



# Multi-Instance Secure Public-Key Encryption <sup>★</sup>

Carlo Brunetta<sup>1</sup>, Hans Heum<sup>2</sup>, and Martijn Stam<sup>1</sup>

<sup>1</sup> Simula UiB, Bergen, Norway.

`carlo, martijn@simula.no`

<sup>2</sup> Department of Mathematical Sciences, NTNU - Norwegian University of Science and Technology, Trondheim, Norway.

`hans.heum@ntnu.no`<sup>\*\*</sup>

**Abstract.** Mass surveillance targets many users at the same time with the goal of learning as much as possible. Intuitively, breaking many users' cryptography simultaneously should be at least as hard as that of only breaking a single one, but ideally security degradation is gradual: an adversary ought to work harder to break more. Bellare, Ristenpart and Tessaro (Crypto'12) introduced the notion of multi-instance security to capture the related concept for password hashing with salts. Auerbach, Giacon and Kiltz (Eurocrypt'20) motivated the study of public key encryption (PKE) in the multi-instance setting, yet their technical results are exclusively stated in terms of key encapsulation mechanisms (KEMs), leaving a considerable gap.

We investigate the multi-instance security of public key encryption. Our contributions are twofold. Firstly, we define and compare possible security notions for multi-instance PKE, where we include PKE schemes whose correctness is not perfect. Secondly, we observe that, in general, a hybrid encryption scheme of a multi-instance secure KEM and an arbitrary data encapsulation mechanism (DEM) is unlikely to inherit the KEM's multi-instance security. Yet, we show how with a suitable information-theoretic DEM, and a computationally secure key derivation function if need be, inheritance is possible. As far as we are aware, ours is the first inheritance result in the challenging multi-bit scenario.

**Keywords:** Multi-Instance Security · Hybrid Encryption · Property Inheritance · Mass Surveillance

## 1 Introduction

Security of cryptographic schemes is increasingly studied concretely. The question changes from whether a scheme is secure or not, to how secure it is. The change in emphasis also results in increased importance in more realistic security notions that model a world where an adversary might have many potential

---

<sup>\*</sup> ©IACR 2023. This article is the final version submitted by the authors to the IACR and to Springer Nature on 15 February 2023.

<sup>\*\*</sup> Work by Hans Heum performed as part of his PhD studies at Simula UiB.

targets. If an adversary simply tries to learn something about one of its  $\kappa$  targets, then intuitively the more targets there are, the easier the adversary’s job becomes. Indeed, using simple hybrid arguments results in a security degradation that is linear in  $\kappa$ . But what happens if the adversary is greedy and wants to learn more, maybe even targets everyone? On the one hand, one could argue that if breaking one instance is hard, then so is breaking many. Yet, on the other hand, one would hope that breaking multiple instances, say  $n$ , is strictly harder than breaking just a single one.

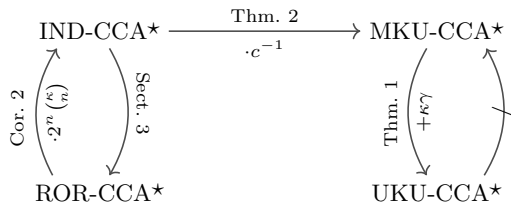
This second perspective made Bellare, Ristenpart and Tessaro [12], henceforth BRT, realize that new security notions are needed to reason about such greedy adversaries. They were motivated by how salts in password hashing protect against attackers re-using precomputation to retrieve multiple passwords. For their study into probabilistic symmetric schemes, they identified left-or-right indistinguishability under xor as the strongest notion. Roughly speaking, there are  $\kappa$  keys in the system each associated with its own left-or-right challenge bit  $b_i$  and the goal of the adversary is to guess the xor of all those bits.

Recently, Auerbach, Giacon and Kiltz [4], henceforth AGK, argued the importance of BRT’s concept to protect against mass surveillance. They introduced the  $(n, \kappa)$  scaling factor as the effort to break  $n$  out of  $\kappa$  instances relative to the effort needed to break a single instance. After recalling several well-known greedy attacks against public key schemes with dubious scaling factors, they set out to provide an encryption scheme with good, non-trivial scaling factor.

They discussed various versions of Hashed ElGamal that differed in whether users shared group parameters and/or generators, plus whether the underlying group was elliptic curve or finite field based. In the programmable random oracle model, they showed that the multi-instance security of Hashed ElGamal tightly relates to a novel multi-instance Gap Computational Diffie–Hellman (MI-GapCDH) assumption, whose validity was further supported by an analysis in the generic group model.

There was, or rather is, just one small problem: Hashed ElGamal is a key encapsulation mechanism (KEM), not a public key encryption (PKE) scheme. Indeed, although AGK use PKE as their motivation, their formalization is entirely centred around KEMs. Of course, Cramer and Shoup [19] already showed how a secure KEM can be combined with a secure data encapsulation mechanism (DEM) to create a secure PKE (for various notions of security). This so-called hybrid encryption paradigm is widely deployed in the real world, yet, can its composition theorem be easily lifted to the multi-instance setting?

For key unrecoverability, all seems fine, but for indistinguishability one quickly uncovers various challenges. Consider an adversary  $\mathbb{A}$  that wants to recover  $n$  out of  $\kappa$  challenge bits  $b_i$ : it can attempt to recover roughly half of its  $b_i$  by somehow breaking the DEM, and recovering the remaining half by breaking the KEM. Intuitively, such a divide-and-conquer strategy essentially rules out inheriting full multi-instance security of both KEM and DEM simultaneously. Instead, perhaps we should aim to bound an adversary’s multi-instance advantage against the hy-



**Fig. 1.** An overview of multi-instance security notions for public-key encryption, where  $\gamma$  relates to imperfect correctness (Def. 1), and the loss factor  $c$  is explained in Thm. 2.

brid encryption in terms of either breaking the full multi-instance security of the KEM or breaking only one of many instances of the DEM.

Special care would have to be taken to ensure that the corresponding multi-user DEM advantage is not overwhelming the multi-instance KEM advantage. After all, already when showing multi-user security of hybrid encryption, ensuring the DEM advantage does not overshadow the multi-user KEM advantage is challenging [23]. Furthermore, the study of multi-user KEMs highlights a second, more technical problem.

For multi-user security, there are essentially two different formalizations possible: one where each user comes with its own challenge bit and one where the users share a global challenge bit. Jager et al. [29] recently observed that only the latter lends itself to an easy adaptation of composition theorems using KEMs, as it allows a simple game-hop where all KEM-derived ephemeral keys are replaced by randomly selected keys (decoupled from the KEM encapsulations). That proof technique fails when there are multiple challenge bits. Unfortunately, for multi-instance security, the only option available is a notion with multiple challenge bits. In such a setting, inheritance of security properties of the KEM to any construction based on the KEM is an open problem.

**Our Contribution.** As mentioned above, multi-instance security was introduced by BRT in the context of probabilistic symmetric primitives and later adapted to key encapsulation mechanisms by AGK, who provide an excellent motivation for the study of multi-instance security in a public key setting. We adapt those notions to multi-instance security for PKE schemes, but make a number of non-trivial changes in the process. Firstly, we observe that the mechanisms used by BRT and AGK to model multi-instance games differ, which seems to have gone unnoticed hitherto. BRT’s mechanism is stronger as it allows for corruptions (denoted by  $\star$ ), yet AGK’s mechanism is more expressive by making explicit how many instances an adversary should break. We use elements of both in our notions, incorporating both BRT’s corruptions and AGK’s explicit expression of the number of targeted instances. Secondly, we allow for correctness to be imperfect, which has ramifications for how to deal with decryption oracles (for chosen-ciphertext attacks) and corruptions. We delve into the differences

between the various mechanisms in Sect. 3.3, furthermore we use our revised mechanism to study a number of related notions, as summarized in Fig. 1.

In more detail, we start out by porting BRT’s notion of key unrecoverability to the public-key setting. In fact, we consider two distinct versions of key unrecoverability: “Universal Key Unrecoverability” (UKU), where the adversary is tasked to recover the exact challenge private key(s) and “Matching Key Unrecoverability” (MKU), where it suffices to recover suitably equivalent private keys, where we leverage our imperfect correctness notion to define “suitably equivalent”. As one would expect, this relaxed key unrecoverability notion implies the stronger, exact notion up to a small loss related to how we model imperfect correctness (Thm. 1).

For our main notion of multi-instance security, we follow BRT’s identification of left-or-right xor-indistinguishability as the strongest notion and adapt it to the public key setting. As for the symmetric encryption setting studied by BRT, this indistinguishability notion implies the above key unrecoverability notions (Thm. 2); however, the differences between perfect symmetric encryption and imperfect PKE affect the corresponding implications and their proofs.

Finally, we explore an alternative notion, namely real-or-random xor-indistinguishability (ROR). Trivially, left-or-right tightly implies real-or-random and in the multi-instance setting BRT showed that the usual factor-2 loss from the single instance implication between real-or-random to left-or right, becomes an exponential factor- $2^\kappa$  loss. A similar loss is possible in our setting, however, we can also achieve a typically preferable bound of  $\binom{\kappa}{n}2^n$  (Cor. 2).

With suitable notions for multi-instance PKE available, we focus on how to turn a suitably multi-instance secure KEM into a multi-instance secure PKE scheme using hybrid encryption. For key unrecoverability, inheritance is immediate, yet we would like to guarantee good multi-instance indistinguishability (the left-hand branch of Fig. 1). We summarize our findings in Fig. 2.

Our first observation is that we can expand the length of the ephemeral key to any desired length using a pseudorandom extendable output function (XOF). The resulting extendable KEM, or XEM, inherits the multi-instance security of the underlying KEM, provided the XOF is secure against multi-challenge adversaries (Thm. 5). To ensure that the XOF does not become the weakest link, its seed will need to be long enough, which in turn implies that the underlying KEM already needs to output a sufficiently long ephemeral key.

The XOF above of course plays the role of key derivation function, but it is more common that it is modelled as part of any key expansion done by the DEM. Moving it into the KEM allows us to use an information-theoretic DEM, read one-time pad (OTP), irrespective of the message length. The OTP’s properties enable a simplified proof for the security of hybrid encryption (Thm. 6), where the PKE does indeed inherit the multi-instance security of the XEM, with two important caveats. Firstly, the OTP is only passively secure, so the PKE only achieves CPA not CCA security, and secondly, standard KEM indistinguishability only tightly provides real-or-random indistinguishability for the PKE (see the top line of Fig. 2).

Switching to the TagKEM framework [2], or in our case TagXEM, takes care of the first shortcoming and tightly achieves multi-instance ROR-CCA secure PKE, or IND-CCA non-tightly (Thm. 7). For the PKE to inherit multi-instance IND-CCA security tightly, we introduce a novel KEM indistinguishability notion that more closely matches PKE’s left-or-right idea, namely real-or-permuted (ROP). Finally, we can show tight multi-instance inheritance for the most desirable PKE notion, based on a ROP-secure TagXEM (Thm. 8).

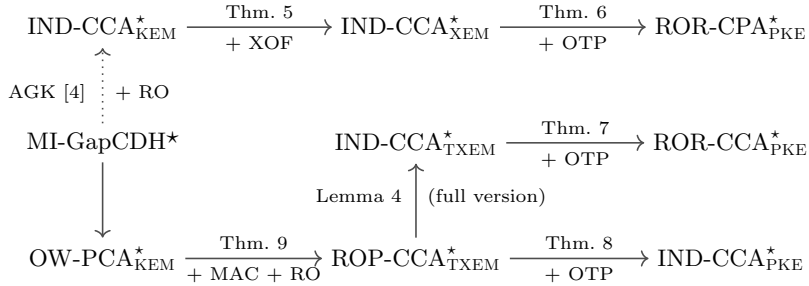
One small hiccup remains, as our KEM-to-XEM result unfortunately only works for classical KEM indistinguishability, not for ROP indistinguishability, nor does it look feasible to convert a KEM or XEM to a TagKEM or TagXEM, respectively, inheriting multi-instance security using standard reductions. Here, the random oracle, as used by AGK to prove their construction secure, comes to the rescue, although rather than looking at Hashed ElGamal under the MI-GapCDH assumption, we consider more general KEMs that are multi-instance one-way under plaintext checking attacks (unfortunately, also at this point we need to restrict to perfect correctness), which we combine with Abe et al.’s TagKEM construction from a KEM and a MAC (message authentication code).

Recalling that the original random oracle [14] was in fact a XOF, we can bake the extendability into the random oracle, including the key needed for an information-theoretic secure MAC. Moreover, the power of the ROM allows proving the stronger ROP indistinguishability just as easily as classical KEM indistinguishability. All in all, with Thm. 9 we achieve a suitably multi-instance secure TagXEM based on a KEM that can be instantiated by Hashed ElGamal. In that case, the security relies on the MI-GapCDH\* assumption, i.e. with corruptions. As an added benefit of using the random oracle, the resulting multi-instance bounds no longer rely on sufficiently long XOF inputs, thus for determining a suitable group size (when instantiating by Hashed ElGamal) the MI-GapCDH\* advantage is leading.

For low granularity, which corresponds to a setting where every user generates its own group as part of its public key, AGK’s technique can easily be extended to include corruptions and in the generic group model we arrive at the same bound for the hardness of MI-GapCDH\*, so with corruptions, as AGK did without corruptions. Unfortunately, for the more realistic high granularity setting, where users share the same (standardized) group, AGK’s proof strategy does not easily allow incorporating corruptions. We provide details in the full version [17].

Thus, we can conclude that XOF-based Hashed ElGamal combined with a suitable information-theoretically secure MAC and the one-time-pad, provides good multi-instance security in the programmable random oracle model and generic group model, provided that users each select their own independent group. We briefly touch upon a concrete interpretation in the full version, where we also informally address AGK’s scaling factor.

**Related Work.** Farshim and Tessaro [20] recently followed up BRT’s line of work on the multi-instance security of password hashing by combining it with the related preprocessing setting. AGK [4] motivated their investigation into



**Fig. 2.** An overview of our constructions achieving various flavours of multi-instance security. The left upwards arrow is dotted, as AGK did not consider corruptions.

multi-instance security by the threat of mass surveillance. The latter had previously motivated Bellare et al. [11] to consider subversion, namely the ease with which a “big brother” might subvert an encryption algorithm by replacing it surreptitiously with a trapdoored one with otherwise identical behaviour.

