ called the blocklength.

FPE SECURITY. We recall the standard prp metric for an FPE scheme F [9, 12]. It coincides with the classic (strong) tweakable-prp metric of [27] in the case that F is a tweakable blockcipher. Let $\mathcal{A}$ be an adversary and define $\mathbf{Adv}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A}) = 2\Pr[\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A})] - 1$, where game $\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A})$ is on the left in Fig. 2. The game picks a random challenge bit b and runs the adversary. The latter gets oracles $\textsc{Enc}, \textsc{Dec}$ for encryption and decryption, and access to an instance $\mathbf{P}$ of the ideal primitive $\mathsf{F.IP}$. It returns a bit b' and wins if $b' = b$. $\textsc{Enc}$ takes a tweak T and message X and returns ciphertext Y, with $\textsc{Dec}$ correspondingly taking tweak T and ciphertext Y to return message X. If $b = 1$, encryption and decryption are done using F with key J. If $b = 0$, each tweak is associated with a random permutation on $\mathsf{F.Dom}$ under which both encryption and decryption are done.

Letting $\mathcal{A}$ again be an adversary, we also define $\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A}) = 2\Pr[\mathbf{G}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A})] - 1$, where game $\mathbf{G}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A})$ is on the right in Fig. 2. This captures what we call adaptive prp security, a notion we will find useful for proofs. Oracles $\textsc{Enc}$ and $\textsc{Dec}$ use F under key J until the adversary calls $\textsc{Ch}$ to switch the game to challenge mode by setting flag ch to true. At that point, for each tweak, the associated permutation starts behaving randomly but consistent with the prior queries and (F-based) answers for that tweak. The prp notion corresponds to the special case where the first query is $\textsc{Ch}()$. We will exploit the following, which says that adaptivity can increase advantage by a factor of at most two in general.

```
Game $\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A})$
$b \leftarrow\!\!\$ \{0,1\}$ ; $J \leftarrow\!\!\$ \{0,1\}^{\mathsf{F.kl}}$ ; $P \leftarrow\!\!\$ \mathsf{F.IP}$
$b' \leftarrow\!\!\$ \mathcal{A}^{\textsc{Enc},\textsc{Dec},P}$ ; Return $(b = b')$

$\textsc{Enc}(T, X)$
If $\mathrm{ET}[T, X] \neq \bot$ then return $\mathrm{ET}[T, X]$
If $b = 0$ then
   $Y \leftarrow\!\!\$ \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[T, Y] = \bot \,\}$
Else $Y \leftarrow \mathsf{F.E}^P(J, T, X)$
$\mathrm{ET}[T, X] \leftarrow Y$ ; $\mathrm{DT}[T, Y] \leftarrow X$ ; Return $Y$

$\textsc{Dec}(T, Y)$
If $\mathrm{DT}[T, Y] \neq \bot$ then return $\mathrm{DT}[T, Y]$
If $b = 0$ then
   $X \leftarrow\!\!\$ \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[T, X] = \bot \,\}$
Else $X \leftarrow \mathsf{F.D}^P(J, T, Y)$
$\mathrm{ET}[T, X] \leftarrow Y$ ; $\mathrm{DT}[T, Y] \leftarrow X$ ; Return $X$
```

```
Game $\mathbf{G}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A})$
$b \leftarrow\!\!\$ \{0,1\}$ ; $J \leftarrow\!\!\$ \{0,1\}^{\mathsf{F.kl}}$ ; $P \leftarrow\!\!\$ \mathsf{F.IP}$ ; ch $\leftarrow$ false
$b' \leftarrow\!\!\$ \mathcal{A}^{\textsc{Enc},\textsc{Dec},\textsc{Ch},P}$ ; Return $(b = b')$

$\textsc{Enc}(T, X)$
If $\mathrm{ET}[T, X] \neq \bot$ then return $\mathrm{ET}[T, X]$
If (ch and $b = 0$) then
   $Y \leftarrow\!\!\$ \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[T, Y] = \bot \,\}$
Else $Y \leftarrow \mathsf{F.E}^P(J, T, X)$
$\mathrm{ET}[T, X] \leftarrow Y$ ; $\mathrm{DT}[T, Y] \leftarrow X$ ; Return $Y$

$\textsc{Dec}(T, Y)$
If $\mathrm{DT}[T, Y] \neq \bot$ then return $\mathrm{DT}[T, Y]$
If (ch and $b = 0$) then
   $X \leftarrow\!\!\$ \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[T, X] = \bot \,\}$
Else $X \leftarrow \mathsf{F.D}^P(J, T, Y)$
$\mathrm{ET}[T, X] \leftarrow Y$ ; $\mathrm{DT}[T, Y] \leftarrow X$ ; Return $X$

$\textsc{Ch}()$
ch $\leftarrow$ true
```

Figure 2: **Games defining security of an FPE scheme F. Left: prp. Right: prpa.**

Proposition 3.1 *Let F be an FPE scheme. Given a prpa adversary $\mathcal{A}_{\mathrm{prp}}$, we can build a prp adversary $\mathcal{A}_{\mathrm{prp}}$ of about the same running time, and making at most as many* Fn *queries, such that $\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}}) \leq 2 \cdot \mathbf{Adv}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}})$.*

Proof of Proposition 3.1: The adversary $\mathcal{A}_{\mathrm{prp}}$ first picks a bit $a \leftarrow\!\!\$ \{0,1\}$ and then runs $\mathcal{A}_{\mathrm{prp}}$. Before the latter calls Ch, the former always use its Enc/Dec oracles to reply to the Enc/Dec queries of $\mathcal{A}_{\mathrm{prp}}$. After $\mathcal{A}_{\mathrm{prp}}$ has called Ch to enter the challenge phase, if $a = 1$ then $\mathcal{A}_{\mathrm{prp}}$ continues to use its Enc/Dec oracles to reply to $\mathcal{A}_{\mathrm{prp}}$'s Enc/Dec queries. However, if $a = 0$ then $\mathcal{A}_{\mathrm{prp}}$ gives answers that are random but still consistent with prior queries and answers. When $\mathcal{A}_{\mathrm{prp}}$ outputs its guess b' then $\mathcal{A}_{\mathrm{prp}}$ outputs 1 if $b' = a$, and outputs 0 otherwise. Let c_{prp} be the challenge bit of game $\mathbf{G}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}})$. We claim that

$$\Pr[\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}}) \Rightarrow \mathsf{true} \mid c_{\mathrm{prp}} = 1] = \Pr[\mathbf{G}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}})] \tag{2}$$

$$\Pr[\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}}) \Rightarrow \mathsf{false} \mid c_{\mathrm{prp}} = 0] = \frac{1}{2}\ . \tag{3}$$

This is because (1) when $c_{\mathrm{prp}} = 0$, the answers for $\mathcal{A}_{\mathrm{prp}}$'s Enc and Dec queries are always simulated via an ideal family of permutations, meaning that whatever $\mathcal{A}_{\mathrm{prp}}$ receives is independent of a, but (2) when $c_{\mathrm{prp}} = 1$, the guess of $\mathcal{A}_{\mathrm{prp}}$ is incorrect if and only if $b' = a$. Subtracting Eq. (2) and Eq. (3) side by side we have

$$\mathbf{Adv}^{\mathsf{prp}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}}) = \frac{1}{2}\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\mathcal{A}_{\mathrm{prp}})$$

as claimed. ▌

Proposition 3.1 says that prpa is an alternative, equivalent (up to a factor two in advantage) characterization of classic (strong) prp security for FPE schemes and tweakable blockciphers. For

```
Game $\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$
$b \leftarrow_{\$} \{0,1\}$ ; $K \leftarrow_{\$} \mathsf{KDF.MKS}$
$\mathrm{XI} \leftarrow \emptyset$ ; $\mathrm{ChI} \leftarrow \emptyset$ ; $P \leftarrow_{\$} \mathsf{F.IP}$
For every $I \in \mathsf{KDF.IS}$ do $J_I \leftarrow \mathsf{KDF}^P(K, I)$
$b' \leftarrow_{\$} \mathcal{A}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$ ; Return $(b = b')$

$\textsc{Enc}(I, T, X)$
If $\mathrm{ET}[I, T, X] \neq \bot$ then return $\mathrm{ET}[I, T, X]$
If $(I \in \mathrm{ChI}$ and $b = 0)$ then
  $Y \leftarrow_{\$} \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[I, T, Y] = \bot \,\}$
Else $Y \leftarrow \mathsf{F.E}^P(J_I, T, X)$
$\mathrm{ET}[I, T, X] \leftarrow Y$ ; $\mathrm{DT}[I, T, Y] \leftarrow X$ ; Return $Y$

$\textsc{Dec}(I, T, Y)$
If $\mathrm{DT}[I, T, Y] \neq \bot$ then return $\mathrm{DT}[I, T, Y]$
If $(I \in \mathrm{ChI}$ and $b = 0)$ then
  $X \leftarrow_{\$} \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[I, T, X] = \bot \,\}$
Else $X \leftarrow \mathsf{F.D}^P(J_I, T, Y)$
$\mathrm{ET}[I, T, X] \leftarrow Y$ ; $\mathrm{DT}[I, T, Y] \leftarrow X$ ; Return $X$

$\textsc{Exp}(I)$
If $I \in \mathrm{ChI}$ then return $\bot$
$\mathrm{XI} \leftarrow \mathrm{XI} \cup \{I\}$ ; Return $J_I$

$\textsc{Ch}(I)$
If $I \in \mathrm{XI}$ then return $\bot$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
```

```
Game $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$ / $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$
$K \leftarrow_{\$} \mathsf{KDF.MKS}$
$\mathrm{XI} \leftarrow \emptyset$ ; $\mathrm{ChI} \leftarrow \emptyset$ ; $\mathrm{ChK} \leftarrow \emptyset$ ; $P \leftarrow_{\$} \mathsf{F.IP}$
For every $I \in \mathsf{KDF.IS}$ do $J_I \leftarrow \mathsf{KDF}^P(K, I)$
$(J', I') \leftarrow_{\$} \mathcal{A}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
Return $((J', I') \in \mathrm{ChK})$ // ib-kr-ti
Return $(\exists\, I : (J', I) \in \mathrm{ChK})$ // ib-kr-ai

$\textsc{Enc}(I, T, X)$
Return $\mathsf{F.E}^P(J_I, T, X)$

$\textsc{Dec}(I, T, Y)$
Return $\mathsf{F.D}^P(J_I, T, Y)$

$\textsc{Exp}(I)$
If $I \in \mathrm{ChI}$ then return $\bot$
$\mathrm{XI} \leftarrow \mathrm{XI} \cup \{I\}$ ; Return $J_I$

$\textsc{Ch}(I)$
If $I \in \mathrm{XI}$ then return $\bot$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
$\mathrm{ChK} \leftarrow \mathrm{ChK} \cup \{(J_I, I)\}$
```

Figure 3: **Games defining security of an IB-FPE scheme** $(\mathsf{F}, \mathsf{KDF})$**. Left: ib-prp. Right: ib-kr-ti and ib-kr-ai.**

untweaked blockciphers, Desai and Miner [17] consider a notion of indistinguishable uniform permutation that is prpa with the adversary restricted to just one post-challenge encryption query and no decryption queries, showing it is equivalent to classic prp security up a factor two in advantage. Our proof extends theirs.

For FPE, we do not need to consider key-recovery security. We will for IB-FPE.

IB-FPE. A *key-derivation function* for FPE scheme F is a function $\mathsf{KDF} : \mathsf{KDF.MKS} \times \mathsf{KDF.IS} \to \{0,1\}^{\mathsf{F.kl}}$ that takes a *master key* K in the master-key space $\mathsf{KDF.MKS}$ and a *user identity* I in the identity-space $\mathsf{KDF.IS}$ to return a key $\mathsf{KDF}(K, I) \in \{0,1\}^{\mathsf{F.kl}}$ for F. An *identity-based FPE* (IB-FPE) scheme is a pair $(\mathsf{F}, \mathsf{KDF})$ consisting of a (base) FPE scheme F and a key-derivation function KDF for F. An IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ is an *identity-based tweakable blockcipher* if F is a tweakable blockcipher.

The key-derivation function KDF may have an associated ideal primitive, denoted $\mathsf{KDF.IP}$, in which case KDF has oracle access to a function $P \in \mathsf{KDF.IP}$. We require that $\mathsf{F.IP} = \mathsf{KDF.IP}$, meaning the ideal primitive of the key-derivation function is the same as that of the FPE scheme, and in games a single instance P of the ideal primitive will be used as the oracle for $\mathsf{F.E}, \mathsf{F.D}$ and KDF. This is not only for simplicity but, more importantly, because the primitive in practice is often instantiated via the same cryptographic function, for example via AES.

IB-FPE SECURITY. Security requires that encryption under the key of some identity remains secure even if the adversary can obtain the keys of other identities. In terms of application and motivation, an identity might represent a point-of-sale terminal as discussed in Section 1, and thus our security requirement ensures that the damage from compromise of a terminal remains local, not affecting the security of encryption performed by other terminals. We give a prp style notion, ib-prp. We also give two variants of key-recovery security, ib-kr-ai and ib-kr-ti. The core notions are adaptive, but each has a corresponding non-adaptive version, obtained by restricting attention to non-adaptive adversaries as defined below. We establish relations between the notions as summarized in Fig. 4. The shown relations hold in both the adaptive and non-adaptive cases.

IB-PRP SECURITY. Let $(\mathsf{F}, \mathsf{KDF})$ be an IB-FPE scheme and $\mathcal{A}$ an adversary we call an ib-prp adversary. Define

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) = 2\Pr[\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})] - 1,$$

where game $\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$ is on the left in Fig. 3. The game picks a random challenge bit b and runs the adversary. The latter gets oracles $\textsc{Enc}, \textsc{Dec}$ for encryption and decryption, an expose oracle $\textsc{Exp}$, a challenge oracle $\textsc{Ch}$ and access to an instance P of the ideal primitive $\mathsf{F.IP} = \mathsf{KDF.IP}$. It returns a bit b' and wins if $b' = b$. XI is the set of exposed identities and ChI is the set of challenge identities. These sets stay disjoint throughout the game. Let us refer to an identity as neutral if it is in neither of these sets. All identities start neutral, since the sets $\mathrm{XI}, \mathrm{ChI}$ are initialized to empty. Encryption oracle $\textsc{Enc}$ takes an identity I, tweak T and message X and returns ciphertext Y, while decryption oracle $\textsc{Dec}$ correspondingly taking identity I, tweak T and ciphertext Y to return message X. For neutral identities (and thus at the start of the game), these oracles behave honestly, meaning use F under keys derived via KDF under master key K, regardless of the value of the challenge bit b. Imagine the adversary querying these for a while. Adaptively, at any point in this process, it can either expose the key of an identity I via a $\textsc{Exp}(I)$ query (this captures real-world compromise of the key of this identity), or switch I to challenge mode via a $\textsc{Ch}(I)$ query. If I is exposed, the encryption and decryption oracles for it continue to behave honestly. If I is switched to a challenge identity, then encryption and decryption continue to behave honestly if $b = 1$, but, if $b = 0$, they use, for any given tweak, a permutation that is random subject to being consistent with prior queries and replies for that identity and tweak.

IB-KR SECURITY. Let $(\mathsf{F}, \mathsf{KDF})$ be an IB-FPE scheme and $\mathcal{A}$ an adversary we call an ib-kr adversary. Define

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) = \Pr[\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})]$$
$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) = \Pr[\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})] ,$$

where the games are defined (together, they differ on just one indicated line) on the right in Fig. 3. There is no challenge bit, and the encryption and decryption oracles are always honest, using F. Oracle $\textsc{Exp}$ again allows key exposure. Choice of challenge identities is again adaptive, meaning an identity can be named as a challenge one after encryption and decryption queries, either to it or to other identities. The adversary returns a key and an identity. In the target-identity case ($\mathsf{ib\text{-}kr\text{-}ti}$), it wins if the key it provides is the correct one for the identity it provides. In the all-identity ($\mathsf{ib\text{-}kr\text{-}ai}$) case, the identity it provides is ignored, and the adversary wins if the key it provides is correct for some (any) identity. In both cases, of course, the adversary can only win if the identity for which it finds the key is not exposed.

KEY-DERIVATION FUNCTIONS. In designs of IB-FPE schemes we will of course want efficient key-derivation functions. But in analyses and for other conceptual purposes, it will be useful to consider

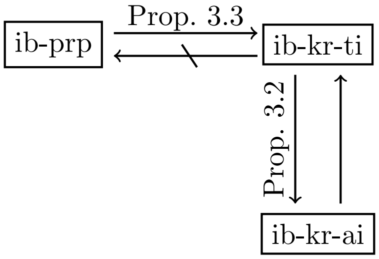

Figure 4: **Relations between notions of IB-PRP security. An arrow A → B is an implication: if an IB-PRP scheme meets A then it also meets B. A barred arrow $\mathrm{A} \not\rightarrow \mathrm{B}$ is a separation: there exists an IB-PRP scheme meeting A but not B. Unannotated lines represent trivial relations. The relations hold in both the adaptive and non-adaptive settings.**

key-derivation functions that are not efficient. In particular we define the *uniform key-derivation function* $\mathsf{U} = \mathbf{U}[\mathsf{F}, \mathsf{ID}]$, associated to F and a set ID of identities, to capture users having random, independent keys. Formally, let the master-key space $\mathsf{U.MKS} = \mathrm{Func}(\mathsf{ID}, \{0,1\}^{\mathsf{F.kl}})$ be the set of all functions from ID to $\{0,1\}^{\mathsf{F.kl}}$, so that a master key $K : \mathsf{ID} \to \{0,1\}^{\mathsf{F.kl}}$ is a function taking an identity and returning the key $K(I)$. Then the function $\mathsf{U} : \mathsf{U.MKS} \times \mathsf{ID} \to \{0,1\}^{\mathsf{F.kl}}$ is defined by $\mathsf{U}(K, I) = K(I)$. Picking K at random means the keys of different identities are random and independent.

<u>Non-adaptive security.</u> Let $\mathcal{A}$ be either an ib-prp or an ib-kr adversary. We say that it is *non-adaptive* if there is no identity I for which $\mathcal{A}$ makes both a $\textsc{Exp}(I)$ query and a non-$\textsc{Exp}(I)$ —that is, $\textsc{Ch}(I)$, $\textsc{Enc}(I, \cdot, \cdot)$ or $\textsc{Dec}(I, \cdot, \cdot)$— query. Thus, the adversary must make its decision to expose the key of an identity I up front, without prior queries to $\textsc{Enc}(I, \cdot, \cdot)$ or $\textsc{Dec}(I, \cdot, \cdot)$. (The definition also excludes post $\textsc{Exp}(I)$ queries $\textsc{Enc}(I, \cdot, \cdot)$, $\textsc{Dec}(I, \cdot, \cdot)$ and $\textsc{Ch}(I)$, but these are redundant anyway.) The security we prove for our constructions of IB-FPE schemes is restricted to non-adaptive adversaries as adaptivity allows selective-opening attacks (SOAs) [19, 7]. We elaborate on this below.

<u>Discussion.</u> In the definition of security for identity-based encryption (IBE) [13], the adversary can pick its (single) challenge not just while querying an exposure oracle (and, in the CCA case, a decryption oracle), but as a function of encryptions under identities of its choice. The latter is captured, trivially, by giving the adversary the master public key up front. Our setting is symmetric, so there is no master public key. As per the paradigm of [4], we accordingly give the adversary an encryption oracle. (To capture CCA, we also give it a decryption oracle. We continue of course to give the exposure oracle.) We allow multiple challenge identities, not just one. Starting encryption (and decryption) for an identity as honest and switching to challenge mode via $\textsc{Ch}$ captures an adaptive (encryption-dependent) choice of challenge identities to mirror IBE security. However, the presence of multiple challenges means that this effectively allows a SOA. SOAs are notoriously subtle, and security against them is known (at least for other primitives) to be hard to achieve [19, 7, 5, 23]. Correspondingly (and unsurprisingly) we find that we are unable to show our schemes meet our ib-prp definition for adaptive adversaries. We prove it, instead, for non-adaptive adversaries. These adversaries are still very powerful. (It is unclear that adaptivity is realistic or possible in practice.) We leave adaptive security as an open question.

```
Game $\mathbf{G}^{\mathsf{fp}}_{\mathsf{F},d}(J, J')$
$P \leftarrow\!\!\$\ \mathsf{F.IP}$ ; $T \leftarrow\!\!\$\ \mathsf{F.TS}$
For $i = 1, \ldots, d$ do $X_i \leftarrow\!\!\$\ \mathsf{F.Dom} \setminus \{X_1, \ldots, X_{i-1}\}$
$V \leftarrow (\mathsf{F.E}^P(J, T, X_1), \ldots, \mathsf{F.E}^P(J, T, X_d))$
$V' \leftarrow (\mathsf{F.E}^P(J', T, X_1), \ldots, \mathsf{F.E}^P(J', T, X_d))$
Return $(V = V')$
```

Figure 5: **Game to define the false positive advantage of F on d random messages, for subkeys J and J'.**

The multi-user (mu) setting [3, 2] considers many users, having keys that are uniformly and independently distributed. Mu security of an FPE scheme F can be viewed as a special case of our setting, as follows. Let n be the number of users, and let $\mathsf{ID} = \{1, \ldots, n\}$. Let $\mathsf{U} = \mathbf{U}[\mathsf{F}, \mathsf{ID}]$ be the uniform key-derivation function for F over this set of identities, and consider the IB-FPE scheme (F, U). Let us call an ib-prp adversary $\mathcal{A}$ a mu adversary if it begins by querying all n identities to its CH oracle, and makes no EXP queries. Then mu security of F is exactly ib-prp security of (F, U) relative to mu adversaries. In this way, certain results about IB-FPE will automatically imply results on the mu security of the base FPE scheme. Also, this lends a different perspective on IB-FPE, viewing it as a generalization of mu security in which keys of different users are not necessarily random and independent, key exposures are permitted and identities can be adaptively and optionally made challenge ones. We thank Stefano Tessaro for pointing out this connection and viewpoint to us.

