On Message Integrity in Symmetric Encryption

Virgil D. Gligor* Pompiliu Donescu

VDG Inc
6009 Brookside Drive
Chevy Chase, Maryland 20815
{gligor, pompiliu}@eng.umd.edu

November 10, 2000

Abstract

Distinct notions of message integrity (authenticity) for block-oriented symmetric encryption are defined by integrity goals to be achieved in the face of different types of attacks. These notions are partially ordered by a "dominance" relation. When chosen-plaintext attacks are considered, most integrity goals form a lattice. The lattice is extended when known-plaintext and ciphertext-only attacks are also included. The practical use of the dominance relation and lattice in defining the relative strength of different integrity notions is illustrated with common modes of encryption, such as the "infinite garble extension" modes, and simple, non-cryptographic, manipulation detection code functions, such as bitwise exclusive-or and constant functions.

1 Introduction

The fact that encryption does not provide message integrity (authenticity) is generally well-understood [19], and so is the fact that often "encryption without integrity-checking is all but useless" [8]. Less well-understood is the fact that message integrity depends intimately on the protection goals of the application environment and the operational threats posed by that environment. Ignoring this fact may lead to performance and usability mismatches. For example, many embedded, low-power, systems and applications can hardly afford to use any of the traditional hash functions or message authentication codes proposed to date to maintain the integrity of encrypted messages, particularly in environments exposed only to limited-scope attacks (e.g., ciphertext-only attacks).

We explore different notions of message integrity for block-oriented symmetric encryption and their relationships. These notions are expressed as a combination of integrity goals to be achieved in the face of different types of attacks, as originally suggested by Naor (viz., attribution [4]). The set of all integrity goals include both goals known to date, such as protection against existential forgery and assurance of plaintext integrity and of non-malleability, and new ones, such as maintenance of plaintext uncertainty, and protection against known- and chosen-plaintext forgery. Attack models include chosen-plaintext, known-plaintext, ciphertext-only, and chosen-ciphertext attacks. The integrity notions defined are partially ordered by a "dominance"

*This work was performed in part while this author was on sabbatical leave from the University of Maryland, Department of Electrical and Computer Engineering, College Park, Maryland 20742.

relation. When chosen-plaintext attacks (CPAs) are considered, most integrity goals form a lattice. This lattice is extended by the inclusion of ciphertext-only attacks (CoAs). Although we do not explicitly show it, the lattice can also be extended by the inclusion of known-plaintext attacks (KPAs). The resulting lattice shows that the strongest notion of integrity is provided by existential forgeries in CPAs and the weakest by chosen-plaintext forgeries in CoAs.

Defining notions of integrity in terms of a "dominance" relation enables us to characterize the relative strength of various symmetric encryption modes precisely. The utility of such characterization extends beyond theory; e.g., it enables us to explore the space of encryption schemes (modes) that can be composed with a variety of Manipulation Detection Code (MDC) functions (e.g., non-cryptographic MDCs such as bitwise exclusive-or, cyclic redundancy code, and even constant, functions), and used in a variety of application environments exposed to well-defined threats. As an example of schemes whose relative strength can be precisely evaluated, we analyze Campbell's "infinite garble extension" (IGE) mode of encryption [9].

The balance of this paper is organized as follows. Section 2 contains some preliminary definitions and notation, Section 3 contains the definition of the integrity notions addressed in this paper. Section 4 contains the relations among integrity notions (i.e., dominance, incomparability, separation) based on the definition of goals and attacks, and the integrity lattice and its extensions. Section 5 contains the lemmas that help characterize the integrity properties of IGE modes when used with very simple manipulation detection code (MDC) functions, and examples of other modes that are vulnerable with respect to different integrity notions when composed with specific MDC functions.

2 Background

In defining the relationships between different notion of integrity for symmetric encryption, we will use encryption modes by the triple $\Pi = (E, D, KG)$, where E is the message encryption function, D is the message decryption function, and KG is the probabilistic key-generation algorithm. These encryption modes are implemented with block ciphers, which can be modeled with finite families of pseudorandom functions (PRFs). A detailed account for the use of such functions in symmetric encryption modes intended to satisfy secrecy goals is provided by Bellare *et al.* [2]. Since most practical encryption schemes use both the encryption and decryption functions of block ciphers, a natural way to model such ciphers is with finite families of super-pseudorandom permutations (SPRPs) [18]. We denote both PRFs and SPRPs by F below and distinguish which we mean in context.