The multi-instance setting is closely related to the multi-user setting, in which the adversary is tasked with breaking only one rather than  $n$  out of  $\kappa$  possible instances. Multi-user security was introduced by Bellare et al. [7] in the public-key setting, with the goal of deriving concrete security parameters in a more realistic setting. There have been many recent follow-up works, including how the hybrid paradigm generalizes to the setting without corruptions [23], and later with corruptions [33], as well as the construction of tightly-secure authenticated key exchange (AKE) from multi-user KEMs [29]. Various versions of the multi-user GapCDH problem with corruptions were recently proposed and analysed in that context [30].

One definitional subtlety of multi-user security is the number of challenge bits: either a single one, as originally conceived, or many, as typical for the multi-instance setting. The various definitions do not appear to imply each other tightly [26], which slightly hinders regarding the multi-user setting as a special case of the multi-instance setting (due to potential tightness losses).

## 2 Preliminaries

### 2.1 Notation

For a positive integer  $n$ , we write  $[n]$  for the set  $\{1, \dots, n\}$ . We use code-based experiments, where  $\leftarrow$  denotes deterministic assignment and  $\leftarrow \$$  denotes probabilistic assignment. By convention, all sets and lists are initialized empty. For a set  $X$ , we use the shorthand  $X \stackrel{\cup}{\leftarrow} x$  for the operation  $X \leftarrow X \cup \{x\}$ . If  $X$  is a list, then  $X \stackrel{\leftarrow}{\leftarrow} x$  denotes appending the element  $x$  to  $X$ ; to retrieve the  $i$ th element of the list, we write  $X[i]$  where by convention  $X[i] = \emptyset$  for out-of-bounds  $i$ .

We use  $\Pr[\textit{Code} : \textit{Event} \mid \textit{Condition}]$  to denote the conditional probability of *Event* occurring when *Code* is executed, conditioned on *Condition*. We omit *Code* when it is clear from the context and *Condition* when it is not needed. For Boolean values, we use  $\{\text{true}, \text{false}\}$  and  $\{0, 1\}$  interchangeably, where by convention 1 corresponds to **true**.

When proving relations between notions and security of constructions, we will often refer to simple fully black box (SFBB) reductions. A reduction is fully black box iff it works for all schemes and adversaries, and only accesses them in a black box manner [6, 38] (we leave the black box dependence implicit in our notation). Moreover, if the reduction only runs its adversary once and without rewinding, then the reduction is simple [34].

Finally, the respective games that the adversary and the reduction are playing often have matching (though not identical) oracles; for instance, both may have access to a decryption oracle or a key corruption oracle. We call a reduction type-preserving with respect to, say, a decryption oracle iff the reduction will make decryption queries iff its black-box adversary makes decryption queries. Type-preservation, without explicit mention of any oracles, is implicitly meant to imply for all meaningfully matching oracles (unless otherwise specified).

Type-preservation of reductions appears folklore and can easily be established by inspection. Intuitively, a type-preserving reduction can be used to show simultaneously that CCA security of some kind implies CCA security of another kind and that CPA security of the same kind implies CPA security of the other kind. In Sect. 3.3 we will encounter several reductions that are only partially type-preserving.

## 2.2 PKE Syntax

A public-key encryption scheme PKE consists of four algorithms: the probabilistic key generation algorithm  $\text{PKE.Kg}$ , which takes as input some system parameter  $\text{pm}$  (see also Remark 1) and outputs a public/private key pair  $(\text{pk}, \text{sk})$ ; the deterministic key validation algorithm  $\text{PKE.Check}$ , which takes as input the system parameters  $\text{pm}$  as well as a purported public/private key pair  $(\text{pk}, \text{sk})$  and returns **true** or **false** (see Remark 2 below), the probabilistic encryption algorithm  $\text{PKE.Enc}$ , which on input a public key  $\text{pk}$  and a message  $m \in \mathcal{M}$  (see Remark 3), outputs a ciphertext  $c$ ; and the deterministic decryption algorithm  $\text{PKE.Dec}$ , which on input of a secret key  $\text{sk}$  and a ciphertext  $c$ , outputs either a message  $m$ , or a special symbol  $\perp$  denoting failure.

*Remark 1.* The system parameters  $\text{pm}$  are implicitly input to  $\text{PKE.Enc}$  and  $\text{PKE.Dec}$  as well; for concreteness, they can for instance be the description of an elliptic curve group with generator for an ECDLP-based system or the dimensions and noise sampling algorithm for an LWE-based system. When one is interested in re-phrasing our results in an asymptotic setting, the parameters  $\text{pm}$  will be generated by a probabilistic, polynomial-time algorithm that only takes the security parameter as input.

*Remark 2.* For various modern cryptosystems, especially schemes targeting post-quantum security or tight multi-user security, the relationship between public and private keys is not one-to-one. For instance, a single public key can have various private keys [23] or a single private key can lead to various public keys [16]. Naively, one could check whether a public key and private key belong together by simply verifying whether encrypting and then decrypting a number of random messages always returns the original messages. With imperfect correctness, such a canonical checking algorithm can produce both false positives and false negatives. Yet, it is usually still possible to check whether a private–public key pair matches more directly, which we model by the key validation algorithm `PKE.Check`. We will define both correctness and key unrecoverability in terms of this key validation algorithm.

*Remark 3.* The message space  $\mathcal{M}$  may depend on the parameters  $\mathbf{pm}$ , but for simplicity we assume it independent of the public key  $\mathbf{pk}$ . Often  $\mathcal{M}$  consists of arbitrary length bitstrings, or at least all bitstrings up to some large length (e.g.  $2^{64}$ ) and messages of the same length are deemed equivalent as they are expected to yield ciphertexts of identical lengths. We will model these equivalences more abstractly by assuming that  $\mathbf{pm}$  implicitly defines a number  $\mathbf{m}$  of equivalence classes, together with an efficient method  $\llbracket \cdot \rrbracket : \mathcal{M} \rightarrow [\mathbf{m}]$  to determine the class (e.g. length) of a message and an efficient algorithm to sample uniformly from a given equivalence class. We write  $\sim$  for the equivalence, so for  $m \in \mathcal{M}$ ,  $m \sim m'$  iff  $\llbracket m \rrbracket = \llbracket m' \rrbracket$ .

**Correctness.** Perfect correctness states that for all parameters  $\mathbf{pm}$ , all key pairs  $(\mathbf{pk}, \mathbf{sk})$  that can be output by `PKE.Kg`( $\mathbf{pm}$ ), and all messages  $m \in \mathcal{M}$ , we always have that `PKE.Dec` <sub>$\mathbf{sk}$</sub> (`PKE.Enc` <sub>$\mathbf{pk}$</sub> ( $m$ )) =  $m$ . Yet modern schemes, especially lattice-based ones, often allow a small decryption error, where occasionally decryption will fail or it will return a wrong message.

Various relaxations of correctness have appeared in the literature in order to argue about such schemes as it turns out that some classical results implicitly or subtly relied on perfect correctness. In order for our work to be meaningful for a large range of both classical and modern schemes, we introduce a stronger version of imperfect correctness based on the key validation algorithm.

**Definition 1 (( $\gamma, \delta$ )-Correctness).** *Let  $\gamma, \delta \in [0, 1]$ . Then a public-key encryption scheme `PKE` is called ( $\gamma, \delta$ )-correct iff for all  $\mathbf{pm}$ ,*

1.  $\Pr[(\mathbf{pk}, \mathbf{sk}) \leftarrow \text{PKE.Kg}(\mathbf{pm}) : \text{PKE.Check}(\mathbf{pm}, \mathbf{pk}, \mathbf{sk}) = \text{false}] \leq \gamma$ ;
2. *for all  $(\mathbf{pk}, \mathbf{sk})$  and all  $m \in \mathcal{M}$ , if  $\text{PKE.Check}(\mathbf{pm}, \mathbf{pk}, \mathbf{sk}) = \text{true}$  then*

$$\Pr[\text{PKE.Dec}_{\mathbf{sk}}(\text{PKE.Enc}_{\mathbf{pk}}(m)) \neq m] \leq \delta .$$

Perfect correctness corresponds to  $(0, 0)$ -correctness and any scheme is trivially both  $(1, 0)$ -correct and  $(0, 1)$ -correct. For good schemes  $\gamma$  and  $\delta$  can simultaneously be chosen small, where typically increasing  $\gamma$  allows for decreasing  $\delta$ . As we will see, both  $\gamma$  and  $\delta$  will appear in various bounds, thus allowing larger  $\gamma$  to enable smaller  $\delta$  (or vice versa) might give preferable bounds.



### 3 Multi-Instance Security of Public-Key Encryption

#### 3.1 Two Flavours of Key Recovery

The minimal requirement for public-key encryption schemes is that, given a public key, it should be difficult to recover the private key. Although key unrecoverability is a very weak notion theoretically, its study has two main motivations: firstly, many multi-instance attacks target key recovery, and secondly, conceptually the notion is relatively simple, allowing both an instructive introduction of formalizing multi-instance security and an initial comparison between BRT’s perfect symmetric encryption and our imperfect public key encryption.

At first sight, the generalization to the multi-instance setting appears immediate: an adversary tries to recover the respective private keys for a number of public keys. BRT introduced universal key unrecoverability (UKU) as a suitable notion for multi-instance security of symmetric encryption. We provide an analogue for public-key encryption, but there are some crucial changes in the game’s mechanics (see also Sect. 3.3).

Let  $0 < n \leq \kappa$  be integer parameters, then the universal key unrecoverability experiment  $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-uku-cca}^\star}(\mathbb{A})$  for public-key encryption scheme PKE and adversary  $\mathbb{A}$  is described in Fig. 3. It generates  $\kappa$  key pairs and provides the public keys to  $\mathbb{A}$ , who is then tasked with recovering exactly  $n$  of the corresponding private keys.

The adversary has access to both a decryption oracle  $\mathcal{D}$  and a key corruption oracle  $\mathcal{K}$ , giving rise to chosen ciphertext attacks with corruptions (CCA $^\star$ ; the  $^\star$  denotes corruptions). The decryption oracle  $\mathcal{D}(i, c)$  takes as input an index  $i$  and a ciphertext  $c$ , and returns the output of the decryption algorithm  $\text{PKE.Dec}$  on input  $\text{sk}_i$  and  $c$ . The corruption oracle  $\mathcal{K}(i)$  simply takes as input a key index  $i$ , and returns the corresponding private key  $\text{sk}_i$ . The game notes that the key pair with index  $i$  has been corrupted by adding it to the global set  $K$ .

Eventually,  $\mathbb{A}$  outputs a set of key indices  $I$  and a list  $(\text{sk}_i)_{i \in I}$  of guesses of the private keys corresponding to those indices. In order for  $I$  to be eligible, it needs to have cardinality  $n$  without containing any corrupted key pairs, that is, the sets of guessed keys  $I$  and corrupted keys  $K$  should be disjoint. If  $I$  is eligible and every guessed private key matches the corresponding sampled one, the adversary wins the game. In that case, the game halts with output 1; otherwise, it halts with output 0. The advantage is the probability that the game outputs 1.

**Definition 2.** *Let PKE be a public-key encryption scheme. Then the universal key unrecoverability advantage of an adversary  $\mathbb{A}$  is*

$$\text{Adv}_{\text{PKE}}^{(n,\kappa)\text{-uku-cca}^\star}(\mathbb{A}) = \Pr \left[ \text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-uku-cca}^\star}(\mathbb{A}) = 1 \right],$$

where the experiment is defined in Fig. 3.

Weaker notions emerge by dropping either or both of the two oracles. Without key corruption, standard CCA security results. Without decryption oracle, chosen plaintext security (CPA $^\star$  resp. CPA) emerges. As usual, an encryption oracle is superfluous in the PKE setting.

Experiment $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-}(u/m)\text{ku-cca}^*}(\mathbb{A})$	Oracle $\mathcal{D}(i, c)$
$(\text{pk}_1, \text{sk}_1), \dots, (\text{pk}_\kappa, \text{sk}_\kappa) \leftarrow \text{PKE.Kg}$ $(\mathbb{I}, (\hat{\text{sk}}_i)_{i \in \mathbb{I}}) \leftarrow \text{A}^{\mathcal{D}, \mathcal{K}}(\text{pk}_1, \dots, \text{pk}_\kappa)$ <b>if</b> $ \mathbb{I}  \neq n \vee \mathbb{I} \cap \mathbb{K} \neq \emptyset$ <b>then return 0</b>	$m \leftarrow \text{PKE.Dec}_{\text{sk}_i}(c)$ <b>return</b> $m$
UKU : <b>return</b> $\bigwedge_{i \in \mathbb{I}} \text{sk}_i = \hat{\text{sk}}_i$	Oracle $\mathcal{K}(i)$
MKU : <b>return</b> $\bigwedge_{i \in \mathbb{I}} \text{PKE.Check}(\text{pk}_i, \hat{\text{sk}}_i)$	$\mathbb{K} \xleftarrow{\cup} i$ <b>return</b> $\text{sk}_i$

**Fig. 3.** The key recovery experiments  $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-uku-cca}^*}(\mathbb{A})$  and  $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-mku-cca}^*}(\mathbb{A})$ ; they only differ in their win condition.

For cryptosystems where a single public key may have many matching private keys (such as Cramer–Shoup [18]), universal key unrecoverability is rather weak. Hence, we consider a second, slightly stronger notion of key recovery, in which the recovered private keys are no longer required to be identical to those sampled in the game. Instead, it suffices that each passes the keypair checking algorithm  $\text{PKE.Check}$ ; here we leverage our correctness definition (Def. 1). We call the resulting notion *matching key unrecoverability* (MKU), whose game is included in Fig. 3. That MKU security indeed implies UKU security is captured by Thm. 1 below, where the error term  $\kappa\gamma$  results from the unique correct keys as output by the key generation not always passing the  $\text{PKE.Check}$  algorithm (see the full version for the proof).