A tweakable blockcipher [28] is the special case of an FPE scheme F in which $\mathsf{F.Dom} = \{0, 1\}^{\mathsf{F.bl}}$ for some $\mathsf{F.bl}$. Mu security for tweakable blockciphers was considered in [40, 26]. The work of LLMM [26], which is concurrent to, and independent of, ours, goes further to allow a key-derivation function so that they consider what in our language is effectively an identity-based tweakable blockcipher. Their definition of security, however, does not allow exposures and does not allow challenge identities to be adaptively determined. It is the special case of our ib-prp in which we restrict attention to what, above, we called mu adversaries.

FALSE POSITIVE RATE. Fix an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$. In some settings, we have an F-key J and an identity I and want to test whether $J = \mathsf{KDF}(K, I)$. We don't have K, or the task is of course easy, but we do have access to an oracle $\mathsf{F.E}^P(\mathsf{KDF}(K, I), \cdot, \cdot)$. The strategy is pick some tweak T and inputs $X_1, \ldots, X_d$, and declare J correct if $\mathsf{F.E}^P(\mathsf{KDF}(K, I), T, X_i) = \mathsf{F.E}^P(J, T, X_i)$ for all $i \in \{1, \ldots, d\}$. This test is not always correct. There may be false positives, meaning it might accept even if $J \neq \mathsf{KDF}(K, I)$. Here we give definitions to quantify this. Consider game $\mathbf{G}^{\mathsf{fp}}_{\mathsf{F},d}(J, J')$ defined in Fig. 5 associated to F, keys $J, J' \in \{0, 1\}^{\mathsf{F.kl}}$ and integer $d \leq |\mathsf{F.Dom}|$. Then define the *false positive advantage*

$$\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} = \max_{J \neq J'} \Pr[\mathbf{G}^{\mathsf{fp}}_{\mathsf{F},d}(J, J')]$$

as the maximum, over all distinct keys $J, J' \in \{0, 1\}^{\mathsf{F.kl}}$, of the probability that the game returns true.

We now compute this advantage for the case that F is ideal. Let $N = |\mathsf{F.Dom}|$ be the size of the domain. If $J \neq J'$ then $\mathsf{F.E}^P(J, T, \cdot)$ and $\mathsf{F.E}^P(J', T, \cdot)$ are independent random permutations, and hence

$$\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} = \frac{1}{N(N-1)\cdots(N-d+1)} \ . \tag{4}$$

The choice of d required to make the bound of Eq. (4) negligible is usually quite small. For example if $N = 2^{32}$ then setting $d = 9$ will be enough, by Eq. (4), to ensure a false positive advantage of only $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \le 2^{-256}$.

When F is not ideal, the false positive advantage depends on the structure of F. It is easy to give artificial examples of F for which $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ remains high even for large d, for example by having two distinct keys J, J' that induce the same encryption function on all tweaks, meaning $\mathsf{F.E}(J, T, X) = \mathsf{F.E}(J', T, X)$ for all T, X, in which case $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} = 1$ for all d. Real and natural designs of FPE schemes, however, are not expected to have such anomalies, and so it is customary to assume that the false positive advantage is about the same as that of an ideal FPE with the same domain, meaning approximated by Eq. (4). We will do this in our estimates.

Equivalence of ib-kr notions. It is clear that ib-kr-ai tightly implies ib-kr-ti; we now prove the converse. Given an ib-kr-ai adversary $\mathcal{A}_{\rm ai}$, one can construct an ib-kr-ai adversary $\mathcal{A}_{\rm ti}$ by running the former to get a candidate (I, J), and then testing J for all identities in the challenge set ChI to find a matching identity. We will use $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ defined above to account for the probability of false positive.

Proposition 3.2 *Let* $(\mathsf{F}, \mathsf{KDF})$ *be an IB-FPE scheme. Suppose that we are given an ib-kr-ai adversary* $\mathcal{A}_{\rm ai}$ *of* q Ch *queries. For a parameter* $d \in \mathbb{N}$*, we can construct an ib-kr-ti adversary* $\mathcal{A}_{\rm ti}$ *of about the same running time plus* qd *calls to* $\mathsf{F.E}$ *such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ai}) \le \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ti}) + q \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \ .$$

Adversary $\mathcal{A}_{\rm ti}$ *uses the same number of queries as* $\mathcal{A}_{\rm ai}$*, plus* qd *additional* Enc *queries. Finally, if* $\mathcal{A}_{\rm ai}$ *is non-adaptive, so is* $\mathcal{A}_{\rm ti}$*.*

Proof of Proposition 3.2: Adversary $\mathcal{A}_{\rm ti}$ runs $\mathcal{A}_{\rm ai}$ and uses its oracles to answer the queries of $\mathcal{A}_{\rm ai}$. When the latter outputs a candidate (I', J), the former samples messages $X_1, \ldots, X_d$ uniformly without replacement from $\mathsf{F.Dom}$, and picks a tweak T at random. Then for every identity I in the challenge set ChI, it compares the answers of $\mathsf{F.E}(J, T, X_i)$ and $\textsc{Enc}(I, T, X_i)$, for all $i \in \{1, \ldots, d\}$, and returns (I, J) if those answers are consistent. Let Hit be the event that J is the subkey of some identity in the challenge set ChI. Let Bad be the event that there is an identity $I \in$ ChI such that J is not the subkey of I, but the testing for I returns consistent answers. Then

$$\begin{aligned} \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ai}) &= \Pr[\mathsf{Hit}], \text{ and} \\ \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ti}) &\ge \Pr[\mathsf{Hit} \wedge \overline{\mathsf{Bad}}] \ge \Pr[\mathsf{Hit}] - \Pr[\overline{\mathsf{Bad}}] \ . \end{aligned}$$

On the other hand, let Bad_i be the event that among the q involved identities, the subkey J' of the ith identity is not J, but the encryption of the messages $X_1, \ldots, X_d$ for tweak T under J' are consistent with those under the subkey J. Then

$$\Pr[\mathsf{Bad}] \le \Pr\Big[\bigcup_{i=1}^{q} \mathsf{Bad}_i\Big] \le \sum_{i=1}^{q} \Pr[\mathsf{Bad}_i] \le q \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \ .$$

Summing up, we have

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ai}) \le \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\rm ti}) + q \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$$

```
Adversary $\mathcal{A}_{\mathrm{prp}}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
$(J,I) \leftarrow_{\$} \mathcal{A}_{\mathrm{kr}}^{\textsc{EncSim},\textsc{DecSim},\textsc{Exp},\textsc{Ch},P}$; $T_I \leftarrow \arg\min_T |Q(I,T)|$
For $i = 1, \ldots, n$ do
  $X \leftarrow_{\$} \mathsf{F.Dom} \setminus Q(I,T_I)$ ; $Q(I,T_I) \leftarrow Q(I,T_I) \cup \{X\}$
  If $\textsc{Enc}(I,T_I,X) \neq \mathsf{F.E}(J,T_I,X)$ then return 0
Return 1

Subroutine $\textsc{EncSim}(I,T,X)$
$Y \leftarrow \textsc{Enc}(I,T,Y); Q(I,T) \leftarrow Q(I,T) \cup \{X\}$
Return $Y$

Subroutine $\textsc{DecSim}(I,T,Y)$
$X \leftarrow \textsc{Dec}(I,T,Y)$; $Q(I,T) \leftarrow Q(I,T) \cup \{X\}$
Return $X$
```

Figure 6: **Adversary $\mathcal{A}_{\mathrm{prp}}$ for Proposition 3.3.**

as claimed. ▌

ib-prp implies ib-kr. One would expect that prpa security of an IB-FPE scheme implies its ib-kr-ti (and thus ib-kr-ai as per Proposition 3.2) security. A basic template for showing that indistinguishability style security implies key-recovery security is given in [35]. The kr adversary is executed to obtain a candidate key J'. To determine its challenge bit, the executing adversary now tests J' by seeing if encryption under it, on some "un-used" input, equals the output of the encryption oracle on the same input, where "un-used" means not already queried to the encryption oracle in the simulation of the kr-adversary. In the $b = 1$ case, the test will succeed. In the $b = 0$ case, the output of the encryption oracle is random and will equal the encryption under J' with a probability inverse in the size of the domain. Since the latter is usually large, the probability p of success in this case will be small.

Here we follow the basic template above. Given an ib-kr-ti adversary $\mathcal{A}_{\mathrm{kr}}$, without loss of generality, we can assume that the adversary makes only a single $\textsc{Ch}$ query, and it is the very last query before the adversary outputs its guess (I,J). We say that $\mathcal{A}_{\mathrm{kr}}$ leaves at least v unused points if there is a tweak T such that $\mathcal{A}_{\mathrm{kr}}$ makes at most $(|\mathsf{F.Dom}| - v)$ $\textsc{Enc}(I,T,\cdot)$ or $\textsc{Dec}(I,T,\cdot)$ queries.

Proposition 3.3 *Let $(\mathsf{F},\mathsf{KDF})$ be an IB-FPE scheme and let $d = |\mathsf{F.Dom}|$ be the size of the domain of F. Suppose we are given an ib-kr adversary $\mathcal{A}_{\mathrm{kr}}$ leaving at least v unused points. Let n be an integer parameter satisfying $1 \leq n \leq v$. Then we build an adversary $\mathcal{A}_{\mathrm{prp}}$ (shown in Fig. 6) such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\mathrm{kr}}) \leq \mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\mathrm{prp}}) + p \,,$$

where

$$p = 2^{\mathsf{F.kl}} \cdot \frac{(v-n)!}{v!} \,.$$

Adversary $\mathcal{A}_{\mathrm{prp}}$ makes the same number of queries as $\mathcal{A}_{\mathrm{kr}}$, plus n additional $\textsc{Enc}$ *queries. The running time of $\mathcal{A}_{\mathrm{prp}}$ is that of $\mathcal{A}_{\mathrm{kr}}$ plus the time for n executions of* $\mathsf{F.E}$. *Finally, if $\mathcal{A}_{\mathrm{kr}}$ is non-adaptive, so is $\mathcal{A}_{\mathrm{prp}}$.*

If d is large then p can be easily made small. The difficult case is when d is small. To illustrate the quality of our bounds in this case, let us consider an example, namely $d = 10^4$, corresponding to the encryption of 4 decimal digits of a credit-card number. With DFF, the key length will be $\mathsf{F.kl} = 256$. Setting $v = d/2 = 5000$ and $n = 30$ we have

$$p = 2^{256} \cdot \frac{(5000 - n)!}{5000!} \leq 2^{256} \cdot 2^{-368} \ ,$$

which is tiny.

Proof of Proposition 3.3: Adversary $\mathcal{A}_{\text{prp}}$, shown in Fig. 6, runs $\mathcal{A}_{\text{kr}}$. It answers all the latter's oracles via its own oracles of the same names, but also records related quantities. The set $Q(I, T)$ holds all X such that either an $\text{Enc}(I, T, X)$ query was made or a $\text{Dec}(I, T, \cdot)$ query returned X. As per our initialization conventions, sets are assumed initialized to empty and integers to 0. Once $\mathcal{A}_{\text{kr}}$ has terminated and returned (J, I), adversary $\mathcal{A}_{\text{prp}}$ tests whether J is the right key for I. It first finds a tweak T_I, called a minimal tweak for I, that minimizes the number of $\text{Enc}(I, T, \cdot)$ and $\text{Dec}(I, T, \cdot)$ queries, thereby allowing the maximum number of possible tests. $\mathcal{A}_{\text{prp}}$ obtains further example encryptions under I and its minimal tweak. This ensures the presence of at least v unused points for tweak T_I. The examples are then compared to the values predicted by J, and b' ends up being 1 if all examples match. Adversary $\mathcal{A}_{\text{prp}}$ returns b'.

For the analysis, let b denote the challenge bit in the execution of $\mathbf{G}^{\text{ib-prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\text{prp}})$. We claim that

$$\Pr[\, b' = 1 \,|\, b = 1 \,] \geq \mathbf{Adv}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\text{kr}}) \tag{5}$$

$$\Pr[\, b' = 1 \,|\, b = 0 \,] \leq p \ . \tag{6}$$

Subtracting we get

$$\begin{aligned} \mathbf{Adv}^{\text{ib-prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\text{prp}}) &= \Pr[\, b' = 1 \,|\, b = 1 \,] - \Pr[\, b' = 1 \,|\, b = 0 \,] \\ &\geq \mathbf{Adv}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\text{kr}}) - p \ , \end{aligned}$$

which establishes the theorem. We now justify the two numbered equations above.

First suppose $b = 1$. The replies that $\mathcal{A}_{\text{kr}}$ gets to its oracle queries match those in game $\mathbf{G}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}_{\text{kr}})$. If $\mathcal{A}_{\text{kr}}$ succeeds in the latter, then $J = \mathsf{KDF}(K, I)$ so b' will be 1 and $\mathcal{A}_{\text{prp}}$ will return 1. This justifies Eq. (5).

Now suppose $b = 0$. Recall that there are n unused points for tweak T_I. If we make Enc queries on those points, the answers are random and distinct, subject to being consistent with prior queries and answers. For any fixed key J, the probability that the tests succeed is at most

$$\prod_{i=0}^{n} \frac{1}{v - i} = \frac{(v - n)!}{v!} \ .$$

Since there are $2^{\mathsf{F.kl}}$ keys, the union bound yields Eq. (6). ∎

<u>Ib-kr doesn't imply ib-prp.</u> Conversely, we claim that ib-kr-ti (and thus ib-kr-ai due to Proposition 3.2) does not imply ib-prp. (This is the hatched arrow in Fig. 4.) This can be shown by counter-example. Thus, consider the FPE scheme F defined by $\mathsf{F.E}(J, T, X) = X$ for all J, T, X. Let ID be some non-empty set of identities and let $\mathsf{KDF} = \mathbf{U}[\mathsf{F}, \mathsf{ID}]$ be the associated uniform key-derivation function as defined above. IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$ is certainly not ib-prp secure. However an adversary $\mathcal{A}$ has $\mathbf{G}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \leq 2^{-\mathsf{F.kl}}$, making it ib-kr-ti secure if the key length of F is large.

4 Attacks

In this section we give generic non-adaptive attacks on *any* IB-FPE $(\mathsf{F}, \mathsf{KDF})$ that show inherent quantitative limits to the security that is achievable. Our attacks recover derived keys (meaning, have good advantage under our ib-kr-ai metric), not just violate ib-prp security. An important implication of these attacks is that for k bits of ib-kr-ai security, the key-length of F (which is the length of derived keys) must be at least $2k$-bits regardless of the length of the master key. The reason, roughly, is that collisions between derived keys can be exploited to violate security. These attacks are important to show that our constructions of IB-FPE schemes in later sections are optimal in security given the key lengths.

<u>Overview and diversity.</u> Let $(\mathsf{F}, \mathsf{KDF})$ be an IB-FPE scheme. For integer parameters $q, p \geq 1$, we will show that there is an attack (adversary) $\mathcal{A}$ that succeeds in key recovery with advantage $\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \geq \Omega(pq) \cdot 2^{-\mathsf{F.kl}}$. The adversary makes $O(q)$ Enc and Ch queries and has offline computation effort about the cost of $O(p)$ encryptions under $\mathsf{F.E}$. In particular, to allow p, q to reach $O(2^k)$, one must have $\mathsf{F.kl} \geq 2k$.

We define the *diversity* $\mathrm{KDiv}_{\mathsf{KDF}}(q)$ of KDF relative to q as the expected size of the set $\{\mathsf{KDF}^P(K, I_1), \ldots, \mathsf{KDF}^P(K, I_q)\}$, where the expectation is over $P \leftarrow\!\!\$\ \mathsf{KDF.IP}$, $K \leftarrow\!\!\$\ \mathsf{KDF.MKS}$, and $I_1, \ldots, I_q$ sampled uniformly without replacement from $\mathsf{KDF.IS}$ (that is, sampled uniformly and independently subject to being distinct). High diversity means that keys of different identities are largely distinct, while low diversity means keys of distinct identities frequently collide. We will give two, separate attacks. The first, called the matching attack (MA) works when the diversity is low. Specifically, it has a high (constant) ib-kr-ai advantage when $\mathrm{KDiv}_{\mathsf{KDF}}(q) \leq q/4$. The second, called the exhaustive-search attack (ESA) works when the diversity is high. Specifically, it has ib-kr-ai advantage around $\Omega(pq) \cdot 2^{-\mathsf{F.kl}}$ when $\mathrm{KDiv}_{\mathsf{KDF}}(q) > q/4$. All cases for the diversity being covered, one or the other attack always applies to get ib-kr-ai advantage of the claimed form. The analyses of the attacks are made more difficult by the fact that F and KDF share the same instance of the ideal primitive.

<u>The matching attack.</u> Let $(\mathsf{F}, \mathsf{KDF})$ and q be given. We associate to them the *matching adversary* $\mathbf{MA}_q$ described in Fig. 7. In this attack, the adversary first samples without replacement q identities $I_1, \ldots, I_q$, and picks a random $\ell \leftarrow\!\!\$\ \{1, \ldots, q\}$. The goal of the adversary is to recover the key of some identity I_i, for $i \in \{1, \ldots, q\} \backslash \{\ell\}$. To achieve this, it queries $\textsc{Exp}(I_\ell)$ to get the key J_ℓ corresponding to I_ℓ, and outputs J_ℓ as its guess. The intuition is that if the set $\{J_1, \ldots, J_q\}$ of keys for all identities involved is small (which happens on the average if the diversity is low) then J_ℓ is likely to equal J_i for some $i \neq \ell$, and the adversary wins. The cost of the attack is q queries to Enc and one query to Exp. The following theorem gives a precise lower bound on adversary advantage.

Theorem 4.1 *Let* $(\mathsf{F}, \mathsf{KDF})$ *be an IB-FPE scheme. Then for any* $q \in \mathbb{N}$ *we have*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{MA}_q) \geq \frac{1}{2} - \frac{\mathrm{KDiv}_{\mathsf{KDF}}(q)}{q} \ .$$

In particular if $\mathrm{KDiv}_{\mathsf{KDF}}(q) \leq q/4$ then $\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{MA}_q) \geq 1/2 - 1/4 = 1/4$, meaning the ib-kr-ai advantage is very high.

Proof of Theorem 4.1: Let K be the master key of KDF and let $P \leftarrow\!\!\$\ \mathsf{KDF.IP}$. From Markov's inequality,

$$\Pr\big[\big|\{\mathsf{KDF}^P(K, I_1), \ldots, \mathsf{KDF}^P(K, I_q)\}\big| \geq q/2\big] \leq \frac{\mathrm{KDiv}_{\mathsf{KDF}}(q)}{q/2} \ .$$

```
Adversary $\mathbf{MA}_q^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
$S \leftarrow \emptyset$ ; $T \leftarrow\!\!\$\ \mathsf{F.TS}$
$X \leftarrow\!\!\$\ \mathsf{F.Dom}$ ; $\ell \leftarrow\!\!\$\ \{1,\ldots,q\}$
For $i \leftarrow 1$ to $q$ do
  $I_i \leftarrow\!\!\$\ \mathsf{KDF.IS}\backslash S$ ; $S \leftarrow S \cup \{I_i\}$
For $i \in \{1,\ldots,q\}\backslash\{\ell\}$ do
  $C \leftarrow \textsc{Ch}(I_i, T, X)$
Pick arbitrary $I \in \{I_1,\ldots,I_q\}\backslash\{I_\ell\}$
Return $(I, \textsc{Exp}(I_\ell))$
```

```
Adversary $\mathbf{ESA}_{q,p,d}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
$S_1, S_2 \leftarrow \emptyset$ ; $T \leftarrow\!\!\$\ \mathsf{F.TS}$
For $\ell \leftarrow 1$ to $d$ do
  $X_\ell \leftarrow\!\!\$\ \mathsf{F.Dom}\backslash S_1$ ; $S_1 \leftarrow S_1 \cup \{X_\ell\}$
For $i \leftarrow 1$ to $q$ do
  $I_i \leftarrow\!\!\$\ \mathsf{KDF.IS}\backslash S_2$ ; $S_2 \leftarrow S_2 \cup \{I_i\}$
  For $\ell \leftarrow 1$ to $d$ do $V_\ell \leftarrow \textsc{Enc}(I_i, T, X_\ell)$
  $Z_i \leftarrow (V_1,\ldots,V_d)$
For $j \leftarrow 1$ to $p$ do
  $J_j \leftarrow\!\!\$\ \{0,1\}^{\mathsf{F.kl}}$
  For $\ell \leftarrow 1$ do $d$ do $U_\ell \leftarrow \mathsf{F.E}^P(J_j, T, X_\ell)$
  $Z \leftarrow (U_1,\ldots,U_\ell)$ ; $i \leftarrow \mathsf{Find}(Z, Z_1,\ldots,Z_q)$
  If $i > 0$ then $\textsc{Ch}(I_i)$; Return $(I_i, J_j)$
```

Figure 7: **Top:** The matching attack. **Bottom:** The exhaustive search attack.

Let $S = \{\mathsf{KDF}^P(K, I_1),\ldots,\mathsf{KDF}^P(K, I_q)\}$ and suppose $|S| \leq q/2$ which occurs with probability at least $1 - 2 \cdot \mathrm{KDiv}_{\mathsf{KDF}}(q)/q$. We say that identity I_i is *bad* if there is some $j \in \{1,\ldots,q\} \setminus \{i\}$ such that I_i and I_j have the same derived key, meaning $\mathsf{KDF}^P(K, I_i) = \mathsf{KDF}^P(K, I_j)$. Note that if there are at most r bad identities then the set S must have size at least $q - r$. Since we assumed $|S| \leq q/2$, there are at least $q/2$ bad identities. Since we pick ℓ at random, the chance that I_ℓ is bad is at least $1/2$. Hence $\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{MA}_q) \geq 1/2 - \mathrm{KDiv}_{\mathsf{KDF}}(q)/q$ as claimed. ▌

<u>The exhaustive search attack.</u> Let $(\mathsf{F}, \mathsf{KDF})$ be given, as well as integer parameters p, q, d. We associate to them adversary $\mathbf{ESA}_{q,p,d}$ described in Fig. 7. Algorithm $\mathsf{Find}(Z, Z_1,\ldots,Z_q)$, used in the attack as a subroutine, returns an index i such that $Z = Z_i$ if $Z \in \{Z_1,\ldots,Z_q\}$, and 0 otherwise. The attack somewhat generalizes and extends the NIST/NSA attack on FF2 [21], and also resembles Biham's key-collision attack on DES [11]. Biham's attack can be viewed as a special case of ours, where the key-derivation function is the uniform one, the domain is large, and the parameters p and q are close to $2^{\mathsf{F.kl}/2}$. The main novelty is a rigorous analysis lower-bounding the ib-kr-ai advantage. The attack uses dq queries to $\textsc{Enc}$, q queries to $\textsc{Ch}$, and no $\textsc{Exp}$ queries. The running time is that of dp executions of $\mathsf{F.E}$ plus p executions of Find. With appropriate data structures, the latter should cost about $O(p \log q)$. The value of d will be a small constant that, in estimates above, we absorbed into the big-oh.