Perhaps the most common method used to detect modifications of encrypted messages applies a MDC function g to a plaintext message and concatenates the result with the plaintext before encryption with E. The choice of MDC function g is entirely that of the designer; e.g., g could be a non-keyed hash, cyclic redundancy code (CRC), bitwise exclusive-or, or even a constant, function [19]. A message thus encrypted can be decrypted and accepted as valid only after the integrity check passes; i.e., after decryption with D, the concatenated value of function g is removed from the plaintext, and the check passes only if this value matches that obtained by applying the MDC function to the remaining plaintext [22, 21, 19]. If the integrity check does not pass, a special failure indicator, denoted by $Null$ herein, is returned.[1] The encryption scheme obtained by using this method is denoted by $\Pi \ o \ g = (E \ o \ g, D \ o \ g, KG)$, where Π is said to be *composed* with the MDC function g. In this mode, we denote the use of the key K in the encryption

[1] This method has been used in commercial systems such as Kerberos V5 [22, 23] and DCE [21, 23], among many others. Note that other methods for protecting the integrity of encrypted messages exist; e.g., encrypting the message with a secret key and then taking the keyed MAC of the ciphertext with a separate key [19, 7].

of a plaintext string x by $(E^{F_K}\ o\ g)(x)$, and in the decryption of ciphertext string y by $(D^{F_K}\ o\ g)(y)$. The passing of the integrity check at decryption is denoted by $(D^{F_K}\ o\ g)(y) \neq Null$.

For any key K, a forgery is any ciphertext message that is not the output of $E^{F_K} o\ g$. A "valid" forgery is a forgery that passes the integrity check. Forgeries can be created in many ways, for example (1) by modifying the ciphertexts of legitimate messages whose plaintext may be known by the forgerer, (2) by including arbitrary, never-seen-before, strings into existing ciphertexts, or (3) by combinations of the two. Ciphertexts of legitimate message encryptions can be obtained as a result of different attack scenarios, such as chosen-plaintext attacks (CPA) or ciphertext-only attacks (CoA).

All attacks considered in this paper are characterized by q_e message encryptions by $(E^{F_K} o\ g)$, whose plaintext input may   may not be chosen by, or known, to an adversary, and q_v forgery verifications; i.e., decryptions by $(D^{F_K} o\ g)$ performed by an adversary. The encryptions and decryption total $\mu_e + \mu_v$ bits, and take time $t_e + t_v$. Note that parameters q_e, μ_e, t_e can be bound by the parameters defining the chosen-plaintext security of Π = (E,D,KG) mode in some well-defined sense. (One, but not the only, way to define these bounds is to use the notion of security in the left-or-right sense for adaptive chosen-plaintext attacks [2]). In contrast, parameters $q_e, \mu_e, t_e, q_v, \mu_v, t_v$ are bound by the parameters of the function family F and by the desired probability of adversary's success. Note that $q_v > 0$ since the adversary must be allowed verification queries. Otherwise, the adversary cannot test whether his forgeries are correct, since he does not know key K. For the purposes of this paper, it is sufficient that $q_v = 1$; for other purposes, such as determining the attack complexity and general bounds, q_v may take on other values.

3 Message Integrity Notions

3.1 Goals

We define new integrity goals and interpret extant ones, in the context of $\Pi\ o\ g$ modes of encryption. However, it should be clear that the same goals can be defined in the context of other modes that aim at protecting the integrity (authenticity) of encrypted message, such as those that compute the keyed MAC of a message using a secret key and encrypting the message with a separate secret key [19, 7].