**Theorem 1** (MKU  $\longrightarrow$  UKU). *Let  $0 < n \leq \kappa$  be integer parameters and let  $\text{PKE}$  be a  $(\gamma, \delta)$ -correct encryption scheme. Then, there is a type-preserving SFBB reduction  $\mathbb{B}_{\text{mku}}$ , such that for every adversary  $\mathbb{A}_{\text{uku}}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n,\kappa)\text{-uku-cca}^*}(\mathbb{A}_{\text{uku}}) \leq \text{Adv}_{\text{PKE}}^{(n,\kappa)\text{-mku-cca}^*}(\mathbb{B}_{\text{mku}}) + \kappa\gamma.$$

### 3.2 Left-or-Right XOR Indistinguishability

To capture a stronger notion of security than simply hardness of key recovery, BRT considered various generalizations of indistinguishability to the multi-instance setting. For perfect probabilistic symmetric encryption, they concluded that left-or-right xor-indistinguishability is the strongest notion. Here each key comes with its own challenge bit that determines the left-or-right nature of the corresponding challenge encryption oracle; the adversary is tasked to retrieve the xor of all the challenge bits. In Def. 3, we use our modified game mechanics to adapt left-or-right xor-indistinguishability for potentially non-perfect public-key encryption.

**Definition 3.** *Let  $\text{PKE}$  be a public-key encryption scheme. Then the xor-indistinguishability advantage of an adversary  $\mathbb{A}$  is*

$$\text{Adv}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A}) = 2 \cdot \Pr\left[\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A}) = 1\right] - 1,$$

Experiment $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A})$	Oracle $\mathcal{E}(i, m_0, m_1)$								
$(\text{pk}_1, \text{sk}_1), \dots, (\text{pk}_\kappa, \text{sk}_\kappa) \leftarrow_{\$} \text{PKE.Kg}$ $b_1, \dots, b_\kappa \leftarrow_{\$} \{0, 1\}$ $(\mathbb{I}, \hat{b}) \leftarrow_{\$} \mathbb{A}^{\mathcal{E}, \mathcal{D}, \mathcal{K}, \mathcal{B}}(\text{pk}_1, \dots, \text{pk}_\kappa)$ <b>if</b> $ \mathbb{I}  \neq n \vee \mathbb{I} \cap (\mathbb{K} \cup \mathbb{B}) \neq \emptyset$ <b>then</b> $\hat{b} \leftarrow 0$ <b>return</b> $\oplus_{i \in \mathbb{I}} b_i = \hat{b}$	<b>if</b> $m_0 \neq m_1$ <b>then return</b> $\perp$ $c \leftarrow_{\$} \text{PKE.Enc}_{\text{pk}_i}(m_{b_i})$ $\mathbf{M}_i(c) \leftarrow m_{b_i}$ $\mathbb{C}_i \xleftarrow{\cup} c$ <b>return</b> $c$								
<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border-bottom: 1px solid black; padding-bottom: 5px;">           Oracle <math>\mathcal{K}(i)</math> </td> <td style="width: 50%; border-bottom: 1px solid black; padding-bottom: 5px;">           Oracle <math>\mathcal{B}(i)</math> </td> </tr> <tr> <td style="padding: 5px;"> <math>\mathbb{K} \xleftarrow{\cup} i</math> </td> <td style="padding: 5px;"> <math>\mathbb{B} \xleftarrow{\cup} i</math> </td> </tr> <tr> <td style="padding: 5px;"> <b>return</b> <math>\text{sk}_i</math> </td> <td style="padding: 5px;"> <b>return</b> <math>b_i</math> </td> </tr> </table>	Oracle $\mathcal{K}(i)$	Oracle $\mathcal{B}(i)$	$\mathbb{K} \xleftarrow{\cup} i$	$\mathbb{B} \xleftarrow{\cup} i$	<b>return</b> $\text{sk}_i$	<b>return</b> $b_i$	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="border-bottom: 1px solid black; padding-bottom: 5px;">           Oracle <math>\mathcal{D}(i, c)</math> </td> </tr> <tr> <td style="padding: 5px;"> <math>m \leftarrow \text{PKE.Dec}_{\text{sk}_i}(c)</math>  <b>if</b> <math>c \in \mathbb{C}_i \wedge m = \mathbf{M}_i(c)</math> <b>then return</b> <math>\perp</math>  <b>return</b> <math>m</math> </td> </tr> </table>	Oracle $\mathcal{D}(i, c)$	$m \leftarrow \text{PKE.Dec}_{\text{sk}_i}(c)$ <b>if</b> $c \in \mathbb{C}_i \wedge m = \mathbf{M}_i(c)$ <b>then return</b> $\perp$ <b>return</b> $m$
Oracle $\mathcal{K}(i)$	Oracle $\mathcal{B}(i)$								
$\mathbb{K} \xleftarrow{\cup} i$	$\mathbb{B} \xleftarrow{\cup} i$								
<b>return</b> $\text{sk}_i$	<b>return</b> $b_i$								
Oracle $\mathcal{D}(i, c)$									
$m \leftarrow \text{PKE.Dec}_{\text{sk}_i}(c)$ <b>if</b> $c \in \mathbb{C}_i \wedge m = \mathbf{M}_i(c)$ <b>then return</b> $\perp$ <b>return</b> $m$									

**Fig. 4.** Our main notion of multi-instance indistinguishability. In blue the slightly non-standard strengthening of the decryption oracle in case of imperfect correctness.

where the experiment is defined in Fig. 4.

In the experiment  $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A})$ , the adversary gets access to  $\kappa$  independently drawn public keys and helper oracles  $\mathcal{D}$  and  $\mathcal{K}$  (as described in Sect. 3.1). Furthermore,  $\mathbb{A}$  gets access to a challenge encryption oracle  $\mathcal{E}$  and a separate bit corruption oracle  $\mathcal{B}$ .

On input two equivalent messages  $m_0$  and  $m_1$  and a public key index  $i$ , the challenge encryption oracle returns  $\text{PKE.Enc}_{\text{pk}_i}(m_{b_i})$  where  $b_i$  is the challenge bit associated with the public key indexed by  $i$ . As usual for IND-CCA notions, challenge ciphertexts cannot be queried to the decryption oracle, which we catch on-the-fly [9]. Owing to the imperfect decryption, we allow a slight relaxation: if a challenge ciphertext decrypts incorrectly, we do not suppress the output and essentially allow the query. This relaxation strengthens the notion, but as challenge ciphertexts are honestly generated, the advantage gained by an adversary can be bound by the correctness parameters of the PKE using an identical-until-bad argument; however such a generic approach might not give bounds appropriate for the multi-instance setting.

Eventually, the adversary returns a set  $\mathbb{I}$  of targets and a guess  $\hat{b}$  of the xor of the corresponding challenge bits  $b_i$ . If  $\mathbb{I}$  is a set of  $n$  uncorrupted indices, then intuitively an adversary's uncertainty about any of the  $n$  challenge bits will be affected in the final guess  $\hat{b}$ , so in that sense  $\hat{b}$  neatly captures an adversary's need to break  $n$  instances in order to win. If  $\mathbb{I}$  is not a set of  $n$  uncorrupted indices, the game resets  $\mathbb{A}$ 's guess  $\hat{b}$  to 0, ensuring an adversary gains zero advantage from such a bad  $\mathbb{I}$ .

**The Relationship with Key Recovery.** BRT showed that in their perfect symmetric setting, multi-instance indistinguishability implies multi-instance universal key unrecoverability. While that may sound like a triviality, their proof [13,

App. C] was not entirely straightforward and, to ensure that the advantages carried over neatly, the distinguishing reduction receiving recovered keys needed to amplify its success probability by repeated random challenge encryptions. Their bound ends up with an additive term that corresponds to the likelihood that decrypting using an incorrect key results in the opposite message from the decrypted one.

Our imperfect public key setting is slightly different. On the one hand, the reduction can check the recovered keys with the `PKE.Check` algorithm, yet on the other hand correct keys can still cause incorrect decryptions. As a result, our amplification based on multiple challenge encryptions differs from BRT's, as we move from unanimity to a plurality vote. Furthermore, our reduction can use fixed messages (to match how correctness is defined), which reduces a dependency (in the bound) on the size of the message space. We suspect that our amplification can be tightened further by a combination of exploiting randomness and more fine-tuned voting, coupled with more fine-grained bounding of probabilities.

As is, the complexity of the bound makes its behaviour somewhat opaque and for some parameter choices vacuous (when  $c < 0$ ). The main idea is that  $\mathbb{B}_{\text{ind}}$  can increase  $q$ , the number of challenge encryptions per user, to counteract the losses inferred by large  $n$  and/or large  $\delta$ , with a small penalty to its running time. For  $\delta = 2^{-64}$ ,  $q = 1$  already suffices for  $c > 1/2$  for  $n < 2^{25}$ . In case of perfect correctness for keys that check out, corresponding to  $\delta = 0$ , the bound is completely tight.

**Theorem 2** (IND  $\rightarrow$  MKU). *Let PKE be a  $(\gamma, \delta)$ -correct encryption scheme with  $\delta < 1/2$ . Then there is a type-preserving SFBB reduction  $\mathbb{B}_{\text{ind}}$  such that, for every  $\mathbb{A}_{\text{mku}}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{B}_{\text{ind}}) \geq c \cdot \text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-mku-cca}^*}(\mathbb{A}_{\text{mku}}),$$

with  $c = 2 \left(1 - 2^q (\delta(1 - \delta))^{\frac{q}{2}}\right)^n - 1$  where  $q \in \mathbb{Z}_{>0}$  is an amplification parameter of the reduction;  $\mathbb{B}_{\text{ind}}$ 's overhead consists of  $q \cdot n$  calls to  $\mathcal{E}$ ,  $n$  offline key checks, and  $q \cdot n$  offline decryptions.

*Proof.* Let  $\mathbb{B}_{\text{ind}}$  run adversary  $\mathbb{A}_{\text{mku}}$  on the same  $\kappa$  public keys as it received itself. Whenever  $\mathbb{A}_{\text{mku}}$  makes a decryption or corruption query,  $\mathbb{B}_{\text{ind}}$  simply forwards the queries to its own oracle, relaying the response back to  $\mathbb{A}_{\text{mku}}$ . Eventually,  $\mathbb{A}_{\text{mku}}$  terminates with output  $(\mathbb{I}, (\hat{\mathbf{s}}\mathbf{k}_i)_{i \in \mathbb{I}})$  and  $\mathbb{B}_{\text{ind}}$  first confirms whether  $\mathbb{A}_{\text{mku}}$  won, by checking, for all the returned private keys, whether `PKE.Check`( $\text{pk}_i, \hat{\mathbf{s}}\mathbf{k}_i$ ) holds. If any check fails,  $\mathbb{B}_{\text{ind}}$  halts with output 0.

Let  $m_0$  and  $m_1$  be two distinct yet equivalent messages. Then for all  $i \in \mathbb{I}$ ,  $\mathbb{B}_{\text{ind}}$  creates a guess  $\hat{b}_i$  by querying its challenge encryption oracle  $q$  times on those two messages, so  $q$  queries  $\mathcal{E}(i, m_0, m_1)$  resulting in  $c_{ij}$ , for  $j \in [q]$ . It then decrypts those ciphertexts using the private key  $\hat{\mathbf{s}}\mathbf{k}_i$  it obtained from  $\mathbb{A}_{\text{mku}}$ , resulting in purported messages  $m_{ij} \leftarrow \text{PKE.Dec}_{\hat{\mathbf{s}}\mathbf{k}_i}(c_{ij})$ . If, for a fixed  $i$ , there are strictly more than  $q/2$  appearances of  $m_0$  amongst the  $m_{ij}$ , it sets  $\hat{b}_i$  to 0;

if there are strictly more than  $q/2$  appearances of  $m_1$ , then it sets  $\hat{b}_i$  to 1. If neither message appears more than  $q/2$  times,  $\mathbb{B}_{\text{ind}}$  halts with output 0. Once  $\mathbb{B}_{\text{ind}}$  has created a guess  $\hat{b}_i$  for all  $i \in \mathbb{I}$ , it terminates on output  $(\mathbb{I}, \bigoplus_{i \in \mathbb{I}} \hat{b}_i)$ .

For  $i \in \mathbb{I}$ , let  $\text{Check}_i$  be the event that  $\mathbb{A}_{\text{mku}}$  outputs a key  $\hat{s}k_i$  that passes the test and let  $\text{Good}_i$  be the event that  $\mathbb{B}_{\text{ind}}$ 's guess  $\hat{b}_i$  actually equals  $b_i$ . Let  $\text{Check}_{\mathbb{I}}$  be the event that all  $\text{Check}_i$  hold (for  $i \in \mathbb{I}$ ) and define  $\text{Good}_{\mathbb{I}}$  analogously.

As  $\mathbb{B}_{\text{ind}}$ 's simulation of  $\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-mku-cca}^*}$  is perfect, we know that

$$\text{Adv}^{(n,\kappa)\text{-mku-cca}^*}(\mathbb{A}_{\text{mku}}) = \Pr[\text{Check}_{\mathbb{I}}],$$

moreover,

$$\begin{aligned} \Pr\left[\text{Exp}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{B}_{\text{ind}}) = 1\right] &\geq \Pr[\text{Check}_{\mathbb{I}} \wedge \text{Good}_{\mathbb{I}}] + \Pr[\neg\text{Check}_{\mathbb{I}} \wedge b = 0] \\ &= \Pr[\text{Good}_{\mathbb{I}} \mid \text{Check}_{\mathbb{I}}] \Pr[\text{Check}_{\mathbb{I}}] + \frac{1}{2} (1 - \Pr[\text{Check}_{\mathbb{I}}]) \end{aligned}$$

which implies that

$$\text{Adv}_{\text{PKE}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{B}_{\text{ind}}) \geq (2 \Pr[\text{Good}_{\mathbb{I}} \mid \text{Check}_{\mathbb{I}}] - 1) \text{Adv}^{(n,\kappa)\text{-mku-cca}^*}(\mathbb{A}_{\text{mku}}).$$

To bound  $\Pr[\text{Good}_{\mathbb{I}} \mid \text{Check}_{\mathbb{I}}]$  we exploit the correctness definition, specifically that its quantification (Def. 1) ensures that whenever  $\text{Check}_i$  holds, we have that  $\Pr\left[\text{PKE.Dec}_{\hat{s}k_i}(\text{PKE.Enc}_{pk_i}(m)) = m\right] \geq 1 - \delta$ , irrespective of  $m$  and where the probability is only over the randomness of  $\text{PKE.Enc}$ .