The idea is as follows. The adversary picks distinct identities $I_1,\ldots,I_q$. Let $J'_i = \mathsf{KDF}^P(K, I_i)$, where K is the master key chosen in the overlying key-recovery game $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{ESA}_{q,p,d})$. The adversary aims to find one of the target keys $J'_1,\ldots,J'_q$ via exhaustive search over the space of FPE keys. It picks at random p keys $J_1,\ldots,J_p$ from the key space $\{0,1\}^{\mathsf{F.kl}}$ of F. Now, for each i, j, it aims to test whether $J'_i = J_j$. If any such test succeeds, it can call $\textsc{Ch}(I_i)$, return (I_i, J_j) and win. If the tests are perfectly correct, then it wins with probability about $pm \cdot 2^{-\mathsf{F.kl}}$ where $m = |\{J'_1,\ldots,J'_q\}|$, and if the diversity is high, like $\geq q/4$, then this looks like the ib-kr-ai advantage we want. There are however several difficulties. One is that there is no reasonable way to do perfectly correct testing. We will handle this by using the false positive advantage $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ defined in Section 3. Another difficulty is the analysis. In particular, $\mathrm{KDiv}_{\mathsf{KDF}}(q)$ is an expectation taken over the choice of P, but the same P is used by the encryption algorithm in the tests, so we cannot use independence of the success and false-positive probabilities.

The following gives a lower bound on the ib-kr-ai advantage of the exhaustive search attack.

Theorem 4.2 *Let* $(\mathsf{F},\mathsf{KDF})$ *be an IB-FPE scheme. Then for any* $p,q,d \in \mathbb{N}$ *such that* $pq \le 2^{\mathsf{F}.\mathsf{kl}}$ *we have*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{ESA}_{q,p,d}) \ge \frac{p \cdot \mathrm{KDiv}_{\mathsf{KDF}}(q)}{2^{\mathsf{F}.\mathsf{kl}+1}} - pq \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \ .$$

We saw above that if $\mathrm{KDiv}_{\mathsf{KDF}}(q) \le q/4$ then the matching attack already gives an attack with high (constant) ib-kr-ai advantage. The exhaustive search attack is effective in the complementary case where $\mathrm{KDiv}_{\mathsf{KDF}}(q) > q/4$. In this case, assuming $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ is negligible, Theorem 4.2 says the attack has ib-kr-ai advantage about $pq/2^{\mathsf{F}.\mathsf{kl}+3}$. In particular $p = q \approx 2^{\mathsf{F}.\mathsf{kl}/2}$ yields constant advantage, showing that k bits of security requires $\mathsf{F}.\mathsf{kl} \ge 2k$.

Proof of Theorem 4.2: Let T and $X_1, \ldots, X_d$ be the tweak and test messages that the adversary samples. Let K denote the master key chosen in the overlying key-recovery game $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{ESA}_{q,p,d})$ and let $J'_i = \mathsf{KDF}^P(K, I_i)$ for $1 \le i \le q$. Let Hit be the event that some guess J_j of the adversary is one of the target keys, meaning there are i,j such that $J_j = J'_i$. For $i \in \{1,\ldots,q\}$ and $j \in \{1,\ldots,p\}$ let $\mathsf{Bad}_{i,j}$ be the event that $J_j \ne J'_i$ and $(\mathsf{F}.\mathsf{E}^P(J_j,T,X_1),\ldots,\mathsf{F}.\mathsf{E}^P(J_j,T,X_d)) = (\textsc{Enc}(I_i,T,X_1), \ldots, \textsc{Enc}(I_i,T,X_d))$. Let Bad be the event $\exists\, i,j : \mathsf{Bad}_{i,j}$. If $\mathsf{Hit} \wedge \overline{\mathsf{Bad}}$ happens then one of the adversary's guesses is one of the target keys, and there are no false positive during the testing. So

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{ESA}_{q,p,d}) \ge \Pr[\mathsf{Hit} \wedge \overline{\mathsf{Bad}}] \ge \Pr[\mathsf{Hit}] - \Pr[\mathsf{Bad}] \ .$$

Separating these probabilities allows us to analyze them independently even though the P they use is the same. First we lower bound $\Pr[\mathsf{Hit}]$. Let Y be the random variable taking value the size of the set $\{J'_1, \ldots, J'_q\}$. Then

$$\Pr[\mathsf{Hit}] = \mathbf{E}\left[1 - \left(1 - \frac{Y}{2^{\mathsf{F}.\mathsf{kl}}}\right)^p\right] \ .$$

Let $a = Y/2^{\mathsf{F}.\mathsf{kl}} \ge 0$. We assumed $pq \le 2^{\mathsf{F}.\mathsf{kl}}$, and clearly $Y \le q$, so $ap \le 1$. This means the conditions of Lemma 2.1 for a and p are met. We now apply the lemma to get

$$\mathbf{E}\left[1 - \left(1 - \frac{Y}{2^{\mathsf{F}.\mathsf{kl}}}\right)^p\right] \ge \mathbf{E}\left[\frac{pY}{2 \cdot 2^{\mathsf{F}.\mathsf{kl}}}\right] = \frac{p \cdot \mathbf{E}[Y]}{2^{\mathsf{F}.\mathsf{kl}+1}} = \frac{p \cdot \mathrm{KDiv}_{\mathsf{KDF}}(q)}{2^{\mathsf{F}.\mathsf{kl}+1}} \ .$$

Next we upper bound $\Pr[\mathsf{Bad}]$. For any $i \in \{1,\ldots,q\}$ and $j \in \{1,\ldots,p\}$, if $J_j \ne J'_i$ then the probability that $(\mathsf{F}.\mathsf{E}^P(J_j,T,X_1),\ldots,\mathsf{F}.\mathsf{E}^P(J_j,T,X_d)) = (\textsc{Enc}(I_i,T,X_1),\ldots,\textsc{Enc}(I_i,T,X_d))$ is at most $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$, and hence $\Pr[\mathsf{Bad}_{i,j}] \le \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$. By the union bound we have $\Pr[\mathsf{Bad}] \le pq \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$. Putting all this together we have

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ai}}_{\mathsf{F},\mathsf{KDF}}(\mathbf{ESA}_{q,p,d}) \ge \frac{p \cdot \mathrm{KDiv}_{\mathsf{KDF}}(q)}{2^{\mathsf{F}.\mathsf{kl}+1}} - pq \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$$

as claimed. ▌

5 The PRF construction

We give a modular approach to build IB-FPE schemes. Given a prpa-secure FPE scheme F we set KDF to a PRF to get an ib-prp-secure IB-FPE scheme $(\mathsf{F},\mathsf{KDF})$. Then we instantiate KDF to get IB-FPE schemes with security matching our attacks.

<u>THE **PRF** CONSTRUCTION.</u> Theorem 5.1 below proves ib-prp security of $(\mathsf{F},\mathsf{KDF})$ assuming prpa security of F and prf security of KDF. The different resource metrics for $\mathcal{A}$ referred to below were defined in Section 3. The proof is a standard hybrid argument.

Theorem 5.1 *Let* $(\mathsf{F}, \mathsf{KDF})$ *be an IB-FPE scheme. Suppose we are given a non-adaptive ib-prp adversary* $\mathcal{A}$ *whose* $\textsc{Enc}, \textsc{Dec}, \textsc{Ch}$ *queries involve at most* u *different identities, with at most* q_1 *queries to* $\textsc{Enc}, \textsc{Dec}$ *per identity. Assume* $\mathcal{A}$ *makes* q_e *queries to* $\textsc{Exp}$*. Then we can construct a prpa adversary* $\overline{\mathcal{A}}$ *of* q_1 $\textsc{Enc}/\textsc{Dec}$ *queries, and a prf adversary* $\mathcal{B}$ *making* $u + q_e$ *queries to its* $\textsc{Fn}$ *oracle, such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \le u \cdot \mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) + 2 \cdot \mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{B}) \ . \tag{7}$$

The running time of $\overline{\mathcal{A}}$ *and* $\mathcal{B}$ *is about the same as that of* $\mathcal{A}$ *plus the time for* q_1 *executions of* $\mathsf{F.E}$*.*

Proof of Theorem 5.1: We construct adversary $\mathcal{B}$ as follows. It runs $\mathcal{A}$ and simulates game $\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A})$. However, instead of sampling a random $2k$-bit key for each identity I, it calls its PRF oracle on input I. Then

$$\begin{aligned}
\Pr[\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{A}) \Rightarrow \mathsf{true} \mid b = 1] &= \Pr[\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})], \text{ and} \\
\Pr[\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{A}) \Rightarrow \mathsf{false} \mid b = 0] &= \Pr[\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A})],
\end{aligned}$$

where b is the challenge bit in game $\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{A})$ and $\mathsf{U} = \mathsf{U}[\mathsf{F}, \mathsf{KDF.IS}]$ is the uniform key-derivation function. Subtracting, we obtain

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{B}) &= \Pr[\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})] - \Pr[\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A})] \\
&= \frac{1}{2}\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) - \frac{1}{2}\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A}) \ .
\end{aligned}$$

We now construct the prp adversary $\overline{\mathcal{A}}$ via a standard hybrid argument. The code of $\overline{\mathcal{A}}$ is shown in Fig. 8. It picks an index $g \leftarrow [1..u]$ and uses its $\textsc{Enc}/\textsc{Dec}$ oracles to respond to queries on the g-th identity. Consider games G_g in Fig. 8, for $0 \le g \le u$. Note that

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A}) = \Pr[G_u] - \Pr[G_0] \ .$$

On the other hand,

$$\begin{aligned}
\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) &= \frac{1}{u}\sum_{g=1}^{u} \Pr[G_g] - \Pr[G_{g-1}] \\
&= \frac{1}{u}(\Pr[G_{u+1}] - \Pr[G_0]) \\
&= \frac{1}{u}\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{U}}(\mathcal{A}) \ .
\end{aligned}$$

Summing up,

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \le u \cdot \mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) + 2\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{B}) \ ,$$

which concludes the proof. ▌

From Theorem 5.1, one can obtain an IB-FPE scheme in the standard model, by setting F to a standard-model FPE scheme such as the Sometimes-Recurse shuffle [30]. This answers in the affirmative the theoretical question of whether IB-FPE is achievable in the standard model.

<u>Tightness of bound.</u> Suppose $\mathsf{F.kl} = 2k$ and F has ideal behavior. Then we would expect $\mathbf{Adv}^{\mathsf{prp}}_{\mathsf{F}}(\overline{\mathcal{A}}) \approx q_1/2^{2k}$, corresponding to exhaustive key search being the best attack on prp security, and consequently from Proposition 3.2, $\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) \lesssim 2q_1/2^{2k}$. Similarly assuming KDF has

```
Adversary $\overline{\mathcal{A}}^{\textsc{Enc},\textsc{Dec},\textsc{Ch},P}$
$g \leftarrow\!\!\$\, [1..u]$; $c \leftarrow 0$
For $i \in [1..u]$ do $J_i \leftarrow\!\!\$\, \{0,1\}^{2k}$
$b' \leftarrow\!\!\$\, \mathcal{A}^{\textsc{EncSim},\textsc{DecSim},\textsc{ExpSim},\textsc{ChSim},P}$
Return $b'$

$\textsc{EncSim}(I,T,X)$
If $\mathrm{ET}[I,T,X] \neq \bot$ then return $\mathrm{ET}[I,T,X]$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I]}$
If $(\mathbf{v}[I] = g)$ then $Y \leftarrow \textsc{Enc}(T,X)$
Else
  If $I \in \mathrm{ChI}$ then
    If $(\mathbf{v}[I] < g)$ then $Y \leftarrow \mathsf{F.E}(J,T,X)$
    Else $Y \leftarrow\!\!\$\, \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[I,T,Y] = \bot \,\}$
  Else $Y \leftarrow \mathsf{F.E}(J,T,X)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$; Return $Y$

$\textsc{DecSim}(I,T,Y)$
If $\mathrm{DT}[I,T,Y] \neq \bot$ then return $\mathrm{DT}[I,T,Y]$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I]}$
If $(\mathbf{v}[I] = g)$ then $X \leftarrow \textsc{Dec}(T,Y)$
Else
  If $I \in \mathrm{ChI}$ then
    If $(\mathbf{v}[I] < g)$ then $X \leftarrow \mathsf{F.D}(J,T,Y)$
    Else $X \leftarrow\!\!\$\, \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[I,T,X] = \bot \,\}$
  Else $X \leftarrow \mathsf{F.D}(J,T,Y)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $X$

$\textsc{ExpSim}(I)$
$J \leftarrow\!\!\$\, \{0,1\}^{2k}$; Return $J$

$\textsc{ChSim}(I)$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
If $\mathbf{v}[I] = g$ then $\textsc{Ch}()$
```

```
Game $\mathrm{G}_g$ $(0 \le g \le u)$
For $i \in [1..u]$ do $J_i \leftarrow\!\!\$\, \{0,1\}^{2k}$
$c \leftarrow 0$; $b' \leftarrow\!\!\$\, \mathcal{A}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
Return $(b' = 1)$

$\textsc{Enc}(I,T,X)$
If $\mathrm{ET}[I,T,X] \neq \bot$ then return $\mathrm{ET}[I,T,X]$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I]}$
If $I \in \mathrm{ChI}$ then
  If $(\mathbf{v}[I] \le g)$ then $Y \leftarrow \mathsf{F.E}(J,T,X)$
  Else $Y \leftarrow\!\!\$\, \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[I,T,Y] = \bot \,\}$
Else $Y \leftarrow \mathsf{F.E}(J,T,X)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$; Return $Y$

$\textsc{Dec}(I,T,Y)$
If $\mathrm{DT}[I,T,Y] \neq \bot$ then return $\mathrm{DT}[I,T,Y]$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
If $I \in \mathrm{ChI}$ then
  If $(\mathbf{v}[I] \le g)$ then $X \leftarrow \mathsf{F.D}(J,T,Y)$
  Else $X \leftarrow\!\!\$\, \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[I,T,X] = \bot \,\}$
Else $X \leftarrow \mathsf{F.D}(J,T,Y)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $X$

$\textsc{Exp}(I)$
$J \leftarrow\!\!\$\, \{0,1\}^{2k}$; Return $J$

$\textsc{Ch}(I)$
If $(\mathbf{v}[I] = \bot)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
```

Figure 8: **The games and constructed adversary $\overline{\mathcal{A}}$ in the proof of Theorem 5.1.**

optimal prf security and $|\mathsf{KDF.MKS}| \ge 2^k$, we would expect $\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{B}) \approx (q_e + u)/2^k$. Then the bound of Eq. (9) becomes

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F,KDF}}(\mathcal{A}) \lesssim \frac{2uq_1}{2^{2k}} + \frac{2(q_e+u)}{2^k} \; . \tag{8}$$

This allows u, q_1, q_e to reach $O(2^k)$, which as per our attacks means the bound from Theorem 6.1 is essentially tight.

Instantiating KDF. Recall that we want to use *only* AES as our cryptographic primitive. Thus one needs to show how to instantiate KDF from a blockcipher $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ such that KDF achieves k-bits of prf security assuming that E has k bits of prp-cpa security. This is non-trivial, and as a stepping stone, we first aim to achieve a good PRF $F : \{0,1\}^k \times \{0,1\}^{k-1} \to \{0,1\}^k$.

```
$\underline{\mathsf{KDF}[E](K, I)}$
$J_0 \leftarrow E_K(I \,\|\, 00) \oplus E_K(I \,\|\, 01)$; $J_1 \leftarrow E_K(I \,\|\, 10) \oplus E_K(I \,\|\, 11)$
Return $J_0 \,\|\, J_1$
```

Figure 9: **Key-derivation function $\mathsf{KDF}[E]$, where $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ is a blockcipher.**

BKR [8] suggest that one can build F by way of

$$F_K(x) = E_K(x \,\|\, 0) \oplus E_K(x \,\|\, 1) \ .$$

The following result by DHT [15] confirms that F indeed has k-bit prf security.

Lemma 5.2 *[15] Let $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ be a blockcipher. Let $F : \{0,1\}^k \times \{0,1\}^{k-1} \to \{0,1\}^k$ be constructed by $E_K(x) = F_K(x \,\|\, 0) \oplus F_K(x \,\|\, 1)$. Then for any prf adversary A making $q \le 2^{k-5}$ queries to $\textsc{Fn}$, we can construct an adversary $\overline{\mathcal{A}}$ of about the same running time and $2q$ oracle queries such that*

$$\mathbf{Adv}^{\mathsf{prf}}_{F}(\mathcal{A}) \le \mathbf{Adv}^{\mathsf{prp\text{-}cpa}}_{E}(\overline{\mathcal{A}}) + \frac{1.5q + 3\sqrt{q}}{2^k} \ .$$

We then can construct a key-derivation function $\mathsf{KDF}[E] : \{0,1\}^k \times \{0,1\}^{k-2} \to \{0,1\}^{2k}$ by

$$\mathsf{KDF}[E](K, I) = F_K(I \,\|\, 0) \,\|\, F_K(I \,\|\, 1)$$

for any k-bit master key K and $(k-2)$-bit identity I. The key-derivation function $\mathsf{KDF}[E]$ can be expressed in terms of E as in Fig. 9. Proposition 5.3 below shows that $\mathsf{KDF}[E]$ also has k-bit prf security.

Proposition 5.3 *Let $E : \{0,1\}^k \times \{0,1\}^k \to \{0,1\}^k$ be a blockcipher. Let $\mathsf{KDF}[E] : \{0,1\}^k \times \{0,1\}^{k-2} \to \{0,1\}^{2k}$ be as in Fig. 9. Then for any adversary $\mathcal{A}$ that makes $q \le 2^{k-6}$ queries, we can construct another adversary $\overline{\mathcal{A}}$ of about the same running time and $4q$ oracle queries such that*

$$\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}[E]}(\mathcal{A}) \le \mathbf{Adv}^{\mathsf{prp\text{-}cpa}}_{E}(\overline{\mathcal{A}}) + \frac{3q + 5\sqrt{q}}{2^k} \ .$$

Proof of Proposition 5.3: Without loss of generality, suppose that $\mathcal{A}$ does not repeat a prior query. We first reduce the prf security of KDF to the prf security of F, by constructing an adversary $\mathcal{B}$ attacking F. Adversary $\mathcal{B}$ runs $\mathcal{A}$. When the latter queries $\textsc{Fn}(I)$, the former queries $I \,\|\, 0$ and $I \,\|\, 1$ to its oracle to get answer Z_0 and Z_1, and returns $Z_0 \,\|\, Z_1$ to $\mathcal{A}$. When $\mathcal{A}$ finally outputs a bit b', $\mathcal{B}$ outputs the same bit. Let a and b be the challenge bit in game $\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}}(\mathcal{A})$ and $\mathbf{G}^{\mathsf{prf}}_{F}(\mathcal{B})$ respectively. Then

$$\begin{aligned}
\Pr[\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}[E]}(\mathcal{A}) \mid a = 1] &= \Pr[\mathbf{G}^{\mathsf{prf}}_{F}(\mathcal{B}) \mid b = 1], \text{ and} \\
\Pr[\mathbf{G}^{\mathsf{prf}}_{\mathsf{KDF}[E]}(\mathcal{A}) \mid a = 0] &= \Pr[\mathbf{G}^{\mathsf{prf}}_{F}(\mathcal{B}) \mid b = 0] \ .
\end{aligned}$$

Adding the equations above side by side we have

$$\mathbf{Adv}^{\mathsf{prf}}_{\mathsf{KDF}[E]}(\mathcal{A}) \le \mathbf{Adv}^{\mathsf{prf}}_{F}(\mathcal{B}) \ .$$

Adversary $\mathcal{B}$ has about the same running time as $\mathcal{A}$, and makes $2q \leq 2^{k-5}$ oracle queries. Using Lemma 5.2, one can construct another adversary $\overline{\mathcal{A}}$ of about the same running time as $\mathcal{B}$, and $4q$ oracle queries such that

$$\begin{aligned}\mathbf{Adv}_F^{\mathsf{prf}}(\mathcal{B}) &\leq \mathbf{Adv}_E^{\mathsf{prp\text{-}cpa}}(\overline{\mathcal{A}}) + \frac{3q + 3\sqrt{2q}}{2^k} \\ &\leq \mathbf{Adv}_E^{\mathsf{prp\text{-}cpa}}(\overline{\mathcal{A}}) + \frac{3q + 5\sqrt{q}}{2^k} \enspace .\end{aligned}$$

Putting all this together we get the claimed result. ▌

Discussion. The KDF construction above uses 4 blockcipher calls. Alternatively, one might consider using Iwata's CENC method [24] that makes only 3 blockcipher calls. Specifically, let $G : \{0,1\}^k \times \{0,1\}^{k-2} \to \{0,1\}^k$ be constructed via

$$G_K(I) = (Z \oplus E(K, I \,\|\, 01)) \,\|\, (Z \oplus E(K, I \,\|\, 10)) \enspace ,$$

for any k-bit master key K and $(k-2)$-bit identity I, where $Z = E(K, I \,\|\, 00)$. IMV [25] claim that G has k-bit prf security, but their analysis is based on a combinatorial result by Patarin [32] whose proof is very involved with some unproven claims [15]. We therefore use the $\mathsf{KDF}[E]$ construction above, as it has a rigorous proof.