The strongest known goal for message integrity is that of protection against *existential forgery (EF)*. This goal has also been known as *existential unforgeability* [15] and *integrity of ciphertext* [7]. To defeat this goal, an adversary only needs to find a "valid" forgery. Knowledge or choice of the plaintext outcome of the forgery is unnecessary to achieve this goal. Formally, an encryption scheme or mode $\Pi\ o\ g$ is secure against existential-forgeries if, for any forgery y,

$$Pr[(D^{F_K} o\ g)(y) \neq Null] \leq \epsilon,$$

where ϵ is a negligible quantity. Throughout this paper, negligibility is used in the traditional sense [2, 20]. In addition to protection against EF goal, two other goals have been defined that have direct applicability to message integrity, namely maintenance of *plaintext integrity (PI)* [7] and assurance *non-malleability (NM)* [10, 4, 15, 7].

The goal of *plaintext integrity* (PI) requires it be infeasible for an adversary to create a "valid" forgery whose decryption is a plaintext not seen before. Formally, an encryption scheme or mode $\Pi\ o\ g$ is secure in the sense of PI if:

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \leq i \leq q_e] \leq \epsilon,$$

where $x^i, 1 \quad i \quad q_e$, are plaintext strings used in encryption and ϵ is a negligible quantity.

The goal of *non-malleability* (NM) formalizes the adversary's inability to create "valid" forgeries that are "meaningfully related" to the unknown plaintext strings corresponding to challenge ciphertext messages. Our interpretation of non-malleability is as follows. Let q_2 be the number of challenge ciphertexts of equal length intercepted by an adversary (i.e., the q_2 plaintexts of the intercepted ciphertexts remain unknown to the adversary). Formally, we say that an encryption scheme $\Pi\ o\ g$ is non-malleable (NM) if, for any message length m and challenge ciphertexts $y^1, \cdot\ \cdot, y^{q_2}$ of unknown plaintext messages $x^1, \cdot\ \cdot, x^{q_2} \in \{0,1\}^m$, and for any forgery $y \neq y^i, 1 \quad i \quad q_2$ and any relationship $\mathcal{R}$,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \cdot\ \cdot, x^{q_2}, (D^{F_K} o\ g)(y))] \quad \epsilon,$$

where ϵ is a negligible quantity.

We define two additional integrity goals for valid forgeries, namely protection against *chosen-plaintext forgery* (CPF), and assurance of *plaintext-uncertainty (PU)*. The rationale for these goals can be summarized as follows. Since different plaintext outcomes of a valid forgery can restrict an adversary's ability to take advantage of forgery to different degrees, it is sensible to examine a variety of constraints placed on these outcomes [12]. Such constraints, which were used to define the NM and PI goals above, can lead to new integrity goals and notions, further refining the integrity design space.

The goal of *chosen-plaintext forgery* (CPF) formalizes the adversary's inability to create a "valid" forgery whose plaintext outcome is an a priori "chosen" challenge for the adversary. In our model, the challenge plaintext string is considered to be "chosen," if every block of the string has a specific value determined prior to the attack. Hence, a plaintext string x is *not chosen* if there is at least a block x_i such that given a *specific* constant a, $Pr[x_i = a] \quad \epsilon$, where ϵ is a negligible quantity. Formally, an encryption scheme $\Pi\ o\ g$ is secure against chosen-plaintext forgeries if, for an a priori chosen challenge x and any forgery y,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is chosen }] \quad \epsilon,$$

where $x^i, 1 \quad i \quad q_e$, are plaintext strings used in encryption and ϵ is a negligible quantity.

The goal of *plaintext uncertainty* (PU) formalizes the adversary's inability to create a "valid" forgery for which the adversary "knows" the underlying plaintext. In our model, a plaintext string x is *unknown* if there is at least a block x_i such that for *any* chosen constant a, $Pr[x_i = a] \quad \epsilon$, where ϵ is a negligible quantity. A plaintext string x is "known" if every block of the string is known. Formally,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \ \Rightarrow\ (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown }] \geq \lambda,$$

where $x^i, 1 \quad i \quad q_e$, are plaintext strings used in encryption and $1 \Leftrightarrow \lambda$ is a negligible quantity.