If, for a given  $i$ , decryption is correct strictly more than  $q/2$  times, then we are guaranteed that  $\text{Good}_i$  occurs. If we let  $B(q, p)$  be the binomial distribution over  $q$  trials and with probability  $p$ , then

$$\Pr[\text{Good}_i \mid \text{Check}_i] \geq \Pr\left[B(q, (1 - \delta)) > \frac{q}{2}\right]$$

and, as this bound only relies on the randomness of the challenge encryption oracle, guaranteed independent for differing  $i$ , we may conclude that

$$\Pr[\text{Good}_{\mathbb{I}} \mid \text{Check}_{\mathbb{I}}] \geq \left(\Pr\left[B(q, (1 - \delta)) > \frac{q}{2}\right]\right)^n.$$

Finally, we note that

$$\Pr\left[B(q, (1 - \delta)) > \frac{q}{2}\right] \geq 1 - 2^q (\delta(1 - \delta))^{\frac{q}{2}}$$

by a standard application of known bounds on binomial tails, requiring  $\delta \leq 1/2$  (see details below). Plugging in all the various bounds recovers the theorem statement.

For the binomial tail bound, we use the Chernoff–Hoeffding bound [27], which states that, for a binomial distribution  $B(q, p)$  over  $q$  trials and with probability  $p$ , and any  $k$  satisfying  $p < \frac{k}{q} < 1$  the tail bound

$$\Pr[B(q, p) \geq k] \leq \exp\left[-qD\left(\frac{k}{q} \parallel p\right)\right]$$

holds, where  $D(a\|b)$  is the Kullback–Leibler divergence defined as  $D(a\|b) = a \ln\left(\frac{a}{b}\right) + (1-a) \ln\left(\frac{1-a}{1-b}\right)$ .

We further use the trick that  $\Pr[B(q, (1-\delta)) > \frac{q}{2}] = 1 - \Pr[B(q, \delta) \leq \frac{q}{2}]$ , so the relevant Kullback–Leibler divergence becomes

$$\begin{aligned} D\left(\frac{1}{2}\left\|\delta\right.\right) &= \frac{1}{2} \ln\left(\frac{\frac{1}{2}}{\delta}\right) + \left(1 - \frac{1}{2}\right) \ln\left(\frac{\left(1 - \frac{1}{2}\right)}{1 - \delta}\right) \\ &= \frac{1}{2} \ln\left(\frac{1}{2\delta}\right) + \frac{1}{2} \ln\left(\frac{1}{2(1-\delta)}\right) \\ &= \ln\left[\left(\frac{1}{4\delta(1-\delta)}\right)^{\frac{1}{2}}\right], \end{aligned}$$

which allows us to compute the bound

$$\begin{aligned} \Pr\left[B(q, (1-\delta)) > \frac{q}{2}\right] &\geq 1 - \exp\left[-qD\left(\frac{1}{2}\left\|\delta\right.\right)\right] \\ &= 1 - \exp\left[-q \ln\left[\left(\frac{1}{4\delta(1-\delta)}\right)^{\frac{1}{2}}\right]\right] \\ &= 1 - 2^q (\delta(1-\delta))^{\frac{q}{2}}. \end{aligned}$$

□

**Corollary 1** (IND  $\rightarrow$  UKU). *Let PKE be a  $(\gamma, \delta)$ -correct encryption scheme with  $\delta < 1/2$ . Then there is a type-preserving SFBB reduction  $\mathbb{B}_{\text{ind}}$  such that, for every  $\mathbb{A}_{\text{uku}}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{B}_{\text{ind}}) \geq c \cdot \text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-uku-cca}^*}(\mathbb{A}_{\text{uku}}) - \kappa\gamma,$$

with  $c, q$ , and  $\mathbb{B}_{\text{ind}}$ 's overhead as above (Thm. 2).

### 3.3 Alternative Mechanisms

As we mentioned before, our mechanism to capture multi-instance security differs slightly from those used by BRT and AGK, respectively, even when accounting for changes in primitive and correctness. At first sight, the differences might appear mostly cosmetic, though there are some subtleties involved.

#### The BRT Notion: Requiring $n = \kappa$ , Possibly Corrupted, Targets.

BRT require an adversary to return the xor of all bits, but allow those bits or corresponding users to be corrupted. Fig. 5 reflects the small change needed in the code of our security experiment to match BRT's mechanism (ignoring a minor, inconsequential difference, as BRT have a single, merged corruption oracle that returns both key and bit). As motivation for including corruptions,

Experiment $\text{Exp}_{\text{PKE}}^{(\leq \kappa, \kappa)\text{-ind-cca}^*}(\mathbb{A})$	Experiment $\text{Exp}_{\text{PKE}}^{(\geq n, \kappa)\text{-ind-cca}^*}(\mathbb{A})$
4: <b>if</b> $ \mathbb{I}  \neq \kappa$ <b>then</b> $\hat{b} \leftarrow 0$	4: <b>if</b> $ \mathbb{I}  < n \vee \mathbb{I} \cap (\mathbb{K} \cup \mathbb{B}) \neq \emptyset$ <b>then</b> $\hat{b} \leftarrow 0$

**Fig. 5.** The main differences between our mechanism for multi-instance indistinguishability (Fig. 4) and prior art revolve around line 4: BRT’s experiment  $\text{Exp}_{\text{PKE}}^{(\leq \kappa, \kappa)\text{-ind-cca}^*}(\mathbb{A})$  (left) and AGK’s experiment  $\text{Exp}_{\text{PKE}}^{(\geq n, \kappa)\text{-ind-cca}^*}(\mathbb{A})$  (right). The differences are highlighted in blue.

BRT discuss the scenario that, say, half of the keys generated are hopelessly insecure: an adversary breaks the insecure half and corrupts the rest, thus being successful. Moreover, they mention that their choice implies security under a corruptionless notion with dynamically chosen  $\mathbb{I}$ .

Although the implication is of course true, and something can be said to target the strongest possible notion, corruptions have a habit of creating complications for reductions and provable security in general. Yet, we believe the inclusion of corruptions, or not, should reflect the threat model of the adversary and that choice should be orthogonal to the number of users being targeted. BRT, instead of having an explicit hardness parameter  $n$ , restrict an adversary to make at most  $q_c$  corruption queries to avoid trivial wins when  $q_c = \kappa$ . Yet, whether the resulting, intuitive hardness will or should then match  $n = \kappa - q_c$ , is unclear.

We address the equivalence between BRT’s mechanism and our general mechanism (with corruptions) in Lemmas 1 and 2. Both lemmas have in common that the respective reductions may make up to  $\kappa - n$  additional bit corruptions. In other words, the reductions are not type-preserving, making the equivalence somewhat sloppy. As an aside, using techniques similar to those to prove Thm. 2, the key corruption oracle could be used (at a loss) to simulate the bit corruption oracle instead(see the full version for the proofs).

**Lemma 1 (main notion  $\implies$  BRT).** *Let  $n \leq \kappa$  and  $q_c \leq \kappa - n$ . Then there is an SFBB reduction  $\mathbb{B}$  such that, for every adversary  $\mathbb{A}$  making at most  $q_c$  corruption oracle calls,*

$$\text{Adv}_{\text{PKE}}^{(\leq \kappa, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{B}),$$

where  $\mathbb{B}$  makes at most  $\kappa - n$  additional bit corruption oracle calls.

**Lemma 2 (BRT  $\implies$  main notion).** *Let  $n \leq \kappa$ . Then there is an SFBB reduction  $\mathbb{B}$  such that, for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{PKE}}^{(\leq \kappa, \kappa)\text{-ind-cca}^*}(\mathbb{B}),$$

where  $\mathbb{B}$  makes at most  $\kappa - n$  additional bit corruption oracle calls.

**The AGK Notion: Allowing More than  $n$  Targets without Corruptions.** When AGK studied KEMs in the multi-instance setting, they used a xor

notion with the  $n$  as the *minimum* number of targets to attack (out of  $\kappa$  possible) as an explicit parameter; moreover, an adversary would not have access to any corruption oracles. Fig. 5 reflects the small change needed in the code of our security experiment to match AGK’s mechanism with corruptions added (where we fixed a minor bug in their code; rather than setting  $\hat{b} \leftarrow 0$  their experiment would immediately return 0 instead, inadvertently granting an adversary that deliberately returns a compromised handle the significant advantage of  $-1$ ).

Absent corruptions, AGK indicated that for some pathological schemes, breaking more targets might paradoxically be easier than breaking fewer [3, App. C]. In those cases, the freedom to return a set  $I$  of cardinality greater than  $n$  would make life easier for an adversary, leading to a stronger notion.

In the presence of corruptions, requiring the adversary to target exactly  $n$  users as we do is without loss of generality. As an example, if an adversary can figure out the xor of  $n + 1$  honest bits, it can bit-corrupt any single one of these  $n + 1$ , and xor the resulting bit out of the initial guess to obtain a final one on  $n$  bits instead. We formalize this intuition below.

**Lemma 3 (main notion  $\implies$  AGK $\star$ ).** *There is an SFBB adversary  $\mathbb{B}$  such that, for every  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(\geq n, \kappa)\text{-ind-cca}\star}(\mathbb{A}) \leq \text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}\star}(\mathbb{B}).$$

*If  $\mathbb{A}$  returns a list of  $n'$  targets,  $\mathbb{B}$  makes  $n' - n$  additional calls to its bit corruption oracle.*

### 3.4 Real-or-Random XOR Indistinguishability

An alternative notion of indistinguishability, known as real-or-random indistinguishability (ROR), sees the adversary tasked with figuring out whether a challenge ciphertext contains the adversarially chosen message  $m$  or an unknown, randomly chosen message. The game  $\text{Exp}_{\text{PKE}}^{(n, \kappa)\text{-ror-cca}\star}$  is exactly as in Fig. 4, apart from the challenge encryption oracle  $\mathcal{E}_{\text{ROR}}(i, m)$ , which sets  $m_0 \leftarrow m$  and  $m_1 \leftarrow \$_[m]$  to then call (left-or-right)  $\mathcal{E}(i, m_0, m_1)$ .

By construction, left-or-right indistinguishability easily implies real-or-random indistinguishability. That statement is as true in the multi-instance setting as it is in the classical single-user setting. Conversely, in the single-user setting, it has long been established that the reduction from ROR to IND loses a factor 2 [8]. However, BRT showed that in the multi-instance setting, the factor 2 blows up exponentially to, in their case,  $2^\kappa$ . Yet, BRT argue that this exponential loss is not as bad as it might seem, given that the multi-instance advantages are supposed to be exponentially smaller than their single-user counterparts. Thus, reductions incurring losses exponential in  $\kappa$  or  $n$  can still be valuable.

To adapt BRT’s reduction to our setting, we require  $n = \kappa$ , implying that  $\mathbb{A}$  cannot access its corruption oracles. Otherwise, corruptions would make the reduction noticeable once at least one  $b_i$  is set to 1, potentially influencing an adversary’s behaviour in unpredictable ways (see the full version for the proof).



**Theorem 3.** *There is an SFBB reduction  $\mathbb{B}$  such that, for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(\kappa, \kappa)\text{-ind-cca}}(\mathbb{A}) \leq 2^\kappa \cdot \text{Adv}_{\text{PKE}}^{(\kappa, \kappa)\text{-ror-cca}}(\mathbb{B}),$$

where  $\mathbb{B}$  additionally draws  $\kappa$  bits uniformly at random.

Furthermore, a reduction playing an  $(n, n)$  game can exploit an adversary playing a  $(n, \kappa)$  game by guessing in advance the set  $I$  of targets that the adversary will return. A correct guess allows the reduction to simulate the remaining keys without being noticed (see the full version for the proof).

**Theorem 4.** *There is an SFBB reduction  $\mathbb{B}$  such that, for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \binom{\kappa}{n} \cdot \text{Adv}_{\text{PKE}}^{(n, n)\text{-ind-cca}}(\mathbb{B}).$$

$\mathbb{B}$ 's overhead consists of generating  $\kappa - n$  fresh keypairs, sampling  $\kappa - n$  bits, and choosing a subset of  $[\kappa]$  of cardinality  $n$  uniformly at random.

Composing Thm. 3 and 4, we obtain the following bound.

**Corollary 2 (ROR  $\implies$  IND).** *There is an SFBB reduction  $\mathbb{B}$  such that, for any adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \binom{\kappa}{n} \cdot 2^n \cdot \text{Adv}_{\text{PKE}}^{(n, n)\text{-ror-cca}}(\mathbb{B}).$$

$\mathbb{B}$ 's overhead consists of generating  $\kappa - n$  fresh keypairs, sampling  $\kappa$  bits, and choosing a subset of  $[\kappa]$  of cardinality  $n$  uniformly at random.

An alternative bound losing a factor  $2^\kappa$  is possible by combining Thm. 3 with Lemma 2, however a simple analysis shows that whenever  $n < \kappa/5$  the corollary above is preferable.

At first glance, an exponential-looking loss of  $2^\kappa$  might seem severe, potentially rendering the resulting bound vacuous. Yet, as BRT already highlighted, the multi-instance advantages themselves might vanish exponentially in  $n$ , making the bounds relevant for the notions being compared. Nonetheless, tighter bounds still matter; unfortunately achieving even tighter bounds in the general case seems challenging [5, 12].