6 The Dbl construction

In Section 5, we followed the natural route to building IB-FPE in which the key-derivation function KDF is a PRF, and showed that one can instantiate KDF using four calls to an underlying blockcipher. In this section, we'll consider how to build a faster key-derivation function KDF for a class of FPE schemes F that we call *square*. It is in fact an abstraction of the key-derivation function of the proposed DFF standard [39]. The key-derivation function makes just two calls to the underlying blockcipher. Interestingly, it has poor (only birthday-bound) prf security, but we'll give a dedicated analysis to justify the (non-adaptive) ib-prp security of the IB-FPE scheme.

We use a *common ideal primitive* framework. All schemes use a common instance of a single ideal primitive—an ideal cipher $\mathbf{IC}(k,k)$ in which the key and block length are the same. In particular we allow the starting square FPE scheme F to use $P \in \mathbf{IC}(k,k)$, and then define KDF using the *same* P. This is because, for efficiency and implementation ease, we aim for all final constructions to be (only) AES-based.

The analysis is made challenging by two elements. First is to not only prove security, but with a good bound. Second is that the ideal primitive being in common means there can be queries made to it (directly, or indirectly via other oracles) by both the key-derivation function and the encryption and decryption functions, so we cannot use independence in a straightforward way.

Square FPE schemes. Let F be an FPE scheme. We say that it is *square* if there is an integer $k \geq 1$ such that $\mathsf{F.kl} = 2k$ and $\mathsf{F.IP} = \mathbf{IC}(k,k)$. That is, the ideal primitive associated to F is the ideal cipher with key and block length both the same value k, and moreover keys for the scheme are of length $2k$. DFF is underlain by such a square scheme [39]. (In contrast, FF2 was not.) The term "square" refers to the fact that the key space has size 2^{2k}, the square of the size 2^k of what, below, will be the master key space of the IB-FPE scheme, which is crucial for getting high security due to the attacks from Section 4.

The **Dbl** construction. Let F be a square FPE scheme with $\mathsf{F.kl} = 2k$. We first define embedding schemes, and then a key-derivation function KDF to turn F into an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$.

```
$\underline{\mathsf{KDF}^P(K, I)}$
$J_0 \leftarrow P(K, \mathsf{M}_0(I), +)$ ; $J_1 \leftarrow P(K, \mathsf{M}_1(I), +)$ ; $J \leftarrow J_0 \| J_1$
Return $J$
```

Figure 10: **Key-derivation function $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$ associated to embedding scheme M, where $P \in \mathbf{IC}(k, k)$.**

An *embedding scheme* M specifies a pair of functions $\mathsf{M}_0, \mathsf{M}_1 : \mathsf{M.IS} \to \{0,1\}^k$ satisfying two conditions: (1) Both M_0 and M_1 are injective and (2) the two maps have disjoint images, meaning $\mathsf{M}_0(I_1) \neq \mathsf{M}_1(I_2)$ for all $I_1, I_2 \in \mathsf{M.IS}$. We refer to $\mathsf{M}_0, \mathsf{M}_1$ as the embedding functions, and $\mathsf{M.IS}$ as the identity space, of M.

Now we define the key-derivation function $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$ to construct an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF})$. We let $\mathsf{KDF.IS} = \mathsf{M.IS}$, meaning the identity space is that of M. We let $\mathsf{KDF.mkl} = k$, so that a master key is a k-bit string. Then the key-derivation function $\mathsf{KDF}^P : \{0,1\}^k \times \mathsf{M.IS} \to \{0,1\}^{2k}$ is as specified in Fig. 10. The key for identity I is the result of applying the ideal cipher, keyed with the master key K, to $\mathsf{M}_0(I)$ and $\mathsf{M}_1(I)$, and concatenating these k-bit strings to get a $2k$-bit key. The key-derivation function, the encryption algorithm and the decryption algorithm all have access to $P \in \mathbf{IC}(k, k)$. We stress that, as discussed above for this common ideal primitive framework, the key derivation uses the *same* instance of the ideal cipher as encryption and decryption.

<u>Resistance to attacks.</u> Let F be a square FPE scheme with $\mathsf{F.kl} = 2k$. We consider how well the attacks of Section 4 do against F under $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$. First, we claim that our choice of KDF renders the matching attack entirely ineffective. Indeed, since $P(K, \cdot, +)$ is a permutation, the keys for distinct identities will be distinct. Thus for any $K \in \{0,1\}^k$ and any identities $I_1, \ldots, I_q \in \mathsf{M.IS}$, the set $\{\mathsf{KDF}^P(K, I_1), \ldots, \mathsf{KDF}^P(K, I_q)\}$ will have size exactly q. So its expected size, which is our diversity metric $\mathrm{KDiv}_{\mathsf{KDF}}(q)$ from Section 4, will equal q. Not only does Theorem 4.1 become vacuous, but, looking at the attack in Fig. 7, we see that it will have ib-kr-ai advantage zero, because the key returned by $\textsc{Exp}(I_\ell)$ will not equal any of the keys corresponding to the other identities. This shows a benefit of using a block cipher as the tool in key derivation for KDF. Had we used even a random oracle, the matching attack would have had at least some success.

The exhaustive search attack does have a non-trivial ib-kr-ai advantage. We noted above that $\mathrm{KDiv}_{\mathsf{KDF}}(q) = q$. Assuming the false positive advantage $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ is negligible, recall that Theorem 4.2 says that the ib-kr-ai advantage of the exhaustive search attack is about $p \cdot \mathrm{KDiv}_{\mathsf{KDF}}(q) \cdot 2^{-\mathsf{F.kl}-1} = pq \cdot 2^{-\mathsf{F.kl}-1} = pq \cdot 2^{-2k-1}$, where q is the number of adversary $\textsc{Enc}$ queries and p is roughly its running time. So the ib-kr-ai advantage stays below 1 as long as p and q each stay below 2^k. This means we have k-bits of security against this attack, and explains the choice of square schemes.

In summary, $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$ has been designed so that the attacks we gave in Section 4 are not threats to the security of F under KDF, in particular because $\mathsf{F.kl} = 2k$ while $\mathsf{KDF.mkl} = k$. However, this does *not* guarantee security, since there may well be other attacks. Moreover, KDF has only birthday-bound prf security, and thus using Theorem 5.1 gives us only $k/2$-bits of ib-prp security for $(\mathsf{F}, \mathsf{KDF})$. The main purpose of this section is to supply proof-based evidence of k-bit security.

<u>Goals and naive reduction.</u> The assumption we make is that the given square FPE scheme F satisfies prpa security. (This is equivalent to conventional prp security as per Proposition 3.1.)

Our goal is thus to reduce the ib-prp security of $(\mathsf{F}, \mathbf{Dbl}[k, \mathsf{M}])$ to the prpa security of F. As F is defined in the ideal-cipher model, this involves something somewhat novel, namely a reduction in the ideal cipher model. (Usually, in idealized models, one directly proves bounds on adversary advantage rather than giving reductions.) Given a non-adaptive ib-prp adversary $\mathcal{A}$ we aim to build another adversary $\overline{\mathcal{A}}$ and bound $\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$ as a function of $\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}})$ and the resources of $\mathcal{A}$, in particular the number u of users (identities) queried. $\overline{\mathcal{A}}$ will simulate $\mathcal{A}$'s P oracle.

The natural approach is a hybrid argument. The naive way of doing this, however, will incur a loss of $u^2/2^k$ in the advantage. This is undesirable since we want to show security up to $u \approx 2^k$, not $u \approx 2^{k/2}$. In more detail, the i-th hybrid game would let the keys of the first i identities be random, and the rest be specified via KDF as per Fig. 10 $(0 \leq i \leq u)$. Adversary $\overline{\mathcal{A}}$ would pick i at random to play the role of its single user, aiming to simulate the other identities for $\mathcal{A}$. Let J_i denote the key (underlying the single identity queried) in $\overline{\mathcal{A}}$'s game. The difficulty is that, for the simulation to be correct in the case that $\overline{\mathcal{A}}$'s challenge bit is 1, the $j = u-i+1$ keys $J_i, \ldots, J_u$ must be consistent with the structure imposed by KDF, meaning be formed by taking $2j$ distinct, random k bit strings and concatenating them in pairs. But J_i is random since $\overline{\mathcal{A}}$ is in the prpa game, and while $\overline{\mathcal{A}}$ can pick $J_{i+1}, \ldots, J_u$ to have the desired structure, this leaves a probability $\epsilon = O(u/2^k)$ that J_i will not have a consistent structure. Specifically, regardless of how $\overline{\mathcal{A}}$ picks distinct $J_{i+1}, \ldots, J_u$, the chance that one of those is J_i is $\epsilon = (u-i)/2^k = O(u/2^k)$. This means a loss of ϵ in each hybrid step, meaning, when $\overline{\mathcal{A}}$ picks i, its advantage is the difference in probabilities from the $(i+1)$-th and i-th hybrid games plus ϵ. When we sum over all hybrids (corresponding to the random choice of i), we get a $u\epsilon$ loss. What we want instead is a reduction with a loss that is $O(u/2^k)$ *globally*. This is what we will provide below, thereby showing security matching our attacks.

KEY USAGE METRIC. When invoked with a particular key J, the algorithms $\mathsf{F.E}, \mathsf{F.D}$ of the square FPE scheme will invoke their ideal cipher instance P with certain keys. Specifically there is a set $T(J) \subseteq \{0,1\}^k$ such that all P-queries of $\mathsf{F.E}$ and $\mathsf{F.D}$ only use keys in $T(J)$, regardless of the inputs to $\mathsf{F.E}, \mathsf{F.D}$ and responses to oracle queries. We let $\mathsf{F.nk}$ be the maximum, over all J, of the size of $T(J)$. This may sound complicated but it is really simple because typical constructions will evaluate the ideal cipher only on some fixed number of keys related to J. For example, for $\mathsf{F} = \mathsf{FF}_{\mathsf{dff}}$, we have $\mathsf{F.nk} = 1$. That is, there is only one blockcipher key used in the construction. We define this because our bounds will depend on it.

REDUCTION THEOREM. We now reduce the non-adaptive ib-prp security of our constructed IB-FPE scheme to the prpa security of the underlying FPE scheme. (The latter can be further reduced to its conventional prp security via Proposition 3.1.) The following theorem gives a good bound, where the global loss (the second term in the bound) is only $O(q/2^k)$ over and above the inevitable linear loss from the hybrid argument, where q is linear (not quadratic) in the number of queries that $\mathcal{A}$ makes to its different oracles. The quality of the bound is the same as that of Theorem 5.1, despite the low prf security of $\mathbf{Dbl}[k, \mathsf{M}]$.

Theorem 6.1 *Let* F *be a square FPE scheme with* $\mathsf{F.kl} = 2k$. *Let* $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$ *be a key-derivation function as per Fig. 10. Suppose we are given a non-adaptive adversary* $\mathcal{A}$ *whose* ENC, DEC, CH *queries involve at most* u *different identities, with at most* q_1 *queries to* ENC, DEC *per identity. Assume* $\mathcal{A}$ *makes* q_e *queries to* EXP *and* p *queries to* IP. *The proof constructs an adversary* $\overline{\mathcal{A}}$ *such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \quad \leq \quad u \cdot \mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) + \frac{8u + 8q_e + 2p + 2u \cdot \mathsf{F.nk} - 6}{2^k} \,. \tag{9}$$

Adversary $\overline{\mathcal{A}}$ *makes at most* q_1 *queries to* ENC, DEC *and* p *queries to* IP. *Its running time is about the same as that of* $\mathcal{A}$.

Starting above, we may use IP as the name of the game procedure that implements the ideal primitive instance. Where Fig. 3 gives $\mathcal{A}$ oracles $\textsc{Enc}, \textsc{Dec}, \textsc{Exp}, \textsc{Ch}, P$, we would now give it oracles $\textsc{Enc}, \textsc{Dec}, \textsc{Exp}, \textsc{Ch}, \mathrm{IP}$, with $\mathrm{IP}(x)$ defined to simply return $P(x)$ in the games of Fig. 3. The reason it helps to name the procedure is that in our proofs it will be programmed, and not always set to P. It will also be useful to define the key-derivation function $\overline{\mathsf{KDF}} : \mathrm{Perm}(\{0,1\}^k) \times \mathsf{M.IS} \to \{0,1\}^{2k}$ by

$$\overline{\mathsf{KDF}}(\overline{\pi}, I) = \overline{\pi}(\mathsf{M}_0(I)) \,\|\, \overline{\pi}(\mathsf{M}_1(I)) \tag{10}$$

for all $\overline{\pi} \in \mathrm{Perm}(\{0,1\}^k)$ and all $I \in \mathsf{M.IS}$. We prove Theorem 6.1 by invoking lemmas that will follow.

Proof of Theorem 6.1: Let $N = u + q_e$. Let $\overline{\mathsf{KDF}}$ be the key derivation function defined by Eq. (10). Using Lemma 6.2 and then Lemma 6.3 we have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) &\leq \mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\overline{\mathsf{KDF}}}(\mathcal{A}) + \frac{2p + 2u \cdot \mathsf{F.nk}}{2^k} \\ &\leq u \cdot \mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) + \frac{8N - 6 + 2p + 2u \cdot \mathsf{F.nk}}{2^k}\ ,\end{aligned}$$

where $\overline{\mathcal{A}}$ is the adversary given by Lemma 6.3. ▮

<u>Lemmas.</u> The first lemma allows a move to a setting where key derivation no longer uses the ideal primitive P that is used by F, but instead generates keys independently, although still with the same distribution as that of the prescribed key-derivation scheme. This lemma holds for both adaptive and non-adaptive adversaries $\mathcal{A}$.

Lemma 6.2 *Let* F *be a square FPE scheme with* $\mathsf{F.kl} = 2k$. *Let* $\mathsf{KDF} = \mathbf{Dbl}[k, \mathsf{M}]$ *be the key-derivation function of Fig. 10. Let* $\overline{\mathsf{KDF}}$ *be the key derivation function defined by Eq. (10). Let* $\mathcal{A}$ *be an adversary. Then*

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \leq \mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\overline{\mathsf{KDF}}}(\mathcal{A}) + \frac{2p + 2u \cdot \mathsf{F.nk}}{2^k} \tag{11}$$

where p *is the number of* IP *queries of* $\mathcal{A}$ *and* u *is the number of different identities involved in the* $\textsc{Enc}, \textsc{Dec}$ *queries of* $\mathcal{A}$.

Note that the reduction does not change the adversary. Our claim is that the ib-prp advantage of $\mathcal{A}$ with respect to the original key-derivation scheme is bounded by its ib-prp advantage with respect to the newly-defined key-distribution scheme plus an error term that is linear in the resources.

Proof of Lemma 6.2: Consider games G_0 and G_1 of Fig. 11. They optimistically imagine that key generation works as per $\overline{\mathsf{KDF}}$, picking a random permutation $\overline{\pi}$ and using it to specify the user keys. The notation $(L, W, s) \leftarrow x$ in the code for IP means this oracle parses its query x as a triple consisting of a key $L \in \{0,1\}^k$, an input $W \in \{0,1\}^k$, and a sign $s \in \{+, -\}$. If L equals the master key K, the bad flag is set to true, and game G_0, which includes the boxed code, corrects by setting $P(K, \cdot, +)$ to $\overline{\pi}$ and its inverse $P(K, \cdot, -)$ to $\overline{\pi}^{-1}$. Game G_1, however, does not include the boxed code. The result is that game G_0 is using KDF for key generation while game G_1 is using $\overline{\mathsf{KDF}}$. Thus

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \;=\; 2\Pr[\mathrm{G}_0] - 1 \tag{12}$$

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\overline{\mathsf{KDF}}}(\mathcal{A}) \;=\; 2\Pr[\mathrm{G}_1] - 1\ . \tag{13}$$

```
Game $\boxed{\mathrm{G}_0}$, $\mathrm{G}_1$
$b \leftarrow_{\$} \{0,1\}$ ; $K \leftarrow_{\$} \{0,1\}^k$ ; $\mathrm{XI} \leftarrow \emptyset$ ; $\mathrm{ChI} \leftarrow \emptyset$
$P \leftarrow_{\$} \mathbf{IC}(k,k)$ ; $\overline{\pi} \leftarrow_{\$} \mathrm{Perm}(\{0,1\}^k)$
For every $I \in \mathsf{M.IS}$ do
    $J_{I,0} \leftarrow \overline{\pi}(\mathsf{M}_0(I))$ ; $J_{I,1} \leftarrow \overline{\pi}(\mathsf{M}_1(I))$
    $J_I \leftarrow J_{I,0} \| J_{I,1}$
$b' \leftarrow_{\$} \mathcal{A}^{\mathrm{Enc},\mathrm{Dec},\mathrm{Exp},\mathrm{Ch},\mathrm{IP}}$
Return $(b = b')$

$\mathrm{Enc}(I,T,X)$
If $\mathrm{ET}[I,T,X] \neq \perp$ then return $\mathrm{ET}[I,T,X]$
If ($I \in \mathrm{ChI}$ and $b = 0$) then
    $Y \leftarrow_{\$} \{\, Y \in \mathsf{F.Dom} \,:\, \mathrm{DT}[I,T,Y] = \perp \,\}$
Else $Y \leftarrow \mathsf{F.E}^{\mathrm{IP}}(J_I,T,X)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $Y$

$\mathrm{Dec}(I,T,Y)$
If $\mathrm{DT}[I,T,Y] \neq \perp$ then return $\mathrm{DT}[I,T,Y]$
If ($I \in \mathrm{ChI}$ and $b = 0$) then
    $X \leftarrow_{\$} \{\, X \in \mathsf{F.Dom} \,:\, \mathrm{ET}[I,T,X] = \perp \,\}$
Else $X \leftarrow \mathsf{F.D}^{\mathrm{IP}}(J_I,T,Y)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $X$

$\mathrm{Exp}(I)$
If $I \in \mathrm{ChI}$ then return $\perp$
$\mathrm{XI} \leftarrow \mathrm{XI} \cup \{I\}$ ; Return $J_I$

$\mathrm{Ch}(I)$
If $I \in \mathrm{XI}$ then return $\perp$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$

$\mathrm{IP}(x)$
$(L,W,s) \leftarrow x$
If $(L = K)$ then
    $\mathsf{bad} \leftarrow \mathsf{true}$ ; $\boxed{P(K,\cdot,+) \leftarrow \overline{\pi} \text{ ; } P(K,\cdot,-) \leftarrow \overline{\pi}^{-1}}$
$y \leftarrow P(x)$ ; Return $y$
```

Figure 11: **Games for proof of Lemma 6.2.**

Games $\mathrm{G}_0, \mathrm{G}_1$ are identical-until-bad, so by the Fundamental Lemma of Game Playing [10] we have

$$\begin{aligned} 2\Pr[\mathrm{G}_0] - 1 &= 2\Pr[\mathrm{G}_1] - 1 + 2\cdot(\Pr[\mathrm{G}_0] - \Pr[\mathrm{G}_1]) \\ &\leq 2\Pr[\mathrm{G}_1] - 1 + 2\Pr[\mathrm{G}_1 \text{ sets } \mathsf{bad}] \ . \end{aligned} \tag{14}$$

Queries to IP may be made directly by the adversary, and there are p such. However, such queries may also be made by the $\mathsf{F.E}$ and $\mathsf{F.D}$ algorithms when invoked in Enc and Dec queries. But we know that for any J the total number of different keys that $\mathsf{F.E}^{\mathrm{IP}}(J,\cdot,\cdot)$ and $\mathsf{F.D}^{\mathrm{IP}}(J,\cdot,\cdot)$ ever use in their oracle queries is limited to $\mathsf{F.nk}$. Since a total of u identities is involved across the Enc and Dec queries we have

$$\Pr[\mathrm{G}_1 \text{ sets } \mathsf{bad}] \leq \frac{p + u\cdot\mathsf{F.nk}}{2^k} \ . \tag{15}$$

Putting together Equations (12)—(15) yields Eq. (11). ▌

The next lemma bounds the ib-prp advantage of a non-adaptive adversary $\mathcal{A}$ relative to $\overline{\mathsf{KDF}}$ via the prpa advantage of a constructed adversary $\overline{\mathcal{A}}$ against the FPE scheme F. This uses a hybrid argument, but done in such a way that the global loss from the structure of the key-derivation scheme remains linear (not quadratic) in the resources.