However, if one takes the view that *any* constraint placed on valid forgeries can be a legitimate integrity goal then, among the additional distinct goals made possible, some may be counterintuitive from an integrity point of view. For example, the goal of *known-plaintext forgery* (KPF) formalizes the adversary's inability to create a "valid" forgery without "knowing" the underlying plaintext.[2] Formally, an encryption scheme $\Pi\ o\ g$ is secure against know-plaintext forgeries if, for any forgery y,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \ \Rightarrow\ (D^{F_K} o\ g)(y) = x \text{ is known }] \geq \lambda,$$

where $1 \Leftrightarrow \lambda$ is a negligible quantity. Security notions using this goal can be related to other integrity notions (e.g., PI-CPA "dominates" KPF-CPA and KPF-CPA is incomparable or separated from other notions, as

[2] This goal is somewhat similar to the goal of "plaintext awareness" [1, 4], except that it is independent of the random-oracle model.

shown in Section 4.4 below). Yet, this goal seems to lack an intuitive justification for possible integrity relevance.

Note that if other constraints of "known/unknown" and "chosen/not chosen" plaintext outcomes for valid forgeries that differ from the ones above are defined, other integrity goals may be obtained. Regardless of the definition chosen, the implication ($x =$ is chosen) ($x =$ is known), and equivalently, ($x =$ is unknown) ($x =$ is not chosen), must hold.

3.2 Goal – Attack Combinations

The first attack model considered here is the *chosen-plaintext attack (CPA)*. In a CPA, an adversary can obtain samples of valid encryptions for plaintext messages of his choice even though the secret encryption key remains unknown to the adversary. In this paper, we assume that the adversary obtains the ciphertext for all his chosen plaintext *before* submitting any of his forgeries for verification (decryption).[3]. This does not represent a restriction of the adversary's power, since it can be shown that the advantage of such an adversary in breaking the integrity of a scheme is at least as high as that of an adversary that is allowed to intersperse encryptions of chosen plaintext with forgery verifications [13, 15]. Although CPAs might appear to be mostly of theoretical interest, they are actually quite practical [23, 24]. In fact, these are some of the oldest known attacks in modern cryptography (viz., the "gardening" attacks of British cryptographers during WWII [14]).

In addition to CPA models, we consider *ciphertext-only attack (CoA)* models; i.e., attacks in which the adversary knows the ciphertext corresponding to plaintext strings encrypted with an unknown key, but does not know the plaintexts strings; i.e., the plaintext strings are random, uniformly distributed and independent of each other. (More general definition for CoA whereby the distribution of the plaintext strings is known is also possible.) In this type of attacks, the adversary can make up his forgeries based on ciphertext of valid but unknown plaintext. These attacks can be mounted very easily in practice since they imply that the adversary only needs to eavesdrop on communication between legitimate parties to obtain the desired ciphertext, which is intuitively easier than obtaining encryptions of chosen plaintext. A stronger attack than CoA but weaker than CPA is the *known-plaintext attack (KPA)*. In this attack model, the adversary is assumed to "know" the entire message plaintext not just its corresponding ciphertext, but cannot choose the plaintext.

Other types of attack models may be used for specific problems that include both secrecy and integrity goals; e.g., *chosen-ciphertext attacks* (CCAs), which often appear in entity authentication and key exchange protocols. Bellare *et al.* [4] use these attack models in establishing relationships among different security notions in asymmetric encryption, and suggest that these relationships among their goal-attack combinations also hold for the symmetric case. Katz and Yung [15] illustrate conditions under which such relationships hold in symmetric encryption. In this paper, we do not address these types of goal-attack combinations. However, we suggest that most goals that are combined with CCAs can be represented within the integrity lattice defined in this paper. From an integrity point of view, such attacks are not stronger than CPAs.

For most goals and attacks, the combination an integrity goal with an attack is straight forward. However, some combinations require care to ensure that specific goals and attacks can be paired. For instance, a question may arise as to whether a goal is or is not satisfied at the end of an attack. More specifically,

[3]In this attack, the adversary can be given an oracle that performs all the q_e encryption queries before all the q_v forgery verification queries. Alternatively, the adversary can be given an encryption-only oracle whose use preceds that of a forgery-verification oracle, the order of use being enforced by a state variable

how can an adversary determine whether he actually "knows" the plaintext outcome of his valid forgery ? In practice, it is sometimes the case that the plaintext outcome is not, or cannot be, returned to the adversary. In such cases, we need to add a "plaintext-outcome extractor" to the definition of the goal-attack combination that plays much the same role as the "plaintext extractor" in the plaintext-awareness definition. Practical examples of plaintext-outcome extractors are available for specific integrity goals defined for Π o g schemes and attacks. For instance, the plaintext outcome extractors for the KPF goal defined for the example schemes of Section 5.3 and CPAs, can be easily derived using the equations of "message splicing and decomposition" invariant of CBC [23] and PCBC modes and simple properties of bitwise exclusive-or functions.