## 4 Inheriting Multi-Instance Security

### 4.1 TagKEM: Definition and Notion of Security

Our goal is to turn the AGK multi-instance secure KEM into a PKE. Yet, for the construction of hybrid encryption, the more general TagKEMs, where encapsulation is split into two algorithms (TKEM.Key and TKEM.Enc) have proven more powerful [2]: intuitively speaking, splitting the algorithm allows the tag and consequently the key encapsulation to depend on the data encapsulation, making

Experiment $\text{Exp}_{\text{TXEM}}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A})$	Oracle $\mathcal{C}(i, \ell)$
$(\text{pk}_1, \text{sk}_1), \dots, (\text{pk}_\kappa, \text{sk}_\kappa) \leftarrow_{\$} \text{TXEM.Kg}$ $b_1, \dots, b_\kappa \leftarrow_{\$} \{0, 1\}$ $(\mathbb{I}, \hat{b}) \leftarrow_{\$} \mathbb{A}^{\mathcal{C}, \mathcal{E}, \mathcal{D}, \mathcal{K}, \mathcal{B}}(\text{pk}_1, \dots, \text{pk}_\kappa)$ <b>if</b> $ \mathbb{I}  \neq n \vee \mathbb{I} \cap (\mathbb{K} \cup \mathbb{B}) \neq \emptyset$ <b>then</b> $\hat{b} \leftarrow 0$ <b>return</b> $\bigoplus_{i \in \mathbb{I}} b_i = \hat{b}$	$(K_0, \sigma) \leftarrow_{\$} \text{TXEM.Key}_{\text{pk}_i}(\ell)$ $\mathbf{E}_i \leftarrow \widehat{\langle \sigma, K_0 \rangle}$ $K_1 \leftarrow_{\$} \{0, 1\}^\ell$ <b>return</b> $K_{b_i}$
Oracle $\mathcal{D}(i, \langle c, \tau \rangle, \ell)$	Oracle $\mathcal{E}(i, j, \tau)$
$K \leftarrow \text{TXEM.Dec}_{\text{sk}_i}(c, \tau, \ell)$ <b>if</b> $\mathbf{P}_i(c, \tau) \neq \emptyset$ $\quad K' \leftarrow \mathbf{P}_i(c, \tau), \ell' \leftarrow \min\{\ell,  K' \}$ <b>else</b> $\quad K' \leftarrow \varepsilon, \ell' \leftarrow 0$ <b>if</b> $\langle c, \tau \rangle \in \mathbf{C}_i \wedge K[\ell'] = K'[\ell']$ $\quad$ <b>return</b> $\perp$ <b>return</b> $K$	<b>if</b> $\mathbf{E}_i[j] = \emptyset$ <b>then return</b> $\perp$ $\langle \sigma, K \rangle \leftarrow \mathbf{E}_i[j], \mathbf{E}_i[j] \leftarrow \emptyset$ $c \leftarrow_{\$} \text{TXEM.Enc}(\sigma, \tau)$ $\mathbf{P}_i(c, \tau) \leftarrow K$ $\mathbf{C}_i \xleftarrow{\cup} \langle c, \tau \rangle$ <b>return</b> $c$
	Oracle $\mathcal{K}(i)$ Oracle $\mathcal{B}(i)$
$\mathbb{K} \xleftarrow{\cup} i$ <b>return</b> $\text{sk}_i$	$\mathbb{B} \xleftarrow{\cup} i$ <b>return</b> $b_i$

**Fig. 6.** Multi-instance indistinguishability notion for TXEM. In blue the same strengthening as in Fig. 4 in the case of imperfect correctness, with a slightly more complex admin to accommodate tags and length extension. We take  $K[\ell]$  to mean the first  $\ell$  bits of  $K$  and  $\varepsilon$  as the empty string.

CCA security of the hybrid construction easier to achieve (cf. the Kurosawa–Desmedt scheme [31]). In Def. 4 we introduce a further generalization, called TagXEM, by allowing extendable output lengths for the ephemeral keys produced by the TagXEM.

**Definition 4 (TagXEM).** A TagXEM is a tuple of algorithms (TXEM.Kg, TXEM.Key, TXEM.Enc, TXEM.Dec, TXEM.Check), where long-term key generation TXEM.Kg on input  $\text{pm}$  outputs a random keypair  $(\text{pk}, \text{sk})$ ; ephemeral key generation TXEM.Key on input  $\text{pk}$  and  $\ell \in \mathbb{Z}_{>0}$ , outputs a random ephemeral key  $K \in \{0, 1\}^\ell$  and an internal state  $\sigma$ , subsequently encapsulation TXEM.Enc on input a state  $\sigma$  and a tag  $\tau \in \mathcal{T}$ , deterministically outputs an encapsulation  $c$ , or a special symbol  $\perp$  denoting failure. The deterministic decapsulation algorithm TXEM.Dec takes input a private key  $\text{sk}$ , an encapsulation  $c$ , a tag  $\tau$ , and a length  $\ell$ , and outputs either a key  $K \in \{0, 1\}^\ell$  or  $\perp$  to denote failure. Finally, the deterministic TXEM.Check takes as input the system parameters  $\text{pm}$  as well as a purported public/private key pair  $(\text{pk}, \text{sk})$  and returns true or false.

If we restrict to a single value  $\ell$ , the usual notion of TagKEMs appears; moreover if we restrict to a single value of  $\tau$ , the TXEM.Key and TXEM.Enc algorithms can

be merged into a single key encapsulation mechanism, leading to normal KEMs (or XEMs if the variable output length is still incorporated). Consequently, the correctness and security definitions for the more general TagXEMs, as discussed throughout this section, imply corresponding definitions for KEM, XEM, and TagKEM.

For correctness, we allow the effective tag space  $\mathcal{T}_\ell$  to depend on the length  $\ell$  of the ephemeral key. Similarly to Def. 1, we define  $(\gamma, \delta)$ -correctness for TagXEM. To ensure correctness for all  $\tau$ , including those that depend on  $K$ ,  $\tau$ 's quantifier sits inside the probability statement.

**Definition 5** ( $(\gamma, \delta)$ -Correctness TagXEM). *Let  $\gamma, \delta \in [0, 1]$ . Then a tag extendable-output key encapsulation mechanism TXEM is called  $(\gamma, \delta)$ -correct iff*

1.  $\Pr[(\text{pk}, \text{sk}) \leftarrow \text{TXEM.Kg}(\text{pm}) : \text{TXEM.Check}(\text{pm}, \text{pk}, \text{sk}) = \text{false}] \leq \gamma;$
  2. *if  $\text{TXEM.Check}(\text{pm}, \text{pk}, \text{sk}) = \text{true}$  then for all  $\ell \in \mathbb{Z}_{>0}$  it holds that*
- $$\Pr \left[ (K, \sigma) \leftarrow \text{TXEM.Key}_{\text{pk}}(\ell) : \exists \tau \in \mathcal{T}_\ell \text{ s.th. } \begin{array}{l} c \leftarrow \text{TXEM.Enc}(\sigma, \tau) \\ \text{TXEM.Dec}_{\text{sk}}(c, \tau, \ell) \neq K \end{array} \right] \leq \delta .$$

For security, Abe et al.'s notion of TagKEM indistinguishability [2] transfers easily to the multi-instance setting. The relevant game is given in Fig. 6, where we also made the necessary changes to deal with the variable output length of TagXEMs, plus the strengthening of  $\mathcal{D}$  in the case of imperfect correctness (cf. Sect. 3.2).

**Definition 6.** *Let TXEM be a TagXEM. Then the xor-indistinguishability advantage of an adversary  $\mathbb{A}$  is*

$$\text{Adv}_{\text{TXEM}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) = 2 \cdot \Pr \left[ \text{Exp}_{\text{TXEM}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) = 1 \right] - 1 ,$$

where the experiment is defined in Fig. 6.

If we fix  $\ell$  and set  $\mathcal{T}_\ell$  to a single element, the notion captures multi-instance security for standard KEMs, which is near equivalent (see Sect. 3.3) the notion that AGK used. In other words, provided MI-gapCDH is hard, their construction achieves  $(n, \kappa)$ -IND-CCA security in the random oracle model, but only for fixed  $\ell$  and trivial  $\mathcal{T}_\ell$  [4, Thm. 2].

## 4.2 Extending the Output of a TagKEM

First, we show how combining a TagKEM with a fixed output length and a suitable pseudorandom extendable output function (XOF), yields a TagXEM that inherits the MI security of the underlying KEM. Recall that a XOF, for instance SHAKE128 and SHAKE256 as standardized by NIST [35], is a function  $F : \mathcal{X} \times \mathbb{Z}_{>0} \rightarrow \{0, 1\}^*$  for some finite domain  $\mathcal{X}$  that on input a seed  $s \in \mathcal{X}$  and a desired output length  $\ell$ , outputs a value  $y \in \{0, 1\}^\ell$ . Moreover, if  $\ell < \ell'$ , then

TXEM.Key <sub>pk</sub> ( $\ell$ )	TXEM.Enc( $\sigma', \tau$ )	TXEM.Dec( $c, \tau, \ell$ )
$(K^{\text{kem}}, \sigma) \leftarrow \text{TKEM.Key}_{\text{pk}}$ $K^{\text{xem}} \leftarrow F(K^{\text{kem}}, \ell)$ $\sigma' \leftarrow \langle \sigma, \ell \rangle$ <b>return</b> $(K^{\text{xem}}, \sigma')$	$\langle \sigma, \ell \rangle \leftarrow \sigma'$ <b>if</b> $\tau \notin \mathcal{T}_\ell$ : <b>return</b> $\perp$ $c \leftarrow \text{TKEM.Enc}(\sigma, \tau)$ <b>return</b> $c$	<b>if</b> $\tau \notin \mathcal{T}_\ell$ : <b>return</b> $\perp_{\text{TAG}}$ $K^{\text{kem}} \leftarrow \text{TKEM.Dec}(c, \tau)$ <b>if</b> $K^{\text{kem}} = \perp$ : <b>return</b> $\perp_{\text{KEM}}$ $K^{\text{xem}} \leftarrow F(K^{\text{kem}}, \ell)$ <b>return</b> $K^{\text{xem}}$

**Fig. 7.** A TagXEM TXEM from a TagKEM TKEM with keyspace  $\{0, 1\}^k$  and a XOF with seed space  $\mathcal{X} = \{0, 1\}^k$ . The key generation algorithm TXEM.Kg is unchanged from TKEM.Kg.

$F(s, \ell)$  is a prefix of  $F(s, \ell')$  for all  $s$ . This prefix preservation is not a requirement of our constructions; rather we model the property to ensure SHAKE128 and SHAKE256 are suitable real-world instantiations.

As security notion for a XOF  $F$  we use its multi-challenge pseudorandomness, which is a standard distinguishing advantage  $\text{Adv}_F^{\text{psrnd}}(\mathbb{A})$ : an adversary needs to distinguish between either a real oracle that, on input a desired length  $\ell$ , samples a seed  $s \leftarrow \mathcal{X}$  uniformly at random and returns  $F(s, \ell)$ , or an ideal oracle that, on input said  $\ell$ , simply returns a uniformly sampled string of length  $\ell$ .

The construction of the TagXEM is given in Fig. 7 and the security claim follows in Thm. 5 (see the full version for the proof). If the PsRND advantage of  $F$  is sufficiently small, then TXEM inherits the multi-instance security of TKEM; moreover, as the result holds for arbitrary  $\mathcal{T}$  and  $\mathcal{T}_\ell$ , it holds for the trivial spaces, yielding a slightly simpler XEM from KEM result.

**Theorem 5.** *Let TKEM be a  $(\gamma, \delta)$ -correct TagKEM sampling keys from  $\{0, 1\}^k$  and with tagspace  $\mathcal{T}$ , let  $F : \{0, 1\}^k \times \mathbb{Z}_{>0} \rightarrow \{0, 1\}^*$  be a XOF, and let TXEM be a TagXEM as given in Fig. 7 for arbitrary  $\mathcal{T}_\ell \subseteq \mathcal{T}$ . Then TXEM is  $(\gamma, \delta)$ -correct, and there are SFBB reductions  $\mathbb{B}$  and  $\mathbb{C}$  such that, for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{TXEM}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{TKEM}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{B}) + 2 \cdot \text{Adv}_F^{\text{psrnd}}(\mathbb{C}).$$

*If  $\mathbb{A}$  calls  $\mathbb{C}$   $q_c$  times and  $\mathbb{D}$   $q_d$  times, then  $\mathbb{B}$ 's overhead consists of at most  $q_c + q_d$  evaluations of  $F$ , while  $\mathbb{C}$ 's overhead consists of doing  $\kappa$  executions of TKEM.Kg, at most  $q_c$  executions of TKEM.Key and TKEM.Enc, and at most  $q_d$  executions of TKEM.Dec.*

One concern is whether the PsRND advantage of  $F$  will be sufficiently small. Suppose  $k$  is the output length of the underlying TagKEM. A generic attacker would always be able to fix  $\ell > k$  and evaluate  $F$  for, say,  $N$  seeds offline in the hope of colliding with any of the challenge evaluations. The PsRND distinguishing advantage of such an adversary is of order  $(q_c + q_d)N/2^k$ , indicating that the

$\frac{\text{PKE.Enc}_{\text{pk}}(m)}{(K, c_1) \leftarrow \$ \text{XEM.Enc}_{\text{pk}}( m )}$ $c_2 \leftarrow K \oplus m$ $\text{return } \langle c_1, c_2 \rangle$	$\frac{\text{PKE.Dec}_{\text{sk}}(\langle c_1, c_2 \rangle)}{K \leftarrow \text{XEM.Dec}_{\text{sk}}(c_1,  c_2 )}$ $\text{if } K = \perp \text{ then return } \perp$ $m \leftarrow K \oplus c_2$ $\text{return } m$
$\frac{\text{PKE}'.\text{Enc}_{\text{pk}}(m)}{(K, \sigma) \leftarrow \$ \text{TXEM.Key}_{\text{pk}}( m )}$ $c_2 \leftarrow K \oplus m$ $c_1 \leftarrow \text{TXEM.Enc}(\sigma, c_2)$ $\text{return } \langle c_1, c_2 \rangle$	$\frac{\text{PKE}'.\text{Dec}_{\text{sk}}(\langle c_1, c_2 \rangle)}{K \leftarrow \text{TXEM.Dec}_{\text{sk}}(c_1, c_2,  c_2 )}$ $\text{if } K = \perp \text{ then return } \perp$ $m \leftarrow K \oplus c_2$ $\text{return } m$

**Fig. 8.** Two hybrid encryption schemes: PKE (top row) is a conventional hybrid scheme combining a XEM with the OTP to yield a CPA-secure PKE, while PKE' (bottom row) combines a TagXEM with the OTP to yield a CCA-secure PKE. The key generation and checking algorithms are equivalent to their XEM resp. TXEM counterparts.

underlying TagKEM already needs to provide keys long enough for Thm. 5 to yield meaningful multi-instance security.