Lemma 6.3 *Let* F *be a square FPE scheme with* $\mathsf{F.kl} = 2k$. *Let* $\overline{\mathsf{KDF}}$ *be the key derivation function defined by Eq. (10). Let* $\mathcal{A}$ *be a non-adaptive adversary whose* Enc, Dec, Ch *queries involve at most* u *different identities, with at most* q_1 *queries to* Enc, Dec *per identity. Assume* $\mathcal{A}$ *makes* q_e *queries to* Exp *and* p *queries to* IP. *The proof constructs an adversary* $\overline{\mathcal{A}}$ *such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\overline{\mathsf{KDF}}}(\mathcal{A}) \leq u\cdot\mathbf{Adv}^{\mathsf{prpa}}_{\mathsf{F}}(\overline{\mathcal{A}}) + \frac{8u + 8q_e - 6}{2^k} \ . \tag{16}$$

```
Game $\boxed{\mathrm{G}_{2,g}}$, $\mathrm{G}_{3,g}$ $(0 \le g \le u)$
$P \leftarrow\!\!\$\; \mathbf{IC}(k,k)$ ; $\mathrm{ChI} \leftarrow \emptyset$ ; $N \leftarrow u + q_e$
$c \leftarrow 0$ ; $e \leftarrow u$
$D \leftarrow \emptyset$ ; $R \leftarrow \{0,1\}^k$
For $i \in [1..N] \setminus \{g\}$ and $j \in \{0,1\}$ do
    $J_{i,j} \leftarrow\!\!\$\; R$ ; $R \leftarrow R \setminus \{J_{i,j}\}$ ; $D \leftarrow D \cup \{J_{i,j}\}$
$J_{g,0} \leftarrow\!\!\$\; \{0,1\}^k$ ; $J_{g,1} \leftarrow\!\!\$\; \{0,1\}^k$
If $(J_{g,0} \in D)$ then
    $\mathsf{bad} \leftarrow \mathsf{true}$ ; $\boxed{J_{g,0} \leftarrow\!\!\$\; R \;;\; R \leftarrow R \setminus \{J_{g,0}\}}$
If $(J_{g,1} \in D \cup \{J_{g,0}\})$ then $\mathsf{bad} \leftarrow \mathsf{true}$ ; $\boxed{J_{g,1} \leftarrow\!\!\$\; R}$
$b' \leftarrow\!\!\$\; \mathcal{A}^{\mathrm{Enc,Dec,Exp,Ch},P}$
Return $(b' = 1)$

$\underline{\mathrm{Enc}(I,T,X)}$
If $\mathrm{ET}[I,T,X] \ne \perp$ then return $\mathrm{ET}[I,T,X]$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$
If $(I \in \mathrm{ChI}$ and $\mathbf{v}[I] > g)$ then
    $Y \leftarrow\!\!\$\; \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[I,T,Y] = \perp \,\}$
Else $Y \leftarrow \mathsf{F.E}^P(J,T,X)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $Y$

$\underline{\mathrm{Dec}(I,T,Y)}$
If $\mathrm{DT}[I,T,Y] \ne \perp$ then return $\mathrm{DT}[I,T,Y]$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$
If $(I \in \mathrm{ChI}$ and $\mathbf{v}[I] > g)$ then
    $X \leftarrow\!\!\$\; \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[I,T,X] = \perp \,\}$
Else $X \leftarrow \mathsf{F.D}^P(J,T,Y)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $X$

$\underline{\mathrm{Exp}(I)}$
If $(\mathbf{v}[I] = \perp)$ then $e \leftarrow e+1$ ; $\mathbf{v}[I] \leftarrow e$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$ ; Return $J$

$\underline{\mathrm{Ch}(I)}$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
```

```
Adversary $\overline{\mathcal{A}}^{\mathrm{Enc,Dec,Ch,IP}}$
$\mathrm{ChI} \leftarrow \emptyset$ ; $N \leftarrow u + q_e$ ; $c \leftarrow 0$ ; $e \leftarrow u$
$g \leftarrow\!\!\$\; [1..u]$ ; $R \leftarrow \{0,1\}^k$
For $i \in [1..N] \setminus \{g\}$ and $j \in \{0,1\}$ do
    $J_{i,j} \leftarrow\!\!\$\; R$ ; $R \leftarrow R \setminus \{J_{i,j}\}$
$J_{g,0} \leftarrow \perp$ ; $J_{g,1} \leftarrow \perp$
$b' \leftarrow\!\!\$\; \mathcal{A}^{\mathrm{EncSim,DecSim,ExpSim,ChSim,IP}}$
Return $b'$

$\underline{\mathrm{EncSim}(I,T,X)}$
If $\mathrm{ET}[I,T,X] \ne \perp$ then return $\mathrm{ET}[I,T,X]$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$
If $(I \in \mathrm{ChI}$ and $\mathbf{v}[I] > g)$ then
    $Y \leftarrow\!\!\$\; \{\, Y \in \mathsf{F.Dom} : \mathrm{DT}[I,T,Y] = \perp \,\}$
If $(\mathbf{v}[I] = g)$ then $Y \leftarrow \mathrm{Enc}(T,X)$
If $(I \notin \mathrm{ChI}$ or $\mathbf{v}[I] < g)$ then $Y \leftarrow \mathsf{F.E}^{\mathrm{IP}}(J,T,X)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $Y$

$\underline{\mathrm{DecSim}(I,T,Y)}$
If $\mathrm{DT}[I,T,Y] \ne \perp$ then return $\mathrm{DT}[I,T,Y]$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$
If $(I \in \mathrm{ChI}$ and $\mathbf{v}[I] > g)$ then
    $X \leftarrow\!\!\$\; \{\, X \in \mathsf{F.Dom} : \mathrm{ET}[I,T,X] = \perp \,\}$
If $(\mathbf{v}[I] = g)$ then $X \leftarrow \mathrm{Dec}(T,Y)$
If $(I \notin \mathrm{ChI}$ or $\mathbf{v}[I] < g)$ then $X \leftarrow \mathsf{F.D}^P(J,T,Y)$
$\mathrm{ET}[I,T,X] \leftarrow Y$ ; $\mathrm{DT}[I,T,Y] \leftarrow X$ ; Return $X$

$\underline{\mathrm{ExpSim}(I)}$
If $(\mathbf{v}[I] = \perp)$ then $e \leftarrow e+1$ ; $\mathbf{v}[I] \leftarrow e$
$J \leftarrow J_{\mathbf{v}[I],0} \| J_{\mathbf{v}[I],1}$ ; Return $J$

$\underline{\mathrm{ChSim}(I)}$
If $(\mathbf{v}[I] = \perp)$ then $c \leftarrow c+1$ ; $\mathbf{v}[I] \leftarrow c$
If $(\mathbf{v}[I] = g)$ then $\mathrm{Ch}(I)$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$
```

Figure 12: **Games and adversary for proof of Lemma 6.3.**

Adversary $\overline{\mathcal{A}}$ makes at most q_1 queries to Enc, Dec *and p queries to* IP. *Its running time is about the same as that of $\mathcal{A}$.*

Proof of Lemma 6.3: Let $N = u + q_e$. Let $I_1, \ldots, I_u$ denote the identities involved in $\mathcal{A}$'s Enc, Dec, Ch queries. Since $\mathcal{A}$ is non-adaptive, these are distinct from the identities, denoted $I_{u+1}, \ldots, I_N$, in $\mathcal{A}$'s Exp queries. To be more precise, since this is how the proof makes crucial use of the non-adaptivity assumption on $\mathcal{A}$, the sets $\{I_1, \ldots, I_u\}$ and $\{I_{u+1}, \ldots, I_N\}$ are disjoint.

We would like to use a hybrid argument in which I_g is viewed as the target for $\overline{\mathcal{A}}$. The difficulty is that the keys of different identities are not independent so we cannot simulate the keys of non-target identities without knowing the target key, and the latter is of course denied to us in the reduction. We could move to a game with random, independent keys, but this would result in an additive security loss involving terms like $N^2/2^k$. The following argument keeps the loss to $N/2^k$.

Consider games $\mathrm{G}_{2,g}, \mathrm{G}_{3,g}$ of Fig. 12, where $g \in [0..u]$ is an associated parameter. Here $J_i = J_{i,0} \| J_{i,1}$ is the key associated to I_i for $i \in [1..N]$. Rather than specify the keys via a permutation $\overline{\pi}$ as prescribed by KDF, we consider sampling them directly, meaning the $2N$ k-bit strings $J_{i,j}$ for $i \in [1..N]$ and $j \in \{0,1\}$ are random subject to being distinct. The games do this, but not quite. The key $J_g = J_{g,0} \| J_{g,1}$ is treated differently, being sampled uniformly at random, independently of other keys. If $J_{g,0}, J_{g,1}$ coincide with some other $J_{i,j}$ or with each other, the distribution is incorrect. Game $\mathrm{G}_{2,g}$, which includes the boxed code, corrects, re-sampling this key to obey the distinctness rule, but game $\mathrm{G}_{3,g}$, which excludes the boxed code, does not correct. The former reflects the real game, the latter the one conducive to doing our hybrid because non-target keys can be sampled without knowing the target key. (It is important that we did not overkill by asking all keys to be independent of each other in the second game, for this would incur the quadratic security loss.) While we have discussed J_i as associated to I_i, the identities to be queried are not known upfront, and the allocation of an index $\mathbf{v}[I]$ to identity I is made dynamically at the time identity I is first queried to Enc or Dec. Queries to Exp are answered directly, simply revealing the created keys. The games do not pick a challenge bit, instead returning true when the output b' of $\mathcal{A}$ is 1, and false otherwise. When $g = u$, all Enc, Dec queries are answered via F, and when $g = 0$ they are answered randomly but consistently with prior replies, so that

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) = \Pr[\mathrm{G}_{2,u}] - \Pr[\mathrm{G}_{2,0}] \ . \tag{17}$$

For each g, the two games $\mathrm{G}_{2,g}, \mathrm{G}_{3,g}$ are identical-until-bad, so by the Fundamental Lemma of Game Playing [10] we have

$$\begin{aligned}
&\Pr[\mathrm{G}_{2,u}] - \Pr[\mathrm{G}_{2,0}] \\
&\quad = \Pr[\mathrm{G}_{3,u}] + (\Pr[\mathrm{G}_{2,u}] - \Pr[\mathrm{G}_{3,u}]) - \Pr[\mathrm{G}_{3,0}] + (\Pr[\mathrm{G}_{3,0}] - \Pr[\mathrm{G}_{2,0}]) \\
&\quad \leq \Pr[\mathrm{G}_{3,u}] - \Pr[\mathrm{G}_{3,0}] + \Pr[\mathrm{G}_{3,u} \text{ sets } \mathsf{bad}] + \Pr[\mathrm{G}_{3,0} \text{ sets } \mathsf{bad}] \ .
\end{aligned} \tag{18}$$

In game $\mathrm{G}_{3,g}$, the set D has size $2N - 2$ at the time of the test "$J_{g,0} \in D$," so bad is set here with probability $(2N-2)/2^k$. Similarly the test involving $J_{g,1}$ sets bad with probability at most $(2N-1)/2^k$, so

$$\forall\, g \in [0..u] \ : \ \Pr[\mathrm{G}_{3,g} \text{ sets } \mathsf{bad}] \leq \frac{4N-3}{2^k} \ . \tag{19}$$

Using Equations (17), (18) and (19), we have

$$\mathbf{Adv}^{\mathsf{ib\text{-}prp}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \leq \Pr[\mathrm{G}_{3,u}] - \Pr[\mathrm{G}_{3,0}] + \frac{8N-6}{2^k} \ . \tag{20}$$

We use a hybrid argument to bound $\Pr[\mathrm{G}_{3,u}] - \Pr[\mathrm{G}_{3,0}]$. Consider adversary $\overline{\mathcal{A}}$ of Fig. 12. It picks g at random from $[0..u]$, and then picks keys $J_{i,j}$ for $i \neq g$ to be random but distinct. It then runs $\mathcal{A}$. It simulates the latter's Enc, Dec, Exp, Ch oracles with the shown sub-routines EncSim, DecSim, ExpSim, ChSim, respectively. For IP, it directly uses its own IP oracle. The ability to do the latter is important and is why we needed Lemma 6.2 to remove all uses of the ideal primitive other than those made by F. In answering Enc, Dec queries of $\mathcal{A}$ for an identity I_i, it uses F under the keys it has created if $i < g$, forwards the queries to its own Enc, Dec oracles if $i = g$ —so that J_g is identified with the hidden key in game $\mathbf{G}^{\mathsf{prp}}_{\mathsf{F}}(\overline{\mathcal{A}})$— and answers randomly if

$i > g$, all this adjusted to take into account whether or not the identity is in ChI. A $\textsc{Exp}(I)$ query of $\mathcal{A}$ can be answered because $\mathbf{v}[I] \neq g$ so $\overline{\mathcal{A}}$ created, and has, the relevant key, and can return it. In answering a $\textsc{Ch}(I)$ query, $\overline{\mathcal{A}}$ calls its own $\textsc{Ch}$ oracle with I if $\mathbf{v}[I] = g$. We have

$$\begin{aligned}\mathbf{Adv}^{\mathsf{prp}}_{\mathsf{F}}(\overline{\mathcal{A}}) &= \frac{1}{u} \cdot \sum_{i=1}^{u} \Pr[\mathrm{G}_{3,i}] - \Pr[\mathrm{G}_{3,i-1}] \\ &= \frac{1}{u} \left(\Pr[\mathrm{G}_{3,u}] - \Pr[\mathrm{G}_{3,0}] \right) . \qquad (21)\end{aligned}$$

Equations (20) and (21) imply Eq. (16). ∎

7 Pre-masking-based IB-FPE

While Theorem 6.1 shows that if we adjoin $\mathbf{Dbl}[k, \mathsf{M}]$ to an ideal square FPE F, the resulting IB-FPE scheme $(\mathsf{F}, \mathbf{Dbl}[k, \mathsf{M}])$ has k-bit ib-prp security, we'd like to have some provable guarantees if F is concretely instantiated from the base FPE scheme of DFF. However, while the Feistel structure of DFF seems to have very strong empirical security, it's notoriously hard to give even a satisfactory prp bound on Feistel networks on small domains. Let us now elaborate on the reason of this difficulty. Recall that in an information-theoretic proof for prp security of Feistel (such as the classic Luby-Rackoff result [29]), all current techniques can only give a bound based on the number of queries of the adversary, but not its running time. However, for a r-round balanced Feistel network on domain $\{0, 1\}^{2n}$, by a simple counting argument, if $r < 2^n$, there *is* an adversary (of astronomical running time) that makes only 2^n $\textsc{Enc}$ queries and wins with advantage very close to 1. But in our setting, n can be any number greater than 3, whereas in practice, r is often at most 36.

Given the huge obstacle in proving ib-prp security as described above, we turn into ib-kr-ti security. We now give a class of square FPE constructions that we call *pre-masking* FPE, such that for *any* F in this class, $(\mathsf{F}, \mathbf{Dbl}[k, \mathsf{M}])$ has nearly k-bit ib-kr-ti security. Members of this class use an ideal cipher $P \in \mathbf{IC}(k, k)$ (which will be instantiated via AES), but we make no other hardness assumption. This class includes the FPE scheme of DFF, and thus justifies the design choice of DFF. We warn that we only claim ib-kr security, and a pre-masking FPE therefore might be subject to different attacks. Thus our security guarantee here doesn't contradict the message-recovery attacks of BHT [6] on Feistel-based FPE schemes, including DFF. (These attacks, however, are easily put out of reach by increasing the number of rounds on small inputs.) Likewise, our security claim for pre-masking FPEs does not contract the recent message-recovery attack of Durak and Vaudenay [18] that exploits a bug in the design of round functions of FF3.

PRE-MASKING FPE. Let F be a square FPE scheme, meaning F has key-length $\mathsf{F.kl} = 2k$ and its ideal primitive is $\mathsf{F.IP} = \mathbf{IC}(k, k)$. We say that F is *pre-masking* if it additionally specifies algorithms $\mathsf{F.EC}, \mathsf{F.DC}$ (we call them encode and decode) such that

$$\mathsf{F.E}^P(J, T, X) = \mathsf{F.EC}^{\mathsf{Round}^P(J, \cdot)}(T, X)$$
$$\mathsf{F.D}^P(J, T, Y) = \mathsf{F.DC}^{\mathsf{Round}^P(J, \cdot)}(T, Y) \; ,$$

where we have defined

```
$\underline{\mathsf{Round}^P(J, U)}$
$J_1 \leftarrow J[1:k]$ ; $J_2 \leftarrow J[k+1:2k]$
Return $P(J_1, U \oplus J_2, +)$.
```

```
$\underline{\mathsf{F.EC}^f(T, X)}$
$L \leftarrow X[1:n]; R \leftarrow X[n+1:2n]$
$\ell \leftarrow k - n - t$
For $i = 1$ to 10 do
    $V \leftarrow f([i]_\ell \,\|\, T \,\|\, R)$
    $L' \leftarrow R; R \leftarrow L \oplus V[1:n]$ ; $L \leftarrow L'$
Return $L \,\|\, R$
```

```
$\underline{\mathsf{F.DC}^f(T, Y)}$
$L \leftarrow Y[1:n]; R \leftarrow Y[n+1:2n]$
$\ell \leftarrow k - n - t$
For $i = 10$ downto 1 do
    $V \leftarrow f([i]_\ell \,\|\, T \,\|\, R)$
    $L' \leftarrow R; R \leftarrow L \oplus V[1:n]$ ; $L \leftarrow L'$
Return $L \,\|\, R$
```

Figure 13: **A 10-round Feistel-based pre-masking FPE scheme F. Here $[i]_\ell$ denotes the ℓ-bit encoding of a number $i \in \{1, \ldots, 10\}$. The oracle $f : \{0,1\}^k \to \{0,1\}^k$ is implemented as $\mathsf{Round}^P(J, \cdot)$.**

That is, $\mathsf{F.E}$ and $\mathsf{F.D}$ use the $2k$-bit key J in a limited way, through Round. The latter takes a k-bit input and implements Rivest's classical DESX construction on top of the ideal-cipher instance P, but omits the post-whitening (meaning that the output is not XOR'ed with J_2). Note the encoding and decoding functions do not have direct access to the key J; they can only access it indirectly through queries to $\mathsf{Round}^P(J, \cdot)$. As an example, if $\mathsf{F.Dom} = \{0,1\}^{2n}$ and $\mathsf{F.TS} = \{0,1\}^t$ for $n + t \le k - 4$, a 10-round Feistel-based pre-masking FPE scheme can be built as in Fig. 13.

The efficiency improvement we obtain (due to dropping the post-whitening in DESX) is based on the fact that Round only calls the forward direction of the ideal cipher.

<u>SECURITY ANALYSIS.</u> As a stepping stone in obtaining the ib-kr-ti security of a pre-masking FPE scheme F, we consider security of the following FPE scheme $\overline{\mathsf{F}}$. Informally, the scheme $\overline{\mathsf{F}}$ is a blockcipher, implementing the DESX variant on top of AES. That is, FPE scheme $\overline{\mathsf{F}}$ has $\overline{\mathsf{F}}.\mathsf{Dom} = \{0,1\}^k$ and $\overline{\mathsf{F}}.\mathsf{TS} = \{\varepsilon\}$. Its encryption scheme $\overline{\mathsf{F}}.\mathsf{E}^P(J, T, X)$ returns $\mathsf{Round}(J, X)$, and the decryption scheme is defined accordingly.

In Lemma 7.1 below, we'll reduce the ib-kr-ti security of F to the ib-kr-ti security of $\overline{\mathsf{F}}$, both relative to $\mathbf{Dbl}[k, \mathsf{M}]$. The constructed adversary however makes no DEC query in attacking $\overline{\mathsf{F}}$. This restriction is crucial, because in $\overline{\mathsf{F}}$, there's pre-whitening but no post-whitening of the output of $P(J_1, \cdot, +)$.

Lemma 7.1 *Let F be a pre-masking FPE scheme of $\mathsf{F.kl} = 2k$ and $\overline{\mathsf{F}}$ be as described above. Let KDF be the key-derivation function $\mathbf{Dbl}[k, \mathsf{M}]$. Suppose that we are given an adversary $\mathcal{A}$ whose* ENC/DEC *queries involve at most q calls to Round. Assume $\mathcal{A}$ makes q_e queries to* EXP *and p queries to* IP. *Then we can construct an adversary $\overline{\mathcal{A}}$ of the same number of* IP *and* EXP *queries such that*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \le \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\overline{\mathcal{A}}) \enspace .$$

Adversary $\overline{\mathcal{A}}$ makes at most q ENC *queries and no* DEC *query.*

Proof of Lemma 7.1: Adversary $\overline{\mathcal{A}}$ runs $\mathcal{A}$ and shares the EXP and P oracles with it. When $\mathcal{A}$ wants to encrypt (I, T, X), adversary $\overline{\mathcal{A}}$ runs $\mathsf{F.EC}^{\text{ENC}(I,\varepsilon,\cdot)}(T, X)$, where ENC is $\overline{\mathcal{A}}$'s own encryption oracle. Likewise, when $\mathcal{A}$ wants to decrypt (I, T, Y), adversary $\overline{\mathcal{A}}$ runs $\mathsf{F.DC}^{\text{ENC}(I,\varepsilon,\cdot)}(T, Y)$. Finally, when $\mathcal{A}$ outputs its guessed key, adversary $\overline{\mathcal{A}}$ returns the same output. Hence game $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\overline{\mathcal{A}})$ coincides with game $\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A})$, and thus

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \le \mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\overline{\mathcal{A}}) \enspace ,$$

as claimed. ▌

Next, we bound the ib-kr-ti security of $\overline{\mathsf{F}}$ relative to $\mathbf{Dbl}[k, \mathsf{M}]$, but the adversary is forbidden from calling Dec. The analysis is challenging, because there's no post-whitening of the output of $P(J_1, \cdot, +)$ in the encryption scheme of $\overline{\mathsf{F}}$, yet the adversary can still query $P(\cdot, \cdot, -)$. The proof is in Appendix A. We note that if q is small, say $q \le 2^k/k^3$, then in Lemma 7.2 the blowup $k/\lg(k)$ can be reduced to $\frac{3k}{k-\lg(q)}$. However, for the practical choice $k = 128$, the blowup $k/\lg(k)$ is smaller than 19 and the bound in Lemma 7.2 is already satisfactory.

Lemma 7.2 *Let $\overline{\mathsf{F}}$ be as described above and let KDF be the key-derivation function $\mathbf{Dbl}[k, \mathsf{M}]$. Assume that $k \ge 128$. Then for any adversary $\mathcal{A}$ that makes at most $q \le 2^{k-2}$ queries to* Enc*, no query to* Dec*, $q_e \le 2^{k-3}$ queries to* Exp*, and p queries to* IP*,*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{A}) \le \frac{2q(p+1)}{2^{2k}} + \frac{4(p+1)k}{2^k \cdot \lg(k)} + \frac{q+q_e+p+5}{2^k} \ .$$

Combining Lemma 7.1 and Lemma 7.2, we immediately obtain the following result.

Theorem 7.3 *Let F be a pre-masking FPE scheme of $\mathsf{F.kl} = 2k$ and let KDF be the key-derivation function $\mathbf{Dbl}[k, \mathsf{M}]$. Assume that $k \ge 128$. Suppose that we are given an adversary $\mathcal{A}$ whose* Enc/Dec *queries involve at most q calls to* Round*. Assume $\mathcal{A}$ makes q_e queries to* Exp *and p queries to* IP*. Then*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \le \frac{2q(p+1)}{2^{2k}} + \frac{4(p+1)k}{2^k \cdot \lg(k)} + \frac{q+q_e+p+5}{2^k} \ .$$

We note that the results of Lemma 7.2 and Theorem 7.3 hold even for adaptive adversaries if the ideal cipher is programmable. If the ideal cipher is non-programmable then these results only work for non-adaptive adversaries.