Care must also be exercised in defining goal-attack combinations whenever a specific goal already includes elements of an attack. For instance, in the NM-CPA combination, the definition of the NM goal already includes some elements of a CoA model; i.e., the ciphertext challenges. For the NM-CPA combination, we allow the adversary to encrypt q_1 plaintext strings whose ciphertext have the same length as that of the challenge ciphertexts. The adversary can issue its encryption queries at any time; e.g., even after he has seen the challenge ciphertext strings. Furthermore, we require that $q_2 > 0, q_1 + q_2 = q_e$, where q_e is the total number of queries that can be encrypted by $E^{F_K} o\ g$, and that forgery y differs from any of the ciphertexts obtained as a result of the q_1 chosen-plaintext encryptions. For the NM-CoA combination, we simply set $q_1 = 0$, thereby removing the adversary's ability to encrypt with $E^{F_K} o\ g$; i.e., encrypt with the same key as that used to generate the q_2 challenge ciphertexts.

Note that combinations of CPA attacks with challenge ciphertexts, as suggested by the NM-CPA attack combination, are fairly common in distributed applications [23]. For example, consider a distributed service that uses a shared key for encrypting messages between two of its components services, S1 and S2. The adversary is one of the legitimate clients of the distributed service, and can obtain q_1 ciphertext messages corresponding to its own chosen plaintext submitted to S1 by eavesdropping on the communication line between S1 and S2. Similarly, the adversary can obtain the q_2 (challenge) ciphertexts produced by the encryption of other clients' plaintexts that remains unknown to the adversary. The distributed service changes the shared key after q_e encryptions performed on behalf of its client, totaling μ_e bits, and taking t_e time.

4 Relationships Among Integrity Notions

The *dominance* relation between integrity notions A and B, denoted by $A > B$, is defined as follows: $A > B$ if $A \quad B$ and $B \not\Rightarrow A$, where $A \quad B$ means that a scheme (mode) that is secure for notion A is also secure for notion B; and $B \not\Rightarrow A$ means that not all schemes that are secure for notion B are secure in notion A (i.e., notions B and A are *separable*). Integrity notions A and B are *incomparable* if $A \not\Rightarrow B$ and $B \not\Rightarrow A$, and *equivalent* if $A \quad B$ and $B \quad A$. These relations have also been used by Katz and Yung [15] for different security notions in symmetric encryption (i.e., indistinguishability and non-malleability in different types of attacks). The relations of implication ( ) and separability ($\not\Rightarrow$) were originally introduced by Bellare *et al.* for security notions in asymmetric encryption, and used later for some integrity notions in symmetric encryption [4, 7].

In proving the dominance, incomparability, and separation relations between different notions of integrity, we use (1) integrity goal definitions, for the (simple) $A \quad B$ proofs, and (2) specific Π o g modes, to provide the necessary counter-examples for $B \not\Rightarrow A$ proofs.

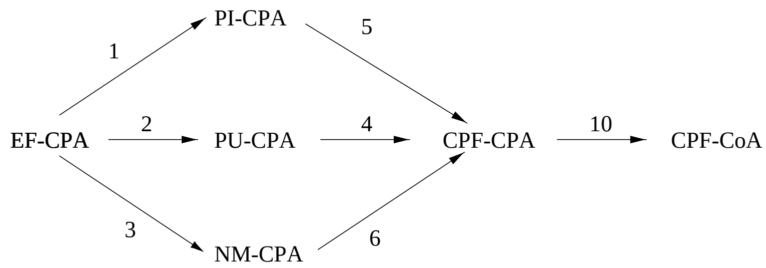

Figure 1: An arrow represents a "dominance" relation (>), and there is a path from **A** to **B** if and only if **A** > **B**. Lack of an arrow and path between two notions indicates incomparable or separated notions. The number on an arrow represents the theorem number that establishes this relationship.