### 4.3 A PKE Inheriting (Tag)XEM Security

As a multi-instance secure XEM provides us with ephemeral keys of any desired length, we can combine it with an information-theoretic DEM in order to achieve PKE. Here we opt for the one-time-pad (OTP), as it is the simplest and best-known primitive providing perfect secrecy. The beauty of the OTP is that whether you switch out the ephemeral key for a uniform random one, or the message for a uniform random one, the resulting ciphertext distribution is the same. It allows the PKE to tightly inherit the MI-security of the XEM, albeit yielding only real-or-random security under chosen-plaintext attacks. The construction is provided in full in Fig. 8 (top row); the security claim is captured in Thm. 6 (see the full version for the proof).

**Theorem 6 (ROR-CPA PKE).** *Let XEM be a  $(\gamma, \delta)$ -correct XEM, and let PKE be a hybrid encryption scheme as given in Fig. 8. Then PKE is  $(\gamma, \delta)$ -correct, and there is a type-preserving SFBB reduction  $\mathbb{B}$  such that for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}}^{(n, \kappa)\text{-ror-cpa}^*}(\mathbb{A}) \leq \text{Adv}_{\text{XEM}}^{(n, \kappa)\text{-ind-cpa}^*}(\mathbb{B}).$$

One might hope that adding information-theoretic MACs to the DEM would result in the inheritance of CCA security, but that is easier said than shown. For instance, the usual proof technique of a game hop where all decryption queries are disallowed does not work: after breaking only a single KEM private

---

Oracle  $\mathcal{C}_{\text{ROP}}(i, \ell, \Pi)$

$(K_0, \sigma) \leftarrow \text{TXEM.Key}_{\text{pk}_i}(\ell)$

$\mathbf{E}_i \leftarrow \sigma$

$K_1 \leftarrow \Pi(K_0)$

**return**  $K_{b_i}$

**Fig. 9.** Fig. 6 is upgraded to  $\text{Exp}_{\text{TXEM}}^{(n, \kappa)\text{-rop-cca}^*}$  by letting  $\mathcal{C}_{\text{ROP}}$  replace  $\mathcal{C}$ .

key, the reduction will be found out as not being faithful. Sadly, a single-instance break (of the reduction) suffices to show that that reduction cannot demonstrate multi-instance security.

Luckily, TagKEMs allow for a modified hybrid scheme for which the DEM no longer needs to satisfy CCA security for the resulting PKE to be guaranteed CCA-secure: in the single-instance setting, if the TagKEM is CCA-secure, then so is the PKE [2]. We upgrade the construction to use TagXEMs and the OTP in Fig. 8 (bottom row) and show its multi-instance inheritance in Thm. 7 (see the full version for the proof).

**Theorem 7 (ROR-CCA PKE).** *Let TXEM be a  $(\gamma, \delta)$ -correct TagXEM, and let PKE' be a hybrid encryption scheme as given in Fig. 8. Then PKE' is  $(\gamma, \delta)$ -correct, and there is a type-preserving SFBB reduction  $\mathbb{B}$  such that for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}'}^{(n, \kappa)\text{-ror-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{TXEM}}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{B}).$$

While encouraging, the claim that the constructed PKE inherits the multi-instance security of the TagXEM is dampened by the exponential separation between the ROR security notion and IND, as argued in Sect. 3.4. Indeed, extrapolating to the latter notion by combining Thm. 7 with Cor. 2, we have only achieved the following bound.

**Corollary 3.** *Let TXEM be a  $(\gamma, \delta)$ -correct TagXEM, and let PKE' be a hybrid encryption scheme as given in Fig. 8. Then PKE' is  $(\gamma, \delta)$ -correct, and there is a type-preserving SFBB reduction  $\mathbb{B}$  such that for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}'}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \binom{\kappa}{n} \cdot 2^n \cdot \text{Adv}_{\text{TXEM}}^{(n, n)\text{-ind-cca}}(\mathbb{B}),$$

where  $\mathbb{B}$ 's overhead is dominated by generating  $\kappa - n$  fresh keypairs, sampling  $\kappa$  bits, and choosing a subset of  $[\kappa]$  of cardinality  $n$  uniformly at random.

#### 4.4 Real-Or-Permuted: A Strengthened Notion for KEM Security

If we want to achieve an IND-CCA PKE more tightly, we seem to need a different notion of security for our TagXEMs. What could such a notion look like?

Our solution is a novel, stronger KEM notion, which we will refer to as “real-or-permuted”, or ROP for short. Fig. 9 provides the crucial new challenge oracle. The adversary has to guess whether a tentative  $K$  is the one encapsulated under  $c$ , or whether an adaptively chosen permutation has been applied to it. As permutations preserve the distribution of the sampling space, there are no choices of  $\Pi$  that make the game generically and trivially winnable.

Technically, we need to specify how the adversary provides  $\Pi$  such that it is guaranteed, or can be checked, to be a permutation. Hence, formally we define ROP with respect to a class of permutations  $\mathcal{P}$ , reminiscent of for instance key-dependent message [24] or related-key attack [10] definitions. We require that membership  $\Pi \in \mathcal{P}$  is easy to check (e.g. ROP can simply index an element in  $\mathcal{P}$ ) and that, by definition,  $\mathcal{P}$  can be verified to indeed only contain permutations. For our main results, it suffices if  $\mathcal{P}$  is the class of one-time pads, in the sense that  $\Pi$  specifies the key (or pad) of the one-time pad enciphering. Henceforth, we will assume that ROP is defined with respect to that class, unless explicitly stated otherwise.

The new notion ROP and IND relate to each other much the same way as IND and ROR for PKE. It is not hard to see that ROP tightly implies IND, whereas the other direction seems to incur the same loss as the ROR-to-IND implication for PKE (see the full version). For completeness, ROP lends itself equally well to XEMs and KEMs, or notions without corruptions or a decryption oracle. Finally, if any of the above primitives are constructed using an IND-secure PKE (e.g. using a Fujisaki–Okamoto style transform [21, 22, 28]), then achieving ROP is as easy as achieving IND: simply let  $K$  be the “left” message, and  $\Pi(K)$  be the “right”!

#### 4.5 PKE' Tightly Inherits IND-CCA Security

Using ROP in place of IND, we are able to show directly that the PKE constructions of Fig. 8 are IND-CPA resp. IND-CCA secure, by (as before) giving a (Tag)XEM reduction that provides a perfect simulation for the PKE adversary.

The crucial observation is that for any pair of messages  $m_0, m_1 \in \{0, 1\}^\ell$ , there exist a permutation  $\Pi_{m_0 \rightarrow m_1}$  on  $\{0, 1\}^\ell$  such that the message encapsulations are related as  $K \oplus m_1 = \Pi_{m_0 \rightarrow m_1}(K) \oplus m_0$ . Namely, the permutation that on input  $K$ , outputs  $m_0 \oplus m_1 \oplus K$  (see the full version for the proof).

**Theorem 8 (IND-CCA PKE).** *Let TXEM be a  $(\gamma, \delta)$ -correct TagXEM, and let PKE' be a hybrid encryption scheme as given in Fig. 8. Then PKE' is  $(\gamma, \delta)$ -correct, and there is a type-preserving SFBB reduction  $\mathbb{B}$  such that for every adversary  $\mathbb{A}$ ,*

$$\text{Adv}_{\text{PKE}'}^{(n, \kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{TXEM}}^{(n, \kappa)\text{-rop-cca}^*}(\mathbb{B}).$$

We leave it to the reader to verify that as before, employing a ROP-CPA XEM in place of the TagXEM yields IND-CPA security for the PKE of Fig. 8 (top row), by adapting the proof of Thm. 6 to the above. We again stress that using an information-theoretically CCA-secure DEM together with a CCA XEM does not seem to yield a proof of CCA inheritance to the PKE (see Sect. 4.3).

$\frac{\text{TXEM.Kg}}{\text{pk}', \text{sk}' \leftarrow_{\$} \text{KEM.Kg}}$ $\text{pk} \leftarrow \text{pk}'$ $\text{sk} \leftarrow \langle \text{pk}', \text{sk}' \rangle$ $\text{return } (\text{pk}, \text{sk})$	$\frac{\text{TXEM.Key}_{\text{pk}}(\ell)}{(K^{\text{kem}}, c) \leftarrow_{\$} \text{KEM.Enc}_{\text{pk}}}$ $\ell' \leftarrow \ell + \ell_{\text{mackey}}$ $K^{\text{mac}} \  K^{\text{xem}} \leftarrow F(\text{pk}, c, K^{\text{kem}}, \ell')$ $\sigma \leftarrow \langle c, K^{\text{mac}} \rangle$ $\text{return } K^{\text{xem}}$
$\frac{\text{TXEM.Check}(\text{pk}, \text{sk})}{\langle \text{pk}', \text{sk}' \rangle \leftarrow \text{sk}}$ $\text{if } \text{pk} \neq \text{pk}' \text{ then return } 0$ $\text{return KEM.Check}(\text{pk}', \text{sk}')$	$\frac{\text{TXEM.Dec}_{\text{sk}}(\langle c, \text{mac} \rangle, \tau, \ell)}{\langle \text{pk}', \text{sk}' \rangle \leftarrow \text{sk}}$ $K^{\text{kem}} \leftarrow \text{KEM.Dec}_{\text{sk}'}(c)$ $\text{if } K^{\text{kem}} = \perp \text{ then return } \perp$ $\ell' \leftarrow \ell + \ell_{\text{mackey}}$ $K^{\text{mac}} \  K^{\text{xem}} \leftarrow F(\text{pk}', c, K^{\text{kem}}, \ell')$ $\text{if } \text{MAC}_{K^{\text{mac}}}(\tau) \neq \text{mac} \text{ then return } \perp$ $\text{return } K^{\text{xem}}$
$\frac{\text{TXEM.Enc}(\sigma, \tau)}{\langle c, K^{\text{mac}} \rangle \leftarrow \sigma}$ $\text{mac} \leftarrow \text{MAC}_{K^{\text{mac}}}(\tau)$ $\text{return } \langle c, \text{mac} \rangle$	

**Fig. 10.** A TagXEM from a KEM, a MAC, and an XOF  $F$ .

#### 4.6 TagXEM from a KEM, a MAC, and a Random Oracle

With Thm. 8, we achieved what we set out to do: demonstrating tight MI inheritance from a TagXEM to an IND-CCA PKE. However, AGK only showed how to construct an IND-CCA KEM, providing a reduction to the MI-GapCDH assumption in the programmable random oracle model. Without the crucial support of tags, our construction only achieves CPA security. Furthermore, Thm. 5 does *not* easily transfer to the ROP setting: it is not clear how to combine a ROP-CCA KEM with a XOF to yield a ROP-CCA XEM.

We complete the picture by providing a TagXEM construction from a KEM, a MAC, and a XOF. Our construction (Fig. 10) is inspired by Abe et al.'s TagKEM construction [2] and we show that with an information-theoretic MAC, if the KEM is perfectly correct, has unique encapsulations [25] and is multi-instance one-way secure under plaintext-checking attacks (OW-PCA), then the TagXEM is ROP-CCA secure in the programmable random oracle model (to model the XOF). Before stating our concrete security result (Thm. 9), we will define the relevant concepts and advantages below.

**Preliminaries.** One-wayness for KEMs tasks an adversary to retrieve the ephemeral key that has been encapsulated, given the public key and the encapsulation. In the multi-instance setting, an adversary has access to many public keys and various encapsulations per public key and endeavours to find ephemeral keys for encapsulations for as many different public keys as possible (no reward for breaking multiple encapsulations under the same public key).



$\text{Exp}_{\text{KEM}}^{(n,\kappa)\text{-ow-pca}^*}(\mathbb{A})$	$\mathcal{E}(i)$
$(\text{pk}_1, \text{sk}_1), \dots, (\text{pk}_\kappa, \text{sk}_\kappa) \leftarrow_{\$} \text{KEM.Kg}$	$(K, c) \leftarrow_{\$} \text{KEM.Enc}_{\text{pk}_i}$
$(\mathbf{I}, (j_i, \hat{K}_i)_{i \in \mathbf{I}}) \leftarrow_{\$} \mathbb{A}^{\mathcal{E}, \mathcal{P}, \mathcal{K}}(\text{pk}_1, \dots, \text{pk}_\kappa)$	$\mathcal{P}_i \leftarrow \widehat{\ } K$
<b>if</b> $ \mathbf{I}  \neq n \vee \mathbf{I} \cap \mathbb{K} \neq \emptyset$ <b>then return</b> 0	<b>return</b> $c$
<b>return</b> $\bigwedge_{i \in \mathbf{I}} \mathcal{P}_i[j_i] = \hat{K}_i$	$\mathcal{K}(i)$
$\mathcal{P}(i, c, K)$	$\mathbb{K} \leftarrow^{\cup} i$
$K' \leftarrow \text{KEM.Dec}_{\text{sk}_i}(c)$	<b>return</b> $\text{sk}_i$
<b>return</b> $K = K'$	

**Fig. 11.** Multi-instance one-way security in the presence of plaintext checking attacks.

Plaintext-checking attacks (PCA) were introduced by Okamoto and Pointcheval [36, Definition 8] in a single-user public key encryption setting. Intuitively, PCA provides the adversary access to an oracle that, on input a pair  $(m, c)$  determines whether  $c$  encrypts  $m$  or not; more formally [1], the oracle checks whether  $c$  decrypts to  $m$  or not. In the context of KEMs, the PCA oracle takes a pair  $(K^{\text{kem}}, c)$  as input and determines whether  $c$  decapsulates to  $K^{\text{kem}}$  or not. The multi-user or multi-instance generalization is straightforward and the definition (in its modern decryption incarnation) inherently deals with imperfect correctness in the decryption.