<u>A matching attack.</u> In Lemma 7.2, at the first glance, the blowup $k/\lg(k)$ looks like an artifact of the analysis. However, Proposition 7.4 shows that it's inherent by demonstrating a matching key-recovery attack. The proof is nontrivial. In both Lemma 7.2 and Proposition 7.4, the term $k/\lg(k)$ comes from some balls-into-bins phenomena.

Proposition 7.4 *Let $\overline{\mathsf{F}}$ be as described above. Let KDF be the key-derivation function $\mathbf{Dbl}[k, \mathsf{M}]$. Assume that $k \ge 128$. Let $r = \lfloor k/9\lg(k) \rfloor$ and let $q = r\lfloor 2^k/9r^2 \rfloor \approx 2^k \lg(k)/k$. Then we can construct a non-adaptive adversary $\mathcal{A}$ making at most $q + r$ queries* Enc *queries and $q + r$ queries to* IP*, a single query to* Ch*, and no query to* Exp *or* Dec *query, yet*

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}}(\mathcal{A}) \ge \frac{(1 - 5 \cdot 2^{-k/9})qr}{2^{k+1}} \ge \frac{1}{19} \ .$$

Proof of Proposition 7.4: The adversary $\mathcal{A}$ is constructed as in Fig. 14. It first picks arbitrary distinct identities $I_1, \ldots, I_r$ and distinct messages $X_1, \ldots, X_{q/r}, X^*$. It then queries $\text{Enc}(I_i, \varepsilon, X_j)$, for every $i \in \{1, \ldots, r\}$ and every $j \in \{1, \ldots, q/r\}$. View each such query as throwing a ball into 2^k possible bins corresponding to strings in $\{0, 1\}^k$. Of course the throws are not independent. For example, the balls corresponding to $\text{Enc}(I_1, \varepsilon, X_1)$ and $\text{Enc}(I_1, \varepsilon, X_2)$ must land into different bins. Note that each bin can have at most r balls. Further below, we'll adapt techniques from classic

```
Adversary $\mathcal{A}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
Pick distinct identities $I_1,\ldots,I_r \in \mathsf{KDF.IS}$
Pick distinct messages $X_1,\ldots,X_{q/r},X^* \in \{0,1\}^k$
For $i=1$ to $r$, $j=1$ to $q/r$ do $B_{i,j} \leftarrow \textsc{Enc}(I_i,\varepsilon,X_j)$
$(C,C',M_1,M'_1,\ldots,M_r,M'_r) \leftarrow \mathsf{Match}(X_1,\ldots,X_{q/r},B_{1,1},\ldots,B_{r,q/r})$
If $C=\perp$ then return $\perp$ // Terminate if there are no two full bins
$J_1,\ldots,J_q \leftarrow\!\!\$\, \{0,1\}^k$
For $j=1$ to $q$ do
   $U_j \leftarrow P(J_j,C,-)$ ; $U'_j \leftarrow P(J_j,C',-)$ ; $V_j \leftarrow U_j \oplus U'_j$
For $i=1$ to $r$ do
   $V \leftarrow M_i \oplus M'_i$ ; $j \leftarrow \mathsf{Find}(V,V_1,\ldots,V_q)$
   If $j>0$ then // Test for false positive
      $R_1 \leftarrow \textsc{Enc}(I_i,\varepsilon,X^*)$ ; $R_2 \leftarrow P(J_j,X^* \oplus U_j \oplus M_i)$
      If $R_1=R_2$ then ($\textsc{Ch}(I_i)$; return $(I_i, J_j \,\|\, (U_j \oplus M_i))$)
```

Figure 14: **Constructed ib-kr-ti adversary for $\overline{\mathsf{F}}$.** Here procedure $\mathsf{Find}(V,V_1,\ldots,V_q)$ returns an index j such that $V=V_j$, and returns 0 if no such j exists. Procedure $\mathsf{Match}(X_1,\ldots,X_{q/r},B_{1,1},\ldots,B_{r,q/r})$ returns $(C,C',M_1,M'_1,\ldots,M_r,M'_r)$ such that for every $i \in \{1,\ldots,r\}$, there are $1 \le \ell_i < s_i \le q/r$ such that $M_i = X_{\ell_i}$, $M'_i = X_{s_i}$, $C = B_{i,\ell_i}$, and $C' = B_{i,s_i}$.

balls-into-bins papers [33] to show that with probability at least $1 - 4 \cdot 2^{-k/9}$, there are two bins that both have r balls.

Suppose that there exist two bins of r balls. That is, for each identity I_i, there are two messages M_i and M'_i whose balls fall into these two bins. Let C and C' be the answer for $\textsc{Enc}(I_i,\varepsilon,M_i)$ and $\textsc{Enc}(I_i,\varepsilon,M'_i)$, respectively. Let L_i be the subkey of I_i, and K the master key. The adversary then picks $J_1,\ldots,J_q \leftarrow\!\!\$\, \{0,1\}^k$, and queries $P(J_j,C,-)$ and $P(J_j,C',-)$ to get answers U_j and U'_j, respectively, for every $j \in \{1,\ldots,q\}$. For each $i \in \{1,\ldots,r\}$, it tries to find an index $j \in \{1,\ldots,q\}$ such that $M_i \oplus M'_i = U_j \oplus U'_j$. For each matched pair (i,j), it is likely that $J_j \,\|\, (U_j \oplus M_i)$ is the subkey of I_i. The adversary will compare $\textsc{Enc}(I_i,\varepsilon,X^*)$ and $P(J_j, X^* \oplus (U_j \oplus M_i))$ to eliminate false positive.

For analysis, let Hit be the event that there are some i and j such that $J_j = L_i[1:k]$. If J_j is $L_i[1:k]$ then $U_j = M_i \oplus L_i[k+1:2k]$ and $U'_j = M'_i \oplus L_i[k+1:2k]$, and thus $U_j \oplus U'_j = M_i \oplus M'_j$ and consequently, the subkey L_i is indeed $J_j \,\|\, (U_j \oplus M_i)$. Under the key-derivation function KDF, the strings $L_1[1:k],\ldots,L_r[1:k]$ are distinct, because they are $P(K,\mathsf{M}_0(I_1),+),\ldots,P(K,\mathsf{M}_0(I_r),+)$ respectively, and $I_1,\ldots,I_r$ are distinct. Hence

$$\Pr[\mathsf{Hit}] = 1 - (1 - r/2^k)^q \ge rq/2^{k+1},$$

where the last inequality is from applying Lemma 2.1 for $a = r/2^k \le 1/q$. For $i \in \{1,\ldots,r\}$ and $j \in \{1,\ldots,q\}$, let $\mathsf{Bad}_{i,j}$ be the event that $J_j \ne L_i[1:k]$, but $M_i \oplus M'_i = U_j \oplus U'_j$ and $\textsc{Enc}(I_i,\varepsilon,X^*_i) = P(J_j, X^*_i \oplus (U_j \oplus M_i))$. Let Bad be the event $\exists\, i,j : \mathsf{Bad}_{i,j}$. If Bad does not happen then the testing eliminates all false positive. From union bound,

$$\Pr[\mathsf{Bad}] \le \sum_{i=1}^{r} \sum_{j=1}^{q} \Pr[\mathsf{Bad}_{i,j}] \le \frac{qr}{(2^k-1)(2^k-2)} \le \frac{qr}{2^{1.5k+1}} \enspace .$$

Note that if $\mathsf{Hit} \wedge \overline{\mathsf{Bad}}$ happens then the adversary wins. So if there are two bins of r balls then the adversary wins with advantage at least

$$\Pr[\mathsf{Hit} \wedge \overline{\mathsf{Bad}}] \geq \Pr[\mathsf{Hit}] - \Pr[\mathsf{Bad}] \geq \frac{(1-2^{-k/2})}{2^{k+1}} \enspace .$$

Putting all this together,

$$\begin{aligned}\mathbf{Adv}^{\text{ib-kr-ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{A}) &\geq \frac{(1-4\cdot 2^{-k/9})(1-2^{-k/2})qr}{2^{k+1}} \\ &\geq \frac{(1-5\cdot 2^{-k/9})qr}{2^{k+1}} \enspace .\end{aligned}$$

What's left is to prove that there are two bins of r balls with probability at least $1-4\cdot 2^{-k/9}$. Let $c = q/r = \lfloor 2^k/9r^2 \rfloor$. For any $1 \leq i < j \leq 2^k$, the chance that bins i and j have r balls is

$$\begin{aligned}\left(\frac{\binom{2^k-2}{c-2}}{\binom{2^k}{c}}\right)^r &= \Big(\frac{c(c-1)}{2^k(2^k-1)}\Big)^r \\ &\geq \Big(\frac{(2^k/9r^2-1)(2^k/9r^2-2)}{2^k(2^k-1)}\Big)^r \\ &\geq \Big(\frac{1}{9r^2}-\frac{1}{2^k}\Big)^r\Big(\frac{1}{9r^2}-\frac{3}{2^k-1}\Big)^r \\ &\geq \Big(\frac{1}{100r^4}\Big)^r \geq \frac{1}{k^{4r}} \\ &\geq \frac{1}{k^{4k/9\log_2(k)}} = \frac{1}{2^{4k/9}} \enspace .\end{aligned}$$

Let $Y_{i,j}$ be the Bernoulli random variable such that $Y_{i,j}=1$ if bins i and j both have r balls, and $Y_{i,j}=0$ otherwise. Hence $\mathbf{E}[Y_{i,j}] = \Pr[Y_{i,j}=1] \geq 2^{-4k/9}$. Let $Y = \sum_{1\leq i<j\leq 2^k} Y_{i,j}$, and thus

$$\mathbf{E}[Y] = \sum_{1\leq i<j\leq 2^k} \mathbf{E}[Y_{i,j}] \geq \frac{(2^k-1)2^{5k/9}}{2} \enspace .$$

Our goal is to prove that $1-\Pr[Y\geq 1] \leq 4\cdot 2^{-k/9}$. Note that

$$1-\Pr[Y\geq 1] = \Pr[Y=0] \leq \Pr\Big[\big|Y-\mathbf{E}[Y]\big| \geq \mathbf{E}[Y]\Big] \leq \frac{\mathbf{Var}[Y]}{(\mathbf{E}[Y])^2},$$

where the last inequality is due to Chebyshev's inequality. Hence what's left is to prove that $\mathbf{Var}[Y] \leq 2^k(2^k-1)^2$. One the one hand, for $1 \leq i < j \leq 2^k$ and $1 \leq \ell < s \leq 2^k$, if $\{i,j\}\cap\{\ell,s\} = \emptyset$ then the random variables $Y_{i,j}$ and $Y_{\ell,s}$ are negatively correlated, because if two bins already have r balls, then it's less likely that two other bins also have r balls. Hence the covariance $\mathbf{Cov}(Y_{i,j},Y_{\ell,s})$ is at most 0. On the other hand, for $1 \leq i < j \leq 2^k$ and $1 \leq \ell < s \leq 2^k$, if $\{i,j\}\cap\{\ell,s\} \neq \emptyset$, since $0 \leq Y_{i,j}, Y_{\ell,s} \leq 1$, we have

$$\mathbf{Cov}(Y_{i,j},Y_{\ell,s}) = \mathbf{E}(Y_{i,j}\cdot Y_{\ell,s}) - \mathbf{E}(Y_{i,j})\mathbf{E}(Y_{\ell,s}) \leq 1 \enspace .$$

Moreover, for each $1 \leq i < j \leq 2^k$, there are exactly $(2^k-i)+j$ pairs (ℓ,s) such that $1 \leq \ell < s \leq 2^k$

and $\{i, j\} \cap \{\ell, s\} \neq \emptyset$. Thus there are at most

$$\begin{aligned}
\sum_{1 \le i < j \le 2^k} (2^k + j - i) &= 2^k \binom{2^k}{2} + \sum_{1 \le i < j \le 2^k} (j - i) \\
&= \frac{2^{2k}(2^k - 1)}{2} + \sum_{t=1}^{2^k - 1} t(2^k - t) \\
&\le \frac{2^{2k}(2^k - 1)}{2} + \sum_{t=1}^{2^k - 1} \frac{(t + 2^k - t)^2}{4} \\
&= \frac{3 \cdot 2^{2k}(2^k - 1)}{4} \le 2^k(2^k - 1)^2
\end{aligned}$$

tuples (i, j, ℓ, s) such that $1 \le i < j \le 2^k$, $1 \le \ell < s \le 2^k$, and $\{i, j\} \cap \{\ell, s\} \neq \emptyset$. Therefore,

$$\begin{aligned}
\mathbf{Var}[Y] &= \mathbf{Var}\Big[\sum_{1 \le i < j \le 2^k} Y_{i,j}\Big] = \sum_{\substack{1 \le i < j \le 2^k \\ 1 \le \ell < s \le 2^k}} \mathbf{Cov}(Y_{i,j}, Y_{\ell,s}) \\
&\le \sum_{\substack{1 \le i < j \le 2^k \\ 1 \le \ell < s \le 2^k \\ \{i,j\} \cap \{\ell,s\} \neq \emptyset}} \mathbf{Cov}(Y_{i,j}, Y_{\ell,s}) \le 2^k(2^k - 1)^2 \ .
\end{aligned}$$

This concludes the proof. ▌

8 Security Analysis of DFF

Here we discuss how to cast DFF as an IB-FPE scheme obtained via the **Dbl** transform and apply the results of Sections 6 and 7 to validate its security, as long as (1) the tweak (identity) space is appropriately restricted and (2) the radix and input length are fixed. Limitation (1) arises because, over the full tweak (identity) space, the M_1 embedding function is not injective: even for a fixed radix and input length, two tweaks may have derived keys with the same second halves. This does not, as far as we know, give rise to a damaging attack (we give below the best attack we could find) but it can be viewed as a design weakness. We suggest modifications to the embedding that restore injectivity and allow our results to apply. Limitation (2) means that a (master) key is used for just one choice of radix and tweak. To prove security for varying radix and input lengths would require that we use the broader definition of FPE from [9] in which the domain is a union of slices, in our case a slice being associated to a choice of radix, input length, and tweak.

DFF AS IB-FPE. We first briefly explain how to view DFF [39] as an IB-FPE scheme $(\mathsf{F}, \mathsf{KDF}_{\mathsf{dff}})$. (See Appendix B for the complete specification.) The DFF specification allows different choices of radix rdx and input length n, but here we fix both, so that $\mathsf{F.Dom} = \mathbb{Z}_{\mathsf{rdx}}^n$. F has 256-bit keys and tweak space the singleton set $\{\varepsilon\}$. The algorithm itself is a 10-round Feistel network. The identity space $\mathsf{KDF}_{\mathsf{dff}}.\mathsf{IS}$ is the set of all $I \in \{0,1\}^*$ such that $|I|$ is at most 13 bytes. The underlying blockcipher $E : \{0,1\}^{128} \times \{0,1\}^{128} \to \{0,1\}^{128}$ is AES. Let $[x]^b$ denote the representation of x as a b-byte string. The embedding scheme $\mathsf{M} = (\mathsf{M}_0, \mathsf{M}_1)$ is specified via $\mathsf{M}_0(I) = [\mathsf{rdx}]^1 \,\|\, [|I|]^1 \,\|\, [n]^1 \,\|\, [I]^{13}$ and $\mathsf{M}_1(I) = [0]^3 \,\|\, [I]^{13}$. Note that M_1 is *not* injective: for example, $\mathsf{M}_1(00) = \mathsf{M}_1(000)$.

SECURITY OVER RESTRICTED IDENTITY SPACES. If the radix and input length are fixed, and one restricts the identities to a subset $S \subset \mathsf{KDF}_{\mathsf{dff}}.\mathsf{IS}$ such that no two strings in S correspond to the

```
Adversary $\mathsf{KR}_{p,d}^{\textsc{Enc},\textsc{Dec},\textsc{Exp},\textsc{Ch},P}$
For $i \leftarrow 0$ to 104 do $I_i \leftarrow 0^i$
$J \leftarrow \textsc{Exp}(I_0)$; $R \leftarrow J[k+1 : :2k]$; $S \leftarrow \emptyset$
For $\ell \leftarrow 1$ to $d$ do $X_\ell \leftarrow\!\!\$\ \mathsf{F.Dom} \backslash S$ ;  $S \leftarrow S \cup \{X_\ell\}$
For $i \leftarrow 1$ to 104 do
  For $\ell \leftarrow 1$ to $d$ do $V_\ell \leftarrow \textsc{Enc}(I_i, \varepsilon, X_\ell)$
  $Z_i \leftarrow (V_1, \ldots, V_d)$
For $j \leftarrow 1$ to $p$ do
  $L_j \leftarrow\!\!\$\ \{0,1\}^{128}$; $J_j \leftarrow L_j \,\|\, R$
  For $\ell \leftarrow 1$ do $d$ do $U_\ell \leftarrow \mathsf{F.E}^P(J_j, T, X_\ell)$
  $Z \leftarrow (U_1, \ldots, U_\ell)$ ; $i \leftarrow \mathsf{Find}(Z, Z_1, \ldots, Z_{104})$
  If $i > 0$ then ($\textsc{Ch}(I_i)$; Return $(I_i, J_j)$)
```

Figure 15: **The attack $\mathsf{KR}_{p,d}$ on the IB-FPE scheme** $(\mathsf{F}, \mathsf{KDF}_{\mathsf{dff}})$**.**

same integer in binary, then the embedding functions M_0 and M_1 above are injective and have disjoint images. Under these restrictions, our results in Sections 6 and 7 apply, and DFF has k-bit non-adaptive ib-prp security, and k-bit adaptive ib-kr-ti security.

SECURITY OVER THE FULL IDENTITY SPACE. The non-injectivity of M_1 allows an adversary to get the second half of the subkey of an identity I without querying $\textsc{Exp}(I)$, by picking another identity I' such that $[I']^{13} = [I]^{13}$, and calling $\textsc{Exp}(I')$. Note that for any $I' \in \mathsf{KDF}_{\mathsf{dff}}.\mathsf{IS}$, there are up to 104 other identities $I \in \mathsf{KDF}_{\mathsf{dff}}.\mathsf{IS}$ such that $[I']^{13} = [I]^{13}$. This leads to the non-adaptive ib-kr-ti adversary $\mathsf{KR}_{p,d}$ shown in Fig. 15. It picks identities $I_0 = \varepsilon, I_1 = 0, I_2 = 0^2, \ldots, I_{104} = 0^{104}$. Note that $[I_0]^{13} = \cdots = [I_{104}]^{13}$. It first queries $J \leftarrow \textsc{Exp}(I_0)$, and let $R \leftarrow J[k+1:2k]$. Note that for any $i \le 104$, R is also the right half of the subkey of identity I_i. The adversary now picks p candidates subkeys $J_1, \ldots, J_p$ such that $J_j[k+1:2k] = L$. Now, for every $i \in \{1, \ldots, 104\}$ and $j \in \{1, \ldots, p\}$, it aims to test whether J_j is the subkey of I_i by comparing $\textsc{Enc}(I_i, \varepsilon, \cdot)$ and $\mathsf{F.E}(J_j, \varepsilon, \cdot)$ on d messages. Proposition 8.1 below shows that this attack achieves ib-kr-ti advantage about $104p/2^{129} - 104p \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$, where the false positive advantage $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$ was defined in Section 3.

Proposition 8.1 *Let* $(\mathsf{F}, \mathsf{KDF}_{\mathsf{dff}})$ *be as above. Then for any* $p, d \in \mathbb{N}$ *such that* $p \le 2^{128}/104$ *we have*

$$\mathbf{Adv}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}_{\mathsf{dff}}}(\mathsf{KR}_{p,d}) \ge \frac{104p}{2^{129}} - 104p \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \ . \ \blacksquare$$

Proof of Proposition 8.1: Let $X_1, \ldots, X_d$ be the test messages that the adversary samples. Let K denote the master key chosen in the overlying key-recovery game $\mathbf{G}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}_{\mathsf{dff}}}(\mathsf{KR}_{p,d})$ and let $J'_i = \mathsf{KDF}^P_{\mathsf{dff}}(K, I_i)$ for $0 \le i \le 104$. Let Hit be the event that some guess J_j of the adversary is one of the target keys, meaning there are $i \in \{1, \ldots, 104\}$ and $j \in \{1, \ldots, p\}$ such that $J_j = J'_i$. For $i \in \{1, \ldots, 104\}$ and $j \in \{1, \ldots, p\}$ let $\mathsf{Bad}_{i,j}$ be the event that $J_j \ne J'_i$ and $(\mathsf{F.E}^P(J_j, \varepsilon, X_1), \ldots, \mathsf{F.E}^P(J_j, \varepsilon, X_d)) = (\textsc{Enc}(I_i, \varepsilon, X_1), \ldots, \textsc{Enc}(I_i, \varepsilon, X_d))$. Let Bad be the event $\exists\, i, j : \mathsf{Bad}_{i,j}$. If $\mathsf{Hit} \wedge \overline{\mathsf{Bad}}$ happens then one of the adversary's guesses is one of the target keys, and there are no false positive during the testing. So

$$\mathbf{Adv}^{\text{ib-kr-ti}}_{\mathsf{F},\mathsf{KDF}_{\mathsf{dff}}}(\mathsf{KR}_{p,d}) \ge \Pr[\mathsf{Hit} \wedge \overline{\mathsf{Bad}}] \ge \Pr[\mathsf{Hit}] - \Pr[\mathsf{Bad}] \ .$$

First we lower bound $\Pr[\mathsf{Hit}]$. Note that J_j and J'_i have the same right half for any i and j, and the target keys $J'_1, \ldots, J'_{104}$ are distinct. Then

$$\Pr[\mathsf{Hit}] = 1 - \left(1 - \frac{104}{2^{128}}\right)^p \geq \frac{104p}{2^{129}},$$

where the inequality is due to Lemma 2.1. Next we upper bound $\Pr[\mathsf{Bad}]$. For any $i \in \{1, \ldots, 104\}$ and $j \in \{1, \ldots, p\}$, if $J_j \neq J'_i$ then the probability that $(\mathsf{F.E}^P(J_j, T, X_1), \ldots, \mathsf{F.E}^P(J_j, T, X_d)) = (\textsc{Enc}(I_i, T, X_1), \ldots, \textsc{Enc}(I_i, T, X_d))$ is at most $\mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$, and hence $\Pr[\mathsf{Bad}_{i,j}] \leq \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$. By the union bound we have $\Pr[\mathsf{Bad}] \leq 104p \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$. Putting all this together we have

$$\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\mathsf{F},\mathsf{KDF}_{\mathsf{dff}}}(\mathsf{KR}_{p,d}) \geq \frac{104p}{2^{129}} - 104p \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d} \ ,$$

completing the proof. ▌

<u>Discussion.</u> While the attack $\mathsf{KR}_{p,d}$ above is impractical and does not significantly affect the 128-bit security claim of DFF, our results, at least, offer no proof that a better attack is not possible. Furthermore, that the right halves of the keys of two different identities can coincide does not feel right. Accordingly, we recommend fixing this. If rdx is fixed, this could be done by setting $\mathsf{M}_1(I) = [0]^1 \parallel [|I|]^1 \parallel [n]^1 \parallel [I]^{13}$. Alternatively one could restrict the identities as mentioned above. If rdx cannot be viewed as fixed and we want a more natural space of identities, we would suggest to let identities be binary strings of at most 12 bytes, let $\mathsf{M}_0(I) = [0]^1 \| [\mathsf{rdx}]^1 \| [|I|]^1 \| [n]^1 \| [I]^{12}$ and $\mathsf{M}_1(I) = [1]^1 \| [\mathsf{rdx}]^1 \| [|I|]^1 \| [n]^1 \| [I]^{12}$. All these choices ensure the embedding functions satisfy our conditions so that our results in Sections 6 and 7 apply.