4.1 Dominance

Theorem 1: EF-ATK > PI-ATK
Proof
(1) EF-ATK PI-ATK.
An encryption scheme (mode) that is secure against existential forgeries (EFs) in an attack (i.e., CPA or CoA) is also secure against integrity of plaintexts (PI) forgeries in the same attack.

Part (1) of the proof follows immediately from the definition of EF and PI goals, as shown by Bellare and Namprempre [7].

(2) PI-ATK $\not\Rightarrow$ EF-ATK.
An encryption scheme (mode) that is PI secure in an attack (i.e., CPA or CoA) is not necessarily secure against EF forgeries in the same attack.

Part (2) of the proof is based on a counter-example. Let scheme $\Pi \; o \; g$ be an arbitrary EF-ATK secure scheme. (Note that such schemes exist [16, 17, 13].) We show that any such scheme can be transformed into a scheme that is PI-ATK secure but not EF-ATK secure. Let us define the modified scheme as $\Pi' \; o \; g = (E' \; o \; g, D' \; o \; g, KG)$ that is obtained as follows:

$$\begin{aligned} (E' \; o \; g)(x) &= ((E \; o \; g)(x))||y_0 \\ (D' \; o \; g)(y||y_0) &= (D \; o \; g)(y); \end{aligned}$$

i.e., the encryption is done by appending a random block y_0 to $y = (E \; o \; g)(x)$ (y_0 is unrelated to the plaintext or the rest of the scheme.) The plaintext is obtained by applying the $D \; o \; g$ function to the ciphertext remaining after the removal of the random block y_0.

It is clear that the scheme is not EF secure, because once the adversary obtains a ciphertext $(E \; o \; g)(x)||y_0$, he generates a forgery in which he replaces the random block y_0 by a different block; i.e., $y' = (E \; o \; g)(x)||y_0', y_0' \neq y_0$. This forgery obviously decrypts correctly. Hence, the scheme is not EF secure.

Now, to show that the scheme $(E' \; o \; g, D' \; o \; g, KG)$ is PI secure, we use the fact the class of all possible forgeries can be divided into two complementary classes as follows:
(a) forgeries of type $y||y_0$, where $y = y^i = (E \; o \; g)(x^i)$ (for some index $i, 1 \quad i \quad q_e$) is the $E \; o \; g$ encrypted part of $x^i, 1 \quad i \quad q_e$. These forgeries have the property that $y_0 \neq y_0^i$, hence the forgery is not the ciphertext of a previous query.
(b) forgeries of type $y||y_0$, where $y \neq y^i = (E \; o \; g)(x^i), \forall i, 1 \quad i \quad q_e$.

Any forgery in class (a) decrypts correctly as follows:

$$(D' \ o \ g)(y||y_0) = (D' \ o \ g)(y^i||y_0) = (D \ o \ g)(y^i) = x^i.$$

Hence, for any forgery from class (a):

$$Pr[(D' \ o \ g)(y||y_0) \neq Null \text{ and } (D' \ o \ g)(y) \neq x^i, \forall i, 1 \quad i \quad q_e] = 0.$$

For any forgery from class (b), we will use the fact that the scheme $(E \ o \ g, D \ o \ g, KG)$ is EF secure. Since $y \neq y^i, \forall i, 1 \quad i \quad q_e$, then y is a valid forgery for the EF secure scheme $(E \ o \ g, D \ o \ g, KG)$. Hence,

$$\begin{aligned} &Pr[(D' \ o \ g)(y||y_0) \neq Null \text{ and } (D' \ o \ g)(y||y_0) \neq x^i, \forall i, 1 \quad i \quad q_e] \\ &\quad Pr[(D' \ o \ g)(y||y_0) \neq Null] = Pr[(D \ o \ g)(y) \neq Null] \quad \epsilon, \end{aligned}$$

where ϵ is negligible. Hence, for any forgery (either from class (a) or class (b)),

$$Pr[(D' \ o \ g)(y||y_0) \neq Null \text{ and } (D' \ o \ g)(y||y_0) \neq x^i, \forall i, 1 \quad i \quad q_e] \quad \epsilon,$$

where ϵ is negligible; i.e., the scheme $(E' \ o \ g, D' \ o \ g, KG)$ is PI secure. □

Theorem 2: EF-CPA > PU-CPA
Proof
(1) EF-CPA PU-CPA
An encryption scheme (mode) that is secure against existential forgeries (EFs) in a CPA is also secure against PU forgeries in the same attack.