Definition 7 considers one-wayness under plaintext checking attacks. For standard ElGamal KEM, where a (multiplicative) discrete-log group with generator  $g$  and of prime order  $q$  is given as part of the parameters, a public key consists of  $h = g^x$  with  $x \leftarrow_{\$} \mathbb{Z}_q$  the private key, and an encapsulation outputs  $(K^{\text{kem}}, c) = (h^r, g^r)$  for random  $r \leftarrow_{\$} \mathbb{Z}_q$ , the one-wayness problem (in the single-user case) is equivalent to the computational Diffie–Hellman (CDH) problem. The plaintext checking oracle allows an adversary to learn, for group elements  $(k, c)$  of its choice, whether  $k = c^x$  or not. The corresponding hardness assumption for OW-PCA is known as the Strong CDH assumption. An even stronger assumption is the GapCDH assumption, where an adversary instead can use an oracle that determines whether a quadruple of group elements is a Diffie–Hellman tuple or not.

**Definition 7 (OW-PCA).** *Let KEM be a key encapsulation mechanism. Then the one-way advantage under plaintext-checking attacks of an adversary  $\mathbb{A}$  is*

$$\text{Adv}_{\text{KEM}}^{(n,\kappa)\text{-ow-pca}^*}(\mathbb{A}) = \Pr \left[ \text{Exp}_{\text{KEM}}^{(n,\kappa)\text{-ow-pca}^*}(\mathbb{A}) = 1 \right],$$

where the experiment is defined in Fig. 11.

In addition to perfect correctness and OW-PCA security, the security reduction for our construction (Thm. 9) relies on two further properties of the underlying KEM. Unique encapsulation captures that for a fixed public key and ephemeral

key, the encapsulation corresponding to that ephemeral key is unique (without saying anything about how to compute it). Unique encapsulations have been used before, for instance by Heuer et al. [25] (see also Remark 4 below).

**Definition 8 (Unique Encapsulation).** *Let KEM be a perfectly correct KEM. Then it has unique encapsulations iff*

$$\Pr \left[ \begin{array}{l} (\text{pk}, \text{sk}) \leftarrow \$ \text{KEM.Kg} \\ (K_0^{kem}, c_0) \leftarrow \$ \text{KEM.Enc}_{\text{pk}} : K_0^{kem} = K_1^{kem} \wedge c_0 \neq c_1 \\ (K_1^{kem}, c_1) \leftarrow \$ \text{KEM.Enc}_{\text{pk}} \end{array} \right] = 0 .$$

The second additional property we require from the KEM is that collisions amongst encapsulations (under a single randomly drawn public key) are suitably rare. Def. 9 captures the relevant probability of a  $k$ -way encapsulation collision. If a KEM is perfectly correct with unique encapsulations, then colliding encapsulations are equivalent to colliding ephemeral keys; if, as is usually the case, these ephemeral keys are furthermore chosen uniformly at random from a finite set  $\mathcal{X}$ , we can upper bound  $\epsilon_k(q)$  by  $q^k/|\mathcal{X}|^{k-1}$  using a standard bound on  $k$ -way collisions (see e.g. [37, Appendix B]).

**Definition 9 (Encapsulation Multi-Collisions).** *Let KEM be a KEM, and let  $q, k \in \mathbb{Z}_{>1}$  be parameters. Then the  $k$ -out-of- $q$  encapsulation multi-collision probability is*

$$\epsilon_k(q) = \Pr \left[ \begin{array}{l} (\text{pk}, \text{sk}) \leftarrow \$ \text{KEM.Kg} \\ \forall_{i \in [q]} (K_i^{kem}, c_i) \leftarrow \$ \text{KEM.Enc}_{\text{pk}} : \exists_{J \subseteq [q], |J|=k} \forall_{i,j \in J} c_i = c_j \end{array} \right] .$$

For completeness, we also present definitions of a deterministic message authentication code, so we dispense with an explicit verification algorithm in Def. 10 (for concreteness, we restrict to bitstrings for both keys and tags, of length  $\ell_{\text{mackey}}$  and  $\ell_{\text{mac}}$  respectively), and an information-theoretic notion of forgeries (Def. 11) where we use the same parameter  $k$  as above (or rather  $k - 1$  in Thm. 9), but this time to denote the number of valid message–tag pairs available to an adversary. The usual choice is  $k = 1$ , e.g. when considering strongly universal<sub>2</sub> hash functions, but Wegman and Carter [40] already investigated  $k > 1$ . Provided  $\ell_{\text{mackey}}$  is large enough (at least  $k \cdot \ell_{\text{mac}}$ ), one can achieve  $\hat{\epsilon}_k = 2^{-\ell_{\text{mac}}}$ , which is optimal.

**Definition 10 (Message Authentication Code (MAC)).** *A message authentication code MAC is a pair of algorithms MAC.Kg and MAC.Mac, where MAC.Kg randomly generates a  $K^{\text{mac}} \in \{0, 1\}^{\ell_{\text{mackey}}}$ , and the deterministic MAC.Mac takes a key  $K^{\text{mac}}$  and a message  $m \in \mathcal{M}$  to output tag  $\text{mac} \leftarrow \text{MAC.Mac}_{K^{\text{mac}}}(m) \in \{0, 1\}^{\ell_{\text{mac}}}$ .*

**Definition 11 (Information-Theoretic MAC Forgeries).** *Let MAC be given and let  $k \in \mathbb{Z}_{\geq 0}$  be a parameter, then the forging advantage after observing  $k$  valid message–tag pairs is defined as*

$$\hat{\epsilon}_k = \max_{\substack{\forall i \in \{0\} \cup [k] \\ (m_i, \text{mac}_i)}} \Pr \left[ \text{MAC.Mac}_{K^{\text{mac}}}(m_0) = \text{mac}_0 \mid \forall_{i \in [k]} \text{MAC.Mac}_{K^{\text{mac}}}(m_i) = \text{mac}_i \right] .$$

**Security Claim.** With all elements in place, we can state the security of Fig. 10's TXEM, in Thm. 9 (see the full version for the proof). The security bound depends on a tuning parameter  $k$  that feeds into both the collision probability of the underlying KEM and the forgery advantage of the MAC, with opposite effects. The ability to tune the bound therefore allows some flexibility when instantiating the three underlying primitives KEM, MAC, and XOF: for fixed  $q_c$ , increasing  $k$  will result in a smaller upper bound on  $\epsilon_k(q_c)$ , but to ensure that  $\hat{\epsilon}_{k-1}$  does not dominate, it might then be necessary to increase the key size  $\ell_{\text{mackey}}$  (and possibly tag size  $\ell_{\text{mac}}$ ) of the information-theoretic MAC (see Cor. 5 for a concrete instantiation). Otherwise, instantiating the information-theoretic MAC and the XOF is relatively straightforward (with the usual ROM caveats for the latter).

**Theorem 9.** *Let TXEM be as in Fig. 10, let KEM be a perfectly correct KEM with unique encapsulations, and let  $k \in \mathbb{Z}_{>1}$ . Then there is an SFBB reduction  $\mathbb{B}$  such that, for all  $\mathbb{A}$  that makes  $q_c$  challenge and  $q_d$  decryption oracle queries,*

$$\text{Adv}_{\text{TXEM}}^{(n,\kappa)\text{-rop-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{KEM}}^{(n,\kappa)\text{-ow-pca}^*}(\mathbb{B}) + 2(q_d \hat{\epsilon}_{k-1} + \epsilon_k(q_c))$$

*in the programmable random oracle model, where  $\hat{\epsilon}_{k-1}$  is the forging advantage after observing  $k - 1$  valid message-tag pairs (Def. 11) and  $\epsilon_k(q_c)$  is the  $k$ -out-of- $q_c$  encapsulation multi-collision probability of KEM (Def. 9). If  $\mathbb{A}$  makes  $q_f$  queries to the random oracle, then  $\mathbb{B}$  makes at most  $q_f$  queries to its plaintext checking oracle.*

The proof borrows some ideas already used to prove AGK's Thm. 2. In fact, it is relatively straightforward to recast AGK's Thm. 2 as the multi-instance version of a OW-PCA KEM plus a programmable random oracle yielding an IND-CCA KEM, although the presence of the error terms  $\hat{\epsilon}_{k-1}$  and especially  $\epsilon_k(q_c)$  render recovery of AGK's Thm. 2 as a special case of our Thm. 9 not immediate.

Combining Thm. 8 and 9 in Cor. 4, we can finally conclude that our construction yields a PKE inheriting the multi-instance security of the underlying KEM (for parameter regimes where the loss term does not dominate).

**Corollary 4.** *Let PKE' be as in Fig. 8, let the underlying TagXEM be as in Fig. 10, let KEM be a perfectly correct KEM with unique encapsulations, and let  $k \in \mathbb{Z}_{>1}$ . Then, there is an SFBB reduction  $\mathbb{B}$  such that, for all  $\mathbb{A}$  that makes  $q_c$  challenge and  $q_d$  decryption oracle queries,*

$$\text{Adv}_{\text{PKE}'}^{(n,\kappa)\text{-ind-cca}^*}(\mathbb{A}) \leq \text{Adv}_{\text{KEM}}^{(n,\kappa)\text{-ow-pca}^*}(\mathbb{B}) + 2(q_d \hat{\epsilon}_{k-1} + \epsilon_k(q_c))$$

*in the programmable random oracle model, where  $\hat{\epsilon}_{k-1}$  is the forging advantage after observing  $k - 1$  valid message-tag pairs (Def. 11) and  $\epsilon_k(q_c)$  is the  $k$ -out-of- $q_c$  encapsulation multi-collision probability of KEM (Def. 9). If  $\mathbb{A}$  makes  $q_f$  queries to the random oracle, then  $\mathbb{B}$  makes at most  $q_f$  queries to its plaintext checking oracle.*

*Remark 4.* The resulting construction is remarkably similar to the PKE studied by Heuer et al. [25] in the context of selective opening attacks (and to a lesser extent its predecessor by Steinfeld et al. [39] and successor by Lai et al. [32]). They too use a random oracle to derive a MAC key and a one-time pad from an ephemeral KEM key. The only two differences are that Heuer et al. do not consider arbitrary length messages and that their random oracle outputs  $K^{\text{xem}}\|K^{\text{mac}}$ , i.e. the opposite order from what we do.

For fixed length messages, the order in which those two keys are output does not matter. However, when moving to arbitrary length-messages, the order of the XOF output does matter. Outputting  $K^{\text{xem}}\|K^{\text{mac}}$  instead would allow a length extension attack enabling the adversary to recover the MAC key, at which point producing forgeries would be trivial.

In a way, the construction is quite brittle that these small details matter. Another example of brittleness is that our reduction for Theorem 9 requires  $\perp$  produced from a KEM decryption error to be indistinguishable from a failed MAC verification. In implementations, a timing attack might well break this requirement.

*Remark 5.* The proof of Thm. 9 does rely on perfect correctness of the underlying KEM, thus excluding many popular post-quantum KEMs based on the hardness of LWE. Having said that, establishing the post-quantum security of TXEM would require a proof in the quantum random oracle model [15]. We leave the construction of a post-quantum TagXEM as an enticing open problem.

**A Concrete Instantiation.** We conclude by providing a concrete bound for the construction when instantiating with low granularity ElGamal KEM on groups of size  $\geq p$ . ElGamal KEM satisfies perfect correctness and unique encapsulation (ensuring compatibility with Thm. 9) and produces uniformly random group elements as ephemeral keys, so  $\epsilon_k(q_c) \leq q^k/p^{k-1}$ . Furthermore, the relevant multi-instance OW-PCA security can be linked to the low granularity MI-GapCDH problem with corruptions (Thm. 12 of the full version). By extending AGK’s low granularity bound [4, Thm. 6] to include corruptions (Thm. 11 of the full version) and combining with Cor. 4, we arrive at a clean information-theoretic bound (Cor. 5) in the generic group and programmable random oracle model. To keep the bound easier to interpret, we assume that the adversary makes at most  $\sqrt{p}$  queries to the encryption and decryption oracles; realistically, an adversary will be able to make far more offline queries  $q$  to its generic group and for  $q \approx \sqrt{p}$  a single discrete logarithm instance can already be broken. In a similar vein, the requirement that each group instance receive at least  $\max\{60 \log_2 p, \sqrt{qT}/2\}$  group operation calls (allowing some simplifications in the MI-GapCDH bound) is a reasonable one, as already argued by AGK, given that the number of group operations performed by an ElGamal adversary is “typically large”.

**Corollary 5.** *Let PKE’ be as in Fig. 8, let the underlying TagXEM be as in Fig 10, let KEM be instantiated as low granularity ElGamal (see the full version for details) and let  $p$  be a lower bound on the generated groups. Let  $k \in \mathbb{Z}_{>1}$ ,*

let MAC be an information-theoretic MAC with key length  $\ell_{\text{mackey}}$  and output length  $\ell_{\text{mac}}$  and satisfying  $\hat{\epsilon}_{k-1} = 2^{-\ell_{\text{mac}}}$ . Then, for any information-theoretic  $\mathbb{A}$  that makes at most  $\sqrt{p}$  challenge oracle queries, at most  $\sqrt{p}$  decryption oracle queries,  $q_f$  queries to the random oracle, and a total of  $q$  queries to the group-operation oracles with at least  $\max\{60 \log_2 p, \sqrt{q_f}/2\}$  queries per group instance, it holds that

$$\text{Adv}_{\text{PKE}'}^{(n,\kappa)\text{-ind-cca}^\star}(\mathbb{A}) \leq \left(\frac{4 \cdot e \cdot q^2}{n^2 \cdot p}\right)^n + 2 \left(\frac{\sqrt{p}}{2^{\ell_{\text{mac}}}} + \frac{1}{p^{\frac{k}{2}-1}}\right)$$

in the programmable random oracle and generic group model.