Earlier, we mentioned that one can generalize our definition for a general domain that is a union of slices, where a slice is associated with a choice of radix, input length, and tweak. There are several ways to do that. For example, one might treat a triple (radix, input length, tweak) as a generalized tweak, so each identity is associated with a single subkey, for all choices of radii and input lengths. The IB-FPE scheme of DFF, as recast above, however, does not follow this approach. Instead, for each identity, there will be a subkey per (radix, tweak) pair. This means that a local device now may have to hold up to s subkeys, where s is the number of supported pairs (radix, input length). The security definitions would be modified so that $\textsc{Exp}(I)$ returns all s subkeys corresponding to identity I. For those generalized notions, the attack $\mathsf{KR}_{p,d}$ above can be improved to have advantage about $\frac{104ps}{2^{129}} - 104ps \cdot \mathbf{Adv}^{\mathsf{fp}}_{\mathsf{F},d}$, while the running time increases only $O(\log(s))$ times. While the DFF specification allows s to be nearly 2^{16}, in real usage, s would be very small, and thus a local device will not have to store too many subkeys, and the improved $\mathsf{KR}_{p,d}$ attack still does not significantly affect the 128-bit security claim of DFF.

References

[1] M. Abdalla and M. Bellare. Increasing the lifetime of a key: a comparative analysis of the security of re-keying techniques. In T. Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 546–559. Springer, Heidelberg, Dec. 2000. 4

[2] M. Bellare, A. Boldyreva, and S. Micali. Public-key encryption in a multi-user setting: Security proofs and improvements. In B. Preneel, editor, *EUROCRYPT 2000*, volume 1807 of *LNCS*, pages 259–274. Springer, Heidelberg, May 2000. 14

[3] M. Bellare, R. Canetti, and H. Krawczyk. Pseudorandom functions revisited: The cascade construction and its concrete security. In *37th FOCS*, pages 514–523. IEEE Computer Society Press, Oct. 1996. 14

[4] M. Bellare, A. Desai, E. Jokipii, and P. Rogaway. A concrete security treatment of symmetric encryption. In *38th FOCS*, pages 394–403. IEEE Computer Society Press, Oct. 1997. 13

[5] M. Bellare, R. Dowsley, B. Waters, and S. Yilek. Standard security does not imply security against selective-opening. In D. Pointcheval and T. Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 645–662. Springer, Heidelberg, Apr. 2012. 4, 13

[6] M. Bellare, V. T. Hoang, and S. Tessaro. Message-recovery attacks on feistel-based format preserving encryption. In E. R. Weippl, S. Katzenbeisser, C. Kruegel, A. C. Myers, and S. Halevi, editors, *ACM CCS 16*, pages 444–455. ACM Press, Oct. 2016. 3, 7, 31

[7] M. Bellare, D. Hofheinz, and S. Yilek. Possibility and impossibility results for encryption and commitment secure under selective opening. In A. Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 1–35. Springer, Heidelberg, Apr. 2009. 4, 13

[8] M. Bellare, T. Krovetz, and P. Rogaway. Luby-Rackoff backwards: Increasing security by making block ciphers non-invertible. In K. Nyberg, editor, *EUROCRYPT'98*, volume 1403 of *LNCS*, pages 266–280. Springer, Heidelberg, May / June 1998. 5, 23

[9] M. Bellare, T. Ristenpart, P. Rogaway, and T. Stegers. Format-preserving encryption. In M. J. Jacobson Jr., V. Rijmen, and R. Safavi-Naini, editors, *SAC 2009*, volume 5867 of *LNCS*, pages 295–312. Springer, Heidelberg, Aug. 2009. 3, 9, 36

[10] M. Bellare and P. Rogaway. The security of triple encryption and a framework for code-based game-playing proofs. In S. Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 409–426. Springer, Heidelberg, May / June 2006. 8, 28, 30

[11] E. Biham. How to decrypt or even substitute DES-encrypted messages in 2^{28} steps. *Information Processing Letters*, 84(3):117–124, 2002. 19

[12] J. Black and P. Rogaway. Ciphers with arbitrary finite domains. In B. Preneel, editor, *CT-RSA 2002*, volume 2271 of *LNCS*, pages 114–130. Springer, Heidelberg, Feb. 2002. 3, 9

[13] D. Boneh and M. K. Franklin. Identity-based encryption from the Weil pairing. In J. Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 213–229. Springer, Heidelberg, Aug. 2001. 4, 7, 13

[14] S. Chen and J. P. Steinberger. Tight security bounds for key-alternating ciphers. In P. Q. Nguyen and E. Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 327–350. Springer, Heidelberg, May 2014. 41

[15] W. Dai, V. T. Hoang, and S. Tessaro. Information-theoretic indistinguishability via the Chi-Squared method. In *CRYPTO 2017*, pages 497–523. Springer, 2017. 5, 23, 24

[16] Y. Dai and J. Steinberger. Tight security bounds for multiple encryption. Cryptology ePrint Archive, Report 2014/096, 2014. `http://eprint.iacr.org/2014/096`. 42

[17] A. Desai and S. Miner. Concrete security characterizations of PRFs and PRPs: Reductions and applications. In T. Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 503–516. Springer, Heidelberg, Dec. 2000. 11

[18] F. B. Durak and S. Vaudenay. Breaking and repairing the FF3 format preserving encryption over small domain. In *CRYPTO 2017*, pages 679–707. Springer, 2017. 3, 7, 31

[19] C. Dwork, M. Naor, O. Reingold, and L. J. Stockmeyer. Magic functions. In *40th FOCS*, pages 523–534. IEEE Computer Society Press, Oct. 1999. 4, 13

[20] M. Dworkin. Recommendation for Block Cipher Modes of Operation: Methods for Format-Preserving Encryption. *NIST Special Publication 800-38G*, Mar. 2016. `http://dx.doi.org/10.6028/NIST.SP.800-38G`. 3, 7

[21] M. Dworkin and R. Perlner. Analysis of VAES3 (FF2). Cryptology ePrint Archive, Report 2015/306, 2015. `http://eprint.iacr.org/2015/306`. 5, 6, 7, 19

[22] V. T. Hoang, B. Morris, and P. Rogaway. An enciphering scheme based on a card shuffle. In R. Safavi-Naini and R. Canetti, editors, *CRYPTO 2012*, volume 7417 of *LNCS*, pages 1–13. Springer, Heidelberg, Aug. 2012. 7

[23] D. Hofheinz, V. Rao, and D. Wichs. Standard security does not imply indistinguishability under selective opening. In M. Hirt and A. D. Smith, editors, *TCC 2016-B, Part II*, volume 9986 of *LNCS*, pages 121–145. Springer, Heidelberg, Oct. / Nov. 2016. 4, 13

[24] T. Iwata. New blockcipher modes of operation with beyond the birthday bound security. In M. J. B. Robshaw, editor, *FSE 2006*, volume 4047 of *LNCS*, pages 310–327. Springer, Heidelberg, Mar. 2006. 24

[25] T. Iwata, B. Mennink, and D. Vizár. CENC is optimally secure. Cryptology ePrint Archive, Report 2016/1087, 2016. `http://eprint.iacr.org/2016/1087`. 24

[26] J. Lee, A. Luykx, B. Mennink, and K. Minematsu. Connecting tweakable and multi-key blockcipher security. *Designs, Codes and Cryptography*, Mar 2017. 14

[27] M. Liskov, R. L. Rivest, and D. Wagner. Tweakable block ciphers. In M. Yung, editor, *CRYPTO 2002*, volume 2442 of *LNCS*, pages 31–46. Springer, Heidelberg, Aug. 2002. 3, 9

[28] M. Liskov, R. L. Rivest, and D. Wagner. Tweakable block ciphers. *Journal of Cryptology*, 24(3):588–613, July 2011. 3, 9, 14

[29] M. Luby and C. Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, 17(2), 1988. 31

[30] B. Morris and P. Rogaway. Sometimes-recurse shuffle - almost-random permutations in logarithmic expected time. In P. Q. Nguyen and E. Oswald, editors, *EUROCRYPT 2014*, volume 8441 of *LNCS*, pages 311–326. Springer, Heidelberg, May 2014. 7, 21

[31] J. Patarin. The “coefficients H” technique (invited talk). In R. M. Avanzi, L. Keliher, and F. Sica, editors, *SAC 2008*, volume 5381 of *LNCS*, pages 328–345. Springer, Heidelberg, Aug. 2009. 41

[32] J. Patarin. Introduction to mirror theory: Analysis of systems of linear equalities and linear non equalities for cryptography. Cryptology ePrint Archive, Report 2010/287, 2010. `http://eprint.iacr.org/2010/287`. 24

[33] M. Raab and A. Steger. “Balls into bins” – a simple and tight analysis. In *RANDOM 1998*, pages 159–170. Springer, 1998. 34, 42

[34] T. Ristenpart and S. Yilek. The mix-and-cut shuffle: Small-domain encryption secure against N queries. In R. Canetti and J. A. Garay, editors, *CRYPTO 2013, Part I*, volume 8042 of *LNCS*, pages 392–409. Springer, Heidelberg, Aug. 2013. 7

[35] P. Rogaway and M. Bellare. Robust computational secret sharing and a unified account of classical secret-sharing goals. In P. Ning, S. D. C. di Vimercati, and P. F. Syverson, editors, *ACM CCS 07*, pages 172–184. ACM Press, Oct. 2007. 16

[36] A. Shamir. Identity-based cryptosystems and signature schemes. In G. R. Blakley and D. Chaum, editors, *CRYPTO'84*, volume 196 of *LNCS*, pages 47–53. Springer, Heidelberg, Aug. 1984. 7

[37] T. Shrimpton and R. S. Terashima. Salvaging weak security bounds for blockcipher-based constructions. In J. H. Cheon and T. Takagi, editors, *ASIACRYPT 2016, Part I*, volume 10031 of *LNCS*, pages 429–454. Springer, Heidelberg, Dec. 2016. 4

[38] J. Vance. VAES3 scheme for FFX: An addendum to The FFX mode of operation for Format Preserving Encryption. Submission to NIST, May 2011. 6, 7

[39] J. Vance and M. Bellare. Delegatable Feistel-based Format Preserving Encryption mode. Submission to NIST, Nov 2015. 3, 6, 7, 24, 36

[40] P. Zhang and H. Hu. On the provable security of the tweakable even-mansour cipher against multi-key and related-key attacks. Cryptology ePrint Archive, Report 2016/1172, 2016. `http://eprint.iacr.org/2016/1172`. 14

A Proof of Lemma 7.2

In this section, we will give the proof of Lemma 7.2 via the H-coefficient technique of Patarin [31, 14]. The proof also uses some technical balls-into-bins results. Therefore, we begin by reviewing the H-coefficient technique, and then describing the balls-into-bins lemmas.

H-COEFFICIENT TECHNIQUE. Let $\mathcal{A}$ be a deterministic, computationally unbounded adversary that tries to distinguish two games $\mathbf{G}_{\text{real}}$ and $\mathbf{G}_{\text{ideal}}$. Let $\mathcal{T}_{\text{real}}$ and $\mathcal{T}_{\text{ideal}}$ be the random variables for the transcript that records everything that the adversary is able to observe during its interaction with $\mathbf{G}_{\text{real}}$ and $\mathbf{G}_{\text{ideal}}$, respectively. We call a transcript τ *valid* if $\Pr[\mathcal{T}_{\text{ideal}} = \tau] > 0$. Partition the set of valid transcripts into Γ_{good} and Γ_{bad}; we refer to them as the set of good and bad transcripts, respectively. The following result bounds the distinguishing advantage of $\mathcal{A}$.

Lemma A.1 *[31, 14] Let $\epsilon, \delta \in [0, 1]$ be such that*

(a) $\Pr[\mathcal{T}_{\text{ideal}} \in \Gamma_{\text{bad}}] \le \epsilon$*, and*

(b) $1 - \frac{\Pr[\mathcal{T}_{\text{real}}=\tau]}{\Pr[\mathcal{T}_{\text{ideal}}=\tau]} \le \delta$ *for every* $\tau \in \Gamma_{\text{good}}$*.*

Then

$$\left|\Pr[\mathbf{G}_{\text{real}}(\mathcal{A}) \Rightarrow 1] - \Pr[\mathbf{G}_{\text{ideal}}(\mathcal{A}) \Rightarrow 1]\right| \le \epsilon + \delta \ .$$

Some balls-into-bins technical results. Consider the following game. Let k and q be integers, and let S be a finite set. An adversary $\mathcal{A}$ is given oracle access to $Q \leftarrow\!\!\$\ \mathbf{IC}(S, \{0,1\}^k)$ and can make at most q queries. If the adversary makes $Q(i, Y, -)$ to get answer X, it's not allowed to call $Q(i, X, +)$. Each query $Q(i, X, +)$ is considered throwing a ball into 2^k possible bins that correspond to the 2^k possible answers. Of course the throws are not independent, since the two balls for $Q(i, 0^k, +)$ and $Q(i, 1^k, +)$ must land in different bins. The goal of the adversary is to maximize the load of the heaviest bin. Let $\mathrm{Balls}(A, q, S, k)$ denote the random variable for the number of balls in the heaviest bin in this game. The following result gives a strong concentration bound on $\mathrm{Balls}(A, q, S, k)$.

Lemma A.2 *Let $k, q \in \mathbb{N}$ such that $k \geq 128$ and $q \leq 2^{k-1}$ and let S be a finite set. Then for any adversary $\mathcal{A}$,*

$$\Pr\Big[\mathrm{Balls}(A, q, S, k) \geq \frac{3.5k}{\log_2(k)}\Big] \leq 2^{1-k} \ .$$

To justify Lemma A.2, we'll need the following technical result of Dai and Steinberger [16] that reduces the problem above to the classic balls-into-bins setting.

Lemma A.3 ([16]) *Let $k, q \in \mathbb{N}$ such that $q \leq 2^{k-1}$ and let S be a finite set. Let X be the random variable for the number of balls in the heaviest bin when one throws q balls uniformly to 2^{k-1} bins. Then for any adversary $\mathcal{A}$ and any $s \in \mathbb{N}$,*

$$\Pr\Big[\mathrm{Balls}(A, q, S, k) \geq s] \leq \Pr[X \geq s] \ .$$

We then need the following classic balls-into-bins result.

Lemma A.4 ([33]) *Let $m, q \in \mathbb{N}$ such that $q \leq 2^m$. Let X be the random variable for the number of balls in the heaviest bin when one throws q balls uniformly to 2^m bins. Then $\Pr[X \geq 3m/\log_2(m)] \leq 2^{-m}$.*

Combining Lemma A.3 and Lemma A.4 leads to Lemma A.2; the proof is given below.

Proof of Lemma A.2: Let X be the random variable for the number of balls in the heaviest bin when one throws q balls uniformly to 2^{k-1} bins. From Lemma A.3,

$$\Pr\Big[\mathrm{Balls}(A, q, S, k) \geq \frac{3.5k}{\log_2(k)}\Big] \quad \leq \quad \Pr\Big[X \geq \frac{3.5k}{\log_2(k)}\Big].$$

On the other hand,

$$\frac{3.5k}{\log_2(k)} \geq \frac{3.5(k-1)}{\log_2(2k-2)} = \frac{3.5(k-1)}{\log_2(k-1)+1} \geq \frac{3(k-1)}{\log_2(k-1)},$$

where the last inequality is due to the hypothesis that $k \geq 128$. Then using Lemma A.4 with $m = k-1$,

$$\Pr\Big[X \geq \frac{3.5k}{\log_2(k)}\Big] \leq \Pr\Big[X \geq \frac{3(k-1)}{\log_2(k-1)}\Big] \leq 2^{1-k} \ .$$

```
Game $\mathbf{G}_{\mathrm{ideal}}(\mathcal{B})$
$K \leftarrow_{\$} \{0,1\}^k$; $\mathrm{XI} \leftarrow \emptyset$
$\mathrm{ChI} \leftarrow \emptyset$; $\pi \leftarrow_{\$} \mathrm{Perm}(\{0,1\}^k)$
For $I \in \mathsf{KDF.IS}$ do $\Pi(I,\cdot) \leftarrow_{\$} \mathrm{Perm}(\{0,1\}^k)$
$b' \leftarrow_{\$} \mathcal{B}^{\textsc{Enc},\textsc{Exp},\textsc{Ch},P}$; Return $(b' = 0)$

$\textsc{Enc}(I,\varepsilon,X)$
$Y \leftarrow \Pi(I,X)$; $L_0 \leftarrow \pi(\mathsf{M}_0(I))$; $L_1 \leftarrow \pi(\mathsf{M}_1(I))$
$\mathrm{EM}[L_0, X \oplus L_1] \leftarrow Y$; $\mathrm{DM}[L_0, Y] \leftarrow X \oplus L_1$
Return $Y$

$\textsc{Exp}(I)$
If $I \in \mathrm{ChI}$ then return $\bot$
$L_0 \leftarrow \pi(\mathsf{M}_0(I))$; $L_1 \leftarrow \pi(\mathsf{M}_1(I))$
$\mathrm{XI} \leftarrow \mathrm{XI} \cup \{I\}$; $\mathrm{KT}[L_0] \leftarrow I$; Return $L_0 \,\|\, L_1$

$\textsc{Ch}(I)$
If $I \in \mathrm{XI}$ then return $\bot$
$\mathrm{ChI} \leftarrow \mathrm{ChI} \cup \{I\}$

$P(L,X,+)$
If $\mathrm{ET}[L,X] \neq \bot$ then return $\mathrm{ET}[L,X]$
$Y \leftarrow \mathrm{EM}[L,X]$
If $\mathrm{KT}[L] = \bot$ or $Y = \bot$ or $\mathrm{DT}[L,Y] \neq \bot$ then
    $Y \leftarrow_{\$} \{R : \mathrm{DT}[L,R] = \bot\}$
$\mathrm{ET}[L,X] \leftarrow Y$; $\mathrm{DT}[L,Y] \leftarrow X$; Return $Y$

$P(L,Y,-)$
If $\mathrm{DT}[L,Y] \neq \bot$ then return $\mathrm{DT}[L,Y]$
$X \leftarrow \mathrm{DM}[L,Y]$
If $\mathrm{KT}[L] = \bot$ or $X = \bot$ or $\mathrm{ET}[L,X] \neq \bot$ then
    $X \leftarrow_{\$} \{R : \mathrm{ET}[L,R] = \bot\}$
$\mathrm{ET}[L,X] \leftarrow Y$; $\mathrm{DT}[L,Y] \leftarrow X$; Return $X$
```

Figure 16: **Game $\mathbf{G}_{\mathrm{ideal}}$ in the proof of Lemma 7.2.**

Putting this all together,

$$\Pr\Big[\mathrm{Balls}(A,q,S,k) \geq \frac{3.5k}{\log_2(k)}\Big] \leq 2^{1-k}$$

as claimed. ▌

<u>MAIN PROOF.</u> Without loss of generality, assume that once the adversary queries $\textsc{Exp}(I)$ to get subkey J, it will not make further $\textsc{Enc}(I,\cdot,\cdot)$ queries. Assume that the adversary always outputs (I,J) such that I is in the challenge set ChI.

We will construct from $\mathcal{A}$ an adversary $\mathcal{B}$ that aims to distinguish the following two games $\mathbf{G}_{\mathrm{real}}$ and $\mathbf{G}_{\mathrm{ideal}}$. Game $\mathbf{G}_{\mathrm{real}}$ is $\mathbf{G}^{\mathsf{ib\text{-}prp}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{B})$ for challenge bit 1, but there is no $\textsc{Dec}$ oracle. Game $\mathbf{G}_{\mathrm{ideal}}$ is instead implemented as in Fig. 16. In this game, the $\textsc{Enc}$ answers and the subkeys are generated independent of the ideal primitive. After $\mathcal{B}$ obtains a subkey $J = L_0 \,\|\, J_1$ of an identity I from the $\textsc{Exp}$ oracle, the answers for the subsequent queries $(L_0, R, \cdot)$ to the ideal primitive are *programmed* to be random but still consistent to (a) the prior ideal-primitive queries, and (b) the prior $\textsc{Enc}$ queries on identity I. In some rare cases, the information in (b) is inconsistent with that in (a). If so, the answers will be consistent to just (a). Note that if $\mathcal{B}$ is non-adaptive then we do not need to program the ideal cipher.

The adversary $\mathcal{B}$ is constructed as follows. It runs $\mathcal{A}$. When the latter makes a query, the former uses its corresponding oracle to answer. When $\mathcal{A}$ outputs its answer (I,J), let $J = L_0 \,\|\, L_1$. Adversary $\mathcal{B}$ picks an arbitrary message M such that there is no prior query $\textsc{Enc}(I,\varepsilon,M)$, no prior query $P(L_0, M \oplus L_1, +)$ and no prior query $P(L_0, Y, -)$ whose answer is $M \oplus L_1$. Since $p + q \leq 2^{k-1} - 2$, there exists such an M. Adversary $\mathcal{B}$ then queries $\textsc{Enc}(I,\varepsilon,M)$, and also queries $(L_0, M \oplus L_1, +)$ to the ideal cipher. If the answers match then $\mathcal{B}$ returns 1; otherwise it returns 0. In the real game, the chance that $\mathcal{B}$ outputs 1 is at least $\Pr[\mathbf{G}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{A})]$. In the ideal game, since the subkey for I was not exposed, so the two answers above are independently generated. Moreover, as $p + q \leq 2^{k-1} - 2$, each of the two answers above has at least 2^{k-1} possible, equally likely choices. Hence the chance that the adversary $\mathcal{B}$ outputs 1 in the ideal game is at most 2^{1-k}. Thus the advantage of $\mathcal{B}$ in distinguishing $\mathbf{G}_{\mathrm{real}}$ and $\mathbf{G}_{\mathrm{ideal}}$ is at least $\mathbf{Adv}^{\mathsf{ib\text{-}kr\text{-}ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{A}) - 2^{1-k}$.

Note that $\mathcal{B}$ queries $\textsc{Exp}(I)$ to get subkey J, it will not make further $\textsc{Enc}(I, \cdot, \cdot)$ queries. Moreover, if $\mathcal{A}$ is non-adaptive then so is $\mathcal{B}$. Adversary $\mathcal{B}$ makes at most $q+1$ $\textsc{Enc}$ and $\textsc{Ch}$ queries, q_e $\textsc{Exp}$ queries, and $p + 1$ ideal-cipher queries, but the $\textsc{Enc}$ and $\textsc{Ch}$ queries involve at most q identities.

We now give an upper bound on the advantage of $\mathcal{B}$ in distinguishing the games $\mathbf{G}_{\mathrm{real}}$ and $\mathbf{G}_{\mathrm{ideal}}$. Since the adversary is computationally unbounded, without loss of generality, assume that it's deterministic. Assume that $\mathcal{B}$ always makes exactly q_e $\textsc{Exp}$ queries. Assume that the adversary doesn't make any redundant queries: it never repeats a prior query, and if it queries $P(L, X, +)$ to get Y, then it won't query $P(L, Y, -)$, and likewise, if it queries $P(L, Y, -)$ to get X then it won't query $P(L, X, +)$. or $P(J[1:k], \cdot, \cdot)$ queries.

Wlog, assume that if the adversary queries $\textsc{Exp}(I)$ to get J, then it will not later queries $P(J[1:k], \cdot, \cdot)$. The reason is that, if the $\textsc{Enc}(I, \varepsilon, \cdot)$ queries and the pre-exposure $P(J[1:k], \cdot, \cdot)$ queries give contradictory information, then the adversary can simply output 0 and wins. Otherwise, the answers will be random but consistent to the information in prior $\textsc{Enc}(I, \varepsilon, \cdot)$ and $P(J[1:k], \cdot, \cdot)$ queries in both games. Moreover, those answers are independent of the answers of $\textsc{Enc}(I', \varepsilon, \cdot)$ and $P(L', \cdot, \cdot)$, for any $I' \neq I$ and $L' \neq J[1:k]$. The adversary thus can sample those answers by itself instead of using the ideal cipher.

We now bound $\mathcal{B}$'s advantage via the H-coefficient technique. After the adversary finishes querying, we'll grant it the master key K and the subkeys $J_1, \ldots, J_\ell$ of all involved identities that are not exposed. We stress that the adversary is forbidden from making further queries after it receives the keys. This key revelation can only help the adversary. A transcript consists of the adversary's queries/answers and the keys (returned via $\textsc{Exp}$ queries or granted at the end). We say that a transcript is *bad* if one of the following properties holds:

(i) The master key K is the left half of some subkey.

(ii) There is some query $P(K, \cdot, \cdot)$ in τ.

(iii) There is a query $\textsc{Enc}(I, \varepsilon, X)$ such that the subkey J of I is not exposed by $\textsc{Exp}$, and there is another query $P(J[1:k], X \oplus J[k+1:2k], +)$.

(iv) There is a query $\textsc{Enc}(I, \varepsilon, X)$ such that the subkey J of I is not exposed by $\textsc{Exp}$, and there is another ideal-cipher query $(J[1:k], Y, -)$ whose answer is $X \oplus J[k+1:2k]$.

(v) There is a query $\textsc{Enc}(I, \varepsilon, X)$ of answer Y such that the subkey J of I is not exposed by $\textsc{Exp}$, and there is another query $P(J[1:k], Y, -)$. We say that this ideal-cipher query *hits* identity I.

If a valid transcript is not bad, we say that it's *good*. Let Γ_{good} and Γ_{bad} be the set of good and bad (valid) transcripts, respectively. We first bound the probability $\Pr[\mathcal{T}_{\mathrm{ideal}} \in \Gamma_{\mathrm{bad}}]$.

- First, the chance that $\mathcal{T}_{\mathrm{ideal}}$ satisfies property (i) is at most $(q + q_e)/2^k$, since in the ideal game, $K \leftarrow_{\$} \{0,1\}^k$ is independent of all other subkeys.

- Next, because $K \leftarrow_{\$} \{0,1\}^k$ is independent of whatever the adversary receives until it is granted K, the chance that $\mathcal{T}_{\mathrm{ideal}}$ satisfies property (ii) is at most $(p+1)/2^k$.

- Moreover, for each granted subkey J, before it is granted, there are at least $(2^k - 2q_e)(2^k - 2q_e - 1) \geq 2^{2k-1}$ choices for the pair $(J[1:k], J[k+1:2k])$ and those choices are equally likely. Hence the chance that $\mathcal{T}_{\mathrm{ideal}}$ satisfies properties (iii) or (iv) is at most $2q(p+1)/2^{2k}$.

- Finally, view each query $\textsc{Enc}(I, \varepsilon, X)$ as throwing a ball into 2^k possible bins. From Lemma A.2, with probability at least $1 - 2/2^k$, in $\mathcal{T}_{\text{ideal}}$, no bin contains more than $3.5k/\lg(k)$ balls. In other words, each query $P(L, Y, -)$ can target at most $3.5k/\lg(k)$ identities to hit. But for each such identity, before the adversary is granted its subkey, there are at least $2^k - 2q_e \geq \frac{7}{8 \cdot 2^k}$ choices for its left half, and all those choices are equally likely. Hence the chance that $\mathcal{T}_{\text{ideal}}$ satisfies (v) is at most $2/2^k + 4k(p+1)/2^k \log_2(k)$.

Summing up,

$$\Pr[\mathcal{T}_{\text{ideal}} \in \Gamma_{\text{bad}}] \leq \frac{4k(p+1)}{2^k \log_2(k)} + \frac{2(p+1)q}{2^{2k}} + \frac{q + q_e + p + 3}{2^k} \quad . \tag{22}$$

Next, fix an arbitrary valid good transcript τ. We claim that

$$1 - \frac{\Pr[\mathcal{T}_{\text{real}} = \tau]}{\Pr[\mathcal{T}_{\text{ideal}} = \tau]} \leq 0 \ . \tag{23}$$

From Eq. (22) and Eq. (23), using Lemma A.1 yields

$$\mathbf{Adv}^{\textsf{ib-prp}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{B}) \leq \frac{2q(p+1)}{2^{2k}} + \frac{4(p+1)k}{2^k \cdot \log_2(k)} + \frac{q + q_e + p + 3}{2^k},$$

and thus

$$\begin{aligned} \mathbf{Adv}^{\textsf{ib-kr-ti}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{A}) &\leq \mathbf{Adv}^{\textsf{ib-prp}}_{\overline{\mathsf{F}},\mathsf{KDF}}(\mathcal{B}) + \frac{2}{2^k} \\ &\leq \frac{2q(p+1)}{2^{2k}} + \frac{4(p+1)k}{2^k \cdot \log_2(k)} + \frac{q + q_e + p + 5}{2^k} \quad . \end{aligned}$$

To justify Eq. (23), consider an arbitrary good transcript τ, and suppose that according to τ, the involved, un-exposed identities are $I_1, \ldots, I_\ell$. Let $I_{\ell+1}, \ldots, I_{\ell+q_e}$ be the identities for the $\textsc{Exp}$ queries, and let $J_{\ell+1}, \ldots, J_{\ell+q_e}$ be the subkeys, respectively. Suppose that according to τ, there are q_i queries to $\textsc{Enc}(I_i, \varepsilon, \cdot)$, and p_i queries to $P(J_i[1:k], \cdot, \cdot)$. We call a query $P(L, \cdot, \cdot)$ *useless* if L is not the left half of some J_i, for $i \in \{1, \ldots, \ell\}$.

Since τ is valid and the adversary is deterministic, at each moment in the real game, as long as the queries/answers that the adversary has received is consistent with τ, the adversary has to make the next query according to τ. Hence the event $\mathcal{T}_{\text{real}} = \tau$ can be decomposed into the following events in the real game:

- $\mathsf{Key}_{\text{real}}$: The master key is K, and querying $P(K, \mathsf{M}_0(I_i), +)$ and $P(K, \mathsf{M}_1(I_i), +)$ return $J_i[1:k]$ and $J_i[k+1, 2k]$ respectively, for every $i \in \{1, \ldots, \ell + q_e\}$.
- $\mathsf{Qr}^0_{\text{real}}$: If we ask the useless queries in τ to P, we'll receive the answers as indicated in τ.
- $\mathsf{Qr}^i_{\text{real}}$, with $i \in \{1, \ldots, \ell\}$: If we query $\textsc{Enc}(I_i, \varepsilon, \cdot)$ and $P(J_i[1:k], \cdot, \cdot)$ according to τ, we'll get the answers as indicated in τ.

Since τ is good, the events $\mathsf{Qr}^0_{\text{real}}, \mathsf{Qr}^1_{\text{real}}, \ldots, \mathsf{Qr}^\ell_{\text{real}}$ are conditionally independent, given $\mathsf{Key}_{\text{real}}$. Moreover, $\mathsf{Qr}^0_{\text{real}}$ is independent of $\mathsf{Key}_{\text{real}}$. Hence

$$\begin{aligned} & \Pr[\mathcal{T}_{\text{real}} = \tau] \\ = \; & \Pr[\mathsf{Key}_{\text{real}}] \cdot \Pr[\mathsf{Qr}^0_{\text{real}} \mid \mathsf{Key}_{\text{real}}] \prod_{i=1}^{\ell} \Pr[\mathsf{Qr}^i_{\text{real}} \mid \mathsf{Key}_{\text{real}}] \\ = \; & \Pr[\mathsf{Key}_{\text{real}}] \cdot \Pr[\mathsf{Qr}^0_{\text{real}}] \prod_{i=1}^{\ell} \frac{1}{2^k \cdots (2^k - q_i - p_i + 1)} \quad . \end{aligned}$$

Likewise, the event $\mathcal{T}_{\text{ideal}} = \tau$ can be decomposed into the following events in the ideal game:

- $\mathsf{Key}_\$$: When we query $\textsc{Exp}(I_j)$, we'll get J_j, for $j \in \{\ell+1, \ldots, \ell+q_e\}$. Moreover, when we sample keys to grant them to the adversary, we get $K, J_1, \ldots, J_\ell$.
- $\mathsf{Qr}^0_\$$: If we ask the useless queries in τ to P, we'll receive the answers as indicated in τ.
- $\mathsf{Qr}^i_\$$, with $i \in \{1, \ldots, \ell\}$: If we make queries $\textsc{Enc}(I_i, \varepsilon, \cdot)$ and $P(J_i[1:k], \cdot, \cdot)$ according to τ, we'll get the answers as indicated in τ.

Note that $\mathsf{Key}_\$, \mathsf{Qr}^0_\$, \mathsf{Qr}^1_\$, \ldots, \mathsf{Qr}^\ell_\$$ are independent, and thus

$$\begin{aligned}
&\Pr[\mathcal{T}_{\text{ideal}} = \tau] \\
&= \Pr[\mathsf{Key}_\$] \cdot \Pr[\mathsf{Qr}^0_\$] \prod_{i=1}^{\ell} \Pr[\mathsf{Qr}^i_\$] \\
&= \Pr[\mathsf{Key}_\$] \cdot \Pr[\mathsf{Qr}^0_\$] \prod_{i=1}^{\ell} \frac{1}{2^k \cdots (2^k - q_i + 1) 2^k \cdots (2^k - p_i + 1)} \\
&\leq \quad \Pr[\mathsf{Key}_\$] \Pr[\mathsf{Qr}^0_\$] \prod_{i=1}^{\ell} \frac{1}{2^k \cdots (2^k - q_i - p_i + 1)} \quad .
\end{aligned}$$

On the other hand, $\Pr[\mathsf{Qr}^0_\$] = \Pr[\mathsf{Qr}^0_{\text{real}}]$, because both events only involve queries to P. Moreover,

$$\Pr[\mathsf{Key}_{\text{real}}] = \Pr[\mathsf{Key}_\$] = \frac{1}{2^k} \cdot \frac{1}{2^k \cdots (2^k - 2\ell - 2q_e + 1)} \quad .$$

Put all this together

$$\frac{\Pr[\mathcal{T}_{\text{real}} = \tau]}{\Pr[\mathcal{T}_{\text{ideal}} = \tau]} \geq 1 \quad .$$

B The FF2 and DFF IB-FPE schemes

FF2 and DFF scheme are conventionally presented as FPE schemes. Here we cast them as IB-FPE schemes $\mathsf{FF2} = (\mathsf{F}_{\mathsf{ff2}}, \mathsf{KDF}_{\mathsf{ff2}})$ and $\mathsf{DFF} = (\mathsf{FF}_{\mathsf{dff}}, \mathsf{KDF}_{\mathsf{dff}})$, respectively, by defining, in each case, the base FPE schemes and the key-derivation functions (cf. Fig. 17). This is done so that the original FF2 and DFF are recovered by viewing identities in the IB-FPE scheme as tweaks in the original scheme.

<u>Notation and conventions.</u> For simplicity, we consider a special case, corresponding to particular choices of parameters for the standard. Thus, we fix

- A radix rdx, in the range $2 \leq \mathsf{rdx} < 2^8$, for example $\mathsf{rdx} = 10$.
- An input length n that is an even integer in the range $2 \leq n \leq 30$, for example $n = 4$.
- Identity space $\mathsf{ID} = \bigcup_{i=0}^{13} \{0,1\}^{8i}$, meaning an identity is a string of at most 13 bytes.
- A number r of rounds in the range $1 \leq r < 2^8$, for example $r = 10$ in the current standards.

We define the alphabet $\Sigma_{\mathsf{rdx}} = \{0, 1, \ldots, \mathsf{rdx} - 1\}$. Members of Σ_{rdx} are referred to as digits. If $X \in \Sigma_2^*$ and $m = |X|$ then $\mathrm{StToNum}(X)$ is the integer representing X, namely

$$\mathrm{StToNum}(X) \;=\; \sum_{i=0}^{m-1} X[m-i] \cdot \mathsf{rdx}^i \;.$$

```
Algorithm $\mathsf{F_{ff2}.E}(J, \varepsilon, X)$
$u \leftarrow \lfloor n/2 \rfloor$ ; $v \leftarrow n - u$
$A \leftarrow X[1:u]$ ; $B \leftarrow X[u+1:n]$
For $i = 0$ to $r - 1$ do
   $Q \leftarrow [i]^1 \| [B]^{15}$
   $Y \leftarrow E(J, Q)$ ; $y \leftarrow \mathrm{StToNum}(Y)$
   If ($i$ is even) then $m \leftarrow u$ else $m \leftarrow v$
   $c \leftarrow (\mathrm{StToNum}(A) + y) \bmod \mathsf{rdx}^m$
   $C \leftarrow \mathrm{NumToSt}^m(c)$
   $A \leftarrow B$ ; $B \leftarrow C$
Return $A \| B$

Algorithm $\mathsf{F_{ff2}.D}(J, \varepsilon, Z)$
$u \leftarrow \lfloor n/2 \rfloor$ ; $v \leftarrow n - u$
$A \leftarrow Z[1:u]$ ; $B \leftarrow Z[u+1:n]$
For $i = r - 1$ downto 0 do
   $Q \leftarrow [i]^1 \| [B]^{15}$
   $Y \leftarrow E(J, Q)$ ; $y \leftarrow \mathrm{StToNum}(Y)$
   If ($i$ is even) then $m \leftarrow u$ else $m \leftarrow v$
   $c \leftarrow (\mathrm{StToNum}(A) - y) \bmod \mathsf{rdx}^m$
   $C \leftarrow \mathrm{NumToSt}^m(c)$
   $A \leftarrow B$ ; $B \leftarrow C$
Return $A \| B$

Algorithm $\mathsf{KDF_{ff2}}(K, I)$
$t \leftarrow |I|$
$P \leftarrow [\mathsf{rdx}]^1 \| [t]^1 \| [n]^1 \| [I]^{13}$
$J \leftarrow E(K, P)$
Return $J$
```

```
Algorithm $\mathsf{FF_{dff}.E}(J_1 \| J_2, \varepsilon, X)$
$u \leftarrow \lfloor n/2 \rfloor$ ; $v \leftarrow n - u$
$A \leftarrow X[1:u]$ ; $B \leftarrow X[u+1:n]$
For $i = 0$ to $r - 1$ do
   $Q \leftarrow [i]^1 \| [B]^{15}$
   $Y \leftarrow E(J_1, Q \oplus J_2)$ ; $y \leftarrow \mathrm{StToNum}(Y)$
   If ($i$ is even) then $m \leftarrow u$ else $m \leftarrow v$
   $c \leftarrow (\mathrm{StToNum}(A) + y) \bmod \mathsf{rdx}^m$
   $C \leftarrow \mathrm{NumToSt}^m(c)$
   $A \leftarrow B$ ; $B \leftarrow C$
Return $A \| B$

Algorithm $\mathsf{FF_{dff}.D}(J_1 \| J_2, \varepsilon, Z)$
$u \leftarrow \lfloor n/2 \rfloor$ ; $v \leftarrow n - u$
$A \leftarrow Z[1:u]$ ; $B \leftarrow Z[u+1:n]$
For $i = r - 1$ downto 0 do
   $Q \leftarrow [i]^1 \| [B]^{15}$
   $Y \leftarrow E(J_1, Q \oplus J_2)$ ; $y \leftarrow \mathrm{StToNum}(Y)$
   If ($i$ is even) then $m \leftarrow u$ else $m \leftarrow v$
   $c \leftarrow (\mathrm{StToNum}(A) - y) \bmod \mathsf{rdx}^m$
   $C \leftarrow \mathrm{NumToSt}^m(c)$
   $A \leftarrow B$ ; $B \leftarrow C$
Return $A \| B$

Algorithm $\mathsf{KDF_{dff}}(K, I)$
$t \leftarrow |I|$
$P_1 \leftarrow [\mathsf{rdx}]^1 \| [t]^1 \| [n]^1 \| [I]^{13}$
$P_2 \leftarrow [0]^1 \| [0]^1 \| [0]^1 \| [I]^{13}$
$J_1 \leftarrow E(K, P_1)$ ; $J_2 \leftarrow E(K, P_2)$
Return $J_1 \| J_2$
```

Figure 17: **The $\mathsf{FF2} = (\mathsf{F_{ff2}}, \mathsf{KDF_{ff2}})$ (left) and $\mathsf{DFF} = (\mathsf{FF_{dff}}, \mathsf{KDF_{dff}})$ (right) IB-FPE schemes.**

By convention, $\mathrm{StToNum}(X) = 0$ if $X = \varepsilon$ is the empty string, meaning the string of length 0. If $0 \leq c < \mathsf{rdx}^m$ is an integer then $\mathrm{NumToSt}^m(c)$ is the string X in Σ^m that represents c, namely such that $\mathrm{StToNum}(X) = c$. If $b \geq 1$ is an integer, then $[x]^b$ denotes the representation of x as a b-byte string. This notation is used both for x an integer in the range $0, \ldots, 2^{8b} - 1$ —for example, $[2]^1 = 00000010$— and for x a binary string of length at most $8b$ —for example, $[101]^1 = 00000101$. By $E : \{0,1\}^{128} \times \{0,1\}^k \rightarrow \{0,1\}^{128}$ we denote the AES blockcipher.

FF2. We describe IB-FPE scheme $\mathsf{FF2} = (\mathsf{F_{ff2}}, \mathsf{KDF_{ff2}})$. The encryption function $\mathsf{F_{ff2}.E} : \{0,1\}^{128} \times \{\varepsilon\} \times \Sigma_2^n \rightarrow \Sigma_2^n$ of the base FPE scheme is a Feistel network, shown at the top left of Fig. 17. Shown right below it is the decryption function $\mathsf{F_{ff2}.D} : \{0,1\}^{128} \times \{\varepsilon\} \times \Sigma_2^n \rightarrow \Sigma_2^n$ of the base FPE scheme. Note the base FPE scheme has trivial tweak space $\{\varepsilon\}$. Below that is the key derivation function $\mathsf{KDF_{ff2}} : \{0,1\}^{128} \times \mathsf{ID} \rightarrow \{0,1\}^{128}$.

DFF. We describe IB-FPE scheme $\mathsf{DFF} = (\mathsf{F}, \mathsf{KDF_{dff}})$. The encryption function $\mathsf{FF_{dff}.E} : \{0,1\}^{256} \times \{\varepsilon\} \times \Sigma_2^n \rightarrow \Sigma_2^n$ of the base FPE scheme is a Feistel network, shown at the top right of Fig. 17. Note that the key length is 256, not 128. Shown right below it is the decryption function $\mathsf{FF_{dff}.D}$:

$\{0,1\}^{256} \times \{\varepsilon\} \times \Sigma_2^n \to \Sigma_2^n$ of the base FPE scheme. Note the base FPE scheme has trivial tweak space $\{\varepsilon\}$. Below that is the key derivation functions $\mathsf{KDF}_{\mathsf{dff}} : \{0,1\}^{128} \times \mathsf{ID} \to \{0,1\}^{256}$. Note the master key length is 128, not 256.