For the construction to exhibit meaningful multi-instance security, we want the upper bound on the adversary’s advantage to diminish with increasing  $n$ . Since the second term on the right hand side of Cor. 5 is independent of  $n$ , the first term has to dominate for advantages of interest. Thus, for a fixed  $p$ , we want to set  $\ell_{\text{mac}}$  and  $k$  so that, irrespective of  $n$ , we do not really care about the other two terms, where  $\ell_{\text{mac}}$  directly corresponds to the PKE’s ciphertext expansion and increasing  $k$  will require longer ephemeral keys as output by the XOF to ensure that  $\ell_{\text{mackey}} \geq k \cdot \ell_{\text{mac}}$ . To minimize overhead, having both terms equal is optimal, corresponding to  $2\ell_{\text{mac}} = (k-1) \log_2 p$ . Some reasonable options are then  $(\ell_{\text{mac}}, k) = (\log_2 p, 3)$  or  $(\ell_{\text{mac}}, k) = (3/2 \log_2 p, 4)$ .

Alternatively, the bound can be interpreted in terms of the scaling factor, which focuses on the minimum resources needed to achieve an overwhelming advantage (see the full version for details). In that case, the second term, being independent of  $n$ , is manifestly of little interest for either of our suggested parameter choices.

## References

1. Abdalla, M., Benhamouda, F., Pointcheval, D.: Public-key encryption indistinguishable under plaintext-checkable attacks. In: Katz, J. (ed.) PKC 2015. LNCS, vol. 9020, pp. 332–352. Springer, Heidelberg (Mar / Apr 2015). [https://doi.org/10.1007/978-3-662-46447-2\\_15](https://doi.org/10.1007/978-3-662-46447-2_15)
2. Abe, M., Gennaro, R., Kurosawa, K.: Tag-KEM/DEM: A new framework for hybrid encryption. *Journal of Cryptology* **21**(1), 97–130 (Jan 2008). <https://doi.org/10.1007/s00145-007-9010-x>
3. Auerbach, B., Giacon, F., Kiltz, E.: Everybody’s a target: Scalability in public-key encryption. *Cryptology ePrint Archive, Report 2019/364* (2019), <https://eprint.iacr.org/2019/364>
4. Auerbach, B., Giacon, F., Kiltz, E.: Everybody’s a target: Scalability in public-key encryption. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020, Part III. LNCS, vol. 12107, pp. 475–506. Springer, Heidelberg (May 2020). [https://doi.org/10.1007/978-3-030-45727-3\\_16](https://doi.org/10.1007/978-3-030-45727-3_16)
5. Bader, C., Jager, T., Li, Y., Schäge, S.: On the impossibility of tight cryptographic reductions. In: Fischlin, M., Coron, J.S. (eds.) EUROCRYPT 2016, Part II. LNCS, vol. 9666, pp. 273–304. Springer, Heidelberg (May 2016). [https://doi.org/10.1007/978-3-662-49896-5\\_10](https://doi.org/10.1007/978-3-662-49896-5_10)

6. Baecher, P., Brzuska, C., Fischlin, M.: Notions of black-box reductions, revisited. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013, Part I. LNCS, vol. 8269, pp. 296–315. Springer, Heidelberg (Dec 2013). [https://doi.org/10.1007/978-3-642-42033-7\\_16](https://doi.org/10.1007/978-3-642-42033-7_16)
7. Bellare, M., Boldyreva, A., Micali, S.: Public-key encryption in a multi-user setting: Security proofs and improvements. In: Preneel, B. (ed.) EUROCRYPT 2000. LNCS, vol. 1807, pp. 259–274. Springer, Heidelberg (May 2000). [https://doi.org/10.1007/3-540-45539-6\\_18](https://doi.org/10.1007/3-540-45539-6_18)
8. Bellare, M., Desai, A., Pointcheval, D., Rogaway, P.: Relations among notions of security for public-key encryption schemes. In: Krawczyk, H. (ed.) CRYPTO'98. LNCS, vol. 1462, pp. 26–45. Springer, Heidelberg (Aug 1998). <https://doi.org/10.1007/BFb0055718>
9. Bellare, M., Hofheinz, D., Kiltz, E.: Subtleties in the definition of IND-CCA: When and how should challenge decryption be disallowed? *Journal of Cryptology* **28**(1), 29–48 (Jan 2015). <https://doi.org/10.1007/s00145-013-9167-4>
10. Bellare, M., Kohno, T.: A theoretical treatment of related-key attacks: RKA-PRPs, RKA-PRFs, and applications. In: Biham, E. (ed.) EUROCRYPT 2003. LNCS, vol. 2656, pp. 491–506. Springer, Heidelberg (May 2003). [https://doi.org/10.1007/3-540-39200-9\\_31](https://doi.org/10.1007/3-540-39200-9_31)
11. Bellare, M., Paterson, K.G., Rogaway, P.: Security of symmetric encryption against mass surveillance. In: Garay, J.A., Gennaro, R. (eds.) CRYPTO 2014, Part I. LNCS, vol. 8616, pp. 1–19. Springer, Heidelberg (Aug 2014). [https://doi.org/10.1007/978-3-662-44371-2\\_1](https://doi.org/10.1007/978-3-662-44371-2_1)
12. Bellare, M., Ristenpart, T., Tessaro, S.: Multi-instance security and its application to password-based cryptography. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 312–329. Springer, Heidelberg (Aug 2012). [https://doi.org/10.1007/978-3-642-32009-5\\_19](https://doi.org/10.1007/978-3-642-32009-5_19)
13. Bellare, M., Ristenpart, T., Tessaro, S.: Multi-instance security and its application to password-based cryptography. *Cryptology ePrint Archive*, Report 2012/196 (2012), <https://eprint.iacr.org/2012/196>
14. Bellare, M., Rogaway, P.: Random oracles are practical: A paradigm for designing efficient protocols. In: Denning, D.E., Pyle, R., Ganesan, R., Sandhu, R.S., Ashby, V. (eds.) ACM CCS 93. pp. 62–73. ACM Press (Nov 1993). <https://doi.org/10.1145/168588.168596>
15. Boneh, D., Dagdelen, Ö., Fischlin, M., Lehmann, A., Schaffner, C., Zhandry, M.: Random oracles in a quantum world. In: Lee, D.H., Wang, X. (eds.) ASIACRYPT 2011. LNCS, vol. 7073, pp. 41–69. Springer, Heidelberg (Dec 2011). [https://doi.org/10.1007/978-3-642-25385-0\\_3](https://doi.org/10.1007/978-3-642-25385-0_3)
16. Bos, J., Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schanck, J.M., Schwabe, P., Seiler, G., Stehle, D.: Crystals - kyber: A cca-secure module-lattice-based kem. In: 2018 IEEE European Symposium on Security and Privacy (EuroSP). pp. 353–367 (2018). <https://doi.org/10.1109/EuroSP.2018.00032>
17. Brunetta, C., Heum, H., Stam, M.: Multi-instance secure public-key encryption. *Cryptology ePrint Archive*, Report 2022/909 (2022), <https://eprint.iacr.org/2022/909>
18. Cramer, R., Shoup, V.: A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In: Krawczyk, H. (ed.) CRYPTO'98. LNCS, vol. 1462, pp. 13–25. Springer, Heidelberg (Aug 1998). <https://doi.org/10.1007/BFb0055717>

19. Cramer, R., Shoup, V.: Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack. *SIAM Journal on Computing* **33**(1), 167–226 (2003)
20. Farshim, P., Tessaro, S.: Password hashing and preprocessing. In: Canteaut, A., Standaert, F.X. (eds.) *EUROCRYPT 2021, Part II*. LNCS, vol. 12697, pp. 64–91. Springer, Heidelberg (Oct 2021). [https://doi.org/10.1007/978-3-030-77886-6\\_3](https://doi.org/10.1007/978-3-030-77886-6_3)
21. Fujisaki, E., Okamoto, T.: Secure integration of asymmetric and symmetric encryption schemes. In: Wiener, M.J. (ed.) *CRYPTO'99*. LNCS, vol. 1666, pp. 537–554. Springer, Heidelberg (Aug 1999). [https://doi.org/10.1007/3-540-48405-1\\_34](https://doi.org/10.1007/3-540-48405-1_34)
22. Fujisaki, E., Okamoto, T.: Secure integration of asymmetric and symmetric encryption schemes. *Journal of Cryptology* **26**(1), 80–101 (Jan 2013). <https://doi.org/10.1007/s00145-011-9114-1>
23. Giacon, F., Kiltz, E., Poettering, B.: Hybrid encryption in a multi-user setting, revisited. In: Abdalla, M., Dahab, R. (eds.) *PKC 2018, Part I*. LNCS, vol. 10769, pp. 159–189. Springer, Heidelberg (Mar 2018). [https://doi.org/10.1007/978-3-319-76578-5\\_6](https://doi.org/10.1007/978-3-319-76578-5_6)
24. Halevi, S., Krawczyk, H.: Security under key-dependent inputs. In: Ning, P., De Capitani di Vimercati, S., Syverson, P.F. (eds.) *ACM CCS 2007*. pp. 466–475. ACM Press (Oct 2007). <https://doi.org/10.1145/1315245.1315303>
25. Heuer, F., Jager, T., Kiltz, E., Schäge, S.: On the selective opening security of practical public-key encryption schemes. In: Katz, J. (ed.) *PKC 2015*. LNCS, vol. 9020, pp. 27–51. Springer, Heidelberg (Mar / Apr 2015). [https://doi.org/10.1007/978-3-662-46447-2\\_2](https://doi.org/10.1007/978-3-662-46447-2_2)
26. Heum, H., Stam, M.: Tightness subtleties for multi-user pke notions. In: Paterson, M.B. (ed.) *Cryptography and Coding*. pp. 75–104. Springer International Publishing, Cham (2021). [https://doi.org/10.1007/978-3-030-92641-0\\_5](https://doi.org/10.1007/978-3-030-92641-0_5)
27. Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58**, 13–30 (1963)
28. Hofheinz, D., Hövelmanns, K., Kiltz, E.: A modular analysis of the Fujisaki-Okamoto transformation. In: Kalai, Y., Reyzin, L. (eds.) *TCC 2017, Part I*. LNCS, vol. 10677, pp. 341–371. Springer, Heidelberg (Nov 2017). [https://doi.org/10.1007/978-3-319-70500-2\\_12](https://doi.org/10.1007/978-3-319-70500-2_12)
29. Jager, T., Kiltz, E., Riepel, D., Schäge, S.: Tightly-secure authenticated key exchange, revisited. In: Canteaut, A., Standaert, F.X. (eds.) *EUROCRYPT 2021, Part I*. LNCS, vol. 12696, pp. 117–146. Springer, Heidelberg (Oct 2021). [https://doi.org/10.1007/978-3-030-77870-5\\_5](https://doi.org/10.1007/978-3-030-77870-5_5)
30. Kiltz, E., Pan, J., Riepel, D., Ringerud, M.: Multi-user CDH problems and the concrete security of NAXOS and HMQV. In: Rosulek, M. (ed.) *CT-RSA 2023* (to appear). Springer, Heidelberg (2023), available as <https://eprint.iacr.org/2023/115>.
31. Kurosawa, K., Desmedt, Y.: A new paradigm of hybrid encryption scheme. In: Franklin, M. (ed.) *CRYPTO 2004*. LNCS, vol. 3152, pp. 426–442. Springer, Heidelberg (Aug 2004). [https://doi.org/10.1007/978-3-540-28628-8\\_26](https://doi.org/10.1007/978-3-540-28628-8_26)
32. Lai, J., Yang, R., Huang, Z., Weng, J.: Simulation-based bi-selective opening security for public key encryption. In: Tibouchi, M., Wang, H. (eds.) *ASIACRYPT 2021, Part II*. LNCS, vol. 13091, pp. 456–482. Springer, Heidelberg (Dec 2021). [https://doi.org/10.1007/978-3-030-92075-3\\_16](https://doi.org/10.1007/978-3-030-92075-3_16)
33. Lee, Y., Lee, D.H., Park, J.H.: Tightly cca-secure encryption scheme in a multi-user setting with corruptions. *Des. Codes Cryptogr.* **88**(11), 2433–2452 (2020)

34. Lewko, A.B., Waters, B.: Why proving HIBE systems secure is difficult. In: Nguyen, P.Q., Oswald, E. (eds.) EUROCRYPT 2014. LNCS, vol. 8441, pp. 58–76. Springer, Heidelberg (May 2014). [https://doi.org/10.1007/978-3-642-55220-5\\_4](https://doi.org/10.1007/978-3-642-55220-5_4)
35. NIST: SHA-3 standard: Permutation-based hash and extendable-output functions. Federal Information Processing Standards Publication 202, NIST (Aug 2015)
36. Okamoto, T., Pointcheval, D.: REACT: Rapid Enhanced-security Asymmetric Cryptosystem Transform. In: Naccache, D. (ed.) CT-RSA 2001. LNCS, vol. 2020, pp. 159–175. Springer, Heidelberg (Apr 2001). [https://doi.org/10.1007/3-540-45353-9\\_13](https://doi.org/10.1007/3-540-45353-9_13)
37. Preneel, B.: Analysis and Design of Cryptographic Hash Functions. Ph.D. thesis, KU Leuven (Feb 1993)
38. Reingold, O., Trevisan, L., Vadhan, S.P.: Notions of reducibility between cryptographic primitives. In: Naor, M. (ed.) TCC 2004. LNCS, vol. 2951, pp. 1–20. Springer, Heidelberg (Feb 2004). [https://doi.org/10.1007/978-3-540-24638-1\\_1](https://doi.org/10.1007/978-3-540-24638-1_1)
39. Steinfeld, R., Baek, J., Zheng, Y.: On the necessity of strong assumptions for the security of a class of asymmetric encryption schemes. In: Batten, L.M., Seberry, J. (eds.) ACISP 02. LNCS, vol. 2384, pp. 241–256. Springer, Heidelberg (Jul 2002). [https://doi.org/10.1007/3-540-45450-0\\_20](https://doi.org/10.1007/3-540-45450-0_20)
40. Wegman, M.N., Carter, L.: New hash functions and their use in authentication and set equality. *Journal of Computer and System Sciences* **22**, 265–279 (1981)