Part (1) of the proof follows immediately from goal definitions.

$$\begin{aligned} &Pr[(D^{F_K} o \ g)(y) \neq Null \qquad (D^{F_K} o \ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown }] \\ &\quad = 1 \Leftrightarrow Pr[(D^{F_K} o \ g)(y) \neq Null \text{ and } (D^{F_K} o \ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e, \text{ is known }] \\ &\qquad 1 \Leftrightarrow Pr[(D^{F_K} o \ g)(y) \neq Null]. \end{aligned}$$

However, if a scheme is EF secure, then for any forgery y, $Pr[(D^{F_K} o \ g)(y) \neq Null] \quad \delta$, where δ is negligible. Thus,

$$Pr[(D^{F_K} o \ g)(y) \neq Null \qquad (D^{F_K} o \ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown }] \quad 1 \Leftrightarrow \delta \stackrel{def}{=} \lambda;$$

i.e., the scheme is PU-CPA secure.

(2) PU-CPA $\not\Rightarrow$ EF-CPA
An encryption scheme (mode) that is secure against plaintext-uncertain (PU) forgeries in an CPA attack is not necessarily secure against EF forgeries in the same attack.

In Part (2) of the proof, we show that there is a scheme that is PU-CPA secure, but is not EF-CPA secure. Let $(E \ o \ g, D \ o \ g, KG)$ be an EF-CPA secure scheme. We show that the derived scheme $(E' \ o \ g, D' \ o \ g, KG)$, where $(E' \ o \ g)(x) = (E^{F_K} o \ g)(w \oplus x)||r, w = f(r), r \stackrel{R}{\leftarrow} \{0,1\}^l$, $w \oplus x \stackrel{def}{=} w \oplus x_1||w \oplus x_2 \quad ||w \oplus x_n$, and $f = F_K$ is a PRF, is PU-CPA, but it is not EF-CPA.

The derived scheme is clearly not EF-CPA secure. For instance, let the adversary issue an encryption query with plaintext x and obtain the corresponding ciphertext string $y = (E \ o \ g)(w \oplus x)||r$. Then the adversary can construct the forgery $y' \neq y$, where $y' = (E \ o \ g)(w \oplus x)||z$ where $z \neq r$. This forgery passes the integrity check, and $(D^{F_K} o \ g)(y') \neq Null$. To see this, let $w' = f(z) \neq w$. Then $D^{F_K}(y') =$

$w \oplus x||g(w \oplus x)$ and, hence, verifies the integrity condition. Furthermore, the plaintext outcome of forgery y' is $(D' \ o \ g)(y') = x' = w \oplus w' \oplus x$. Hence, the scheme is not EF-CPA secure.

We show that the derived scheme is PU-CPA secure. To see this, let y' be the adversary's forgery after q_e encryption queries with chosen plaintext input. Write $y' = \tilde{y}||z$, for some z. Two complementary cases are identified for the values of the forgery prefix $\tilde{y}$, namely:
(a) there exists $i, 1 \quad i \quad q_e : \tilde{y} = (E \ o \ g)(x^i)$;
(b) $\forall i, 1 \quad i \quad q_e : \tilde{y} \neq (E \ o \ g)(x^i)$.

In case (a), $z \neq r^i$, hence w and w^i are random, uniformly distributed, and independent (here, we assume $f \overset{\mathcal{R}}{\quad} R$). The forgery passes the integrity check since the derived scheme is not EF-CPA secure, and the its plaintext outcome is

$$x = (w \oplus w^i) \oplus x^i.$$

Hence, any block $j, 1 \quad j \quad |x^i| = |x|$, of the plaintext outcome can be written as $x_j = (w \oplus w^i) \oplus x^i_j$. Hence, for any arbitrary constant a

$$Pr[x_j = a] = Pr[(w \oplus w^i) \oplus x^i_j = a] = \frac{1}{2^l}$$

because w, w^i are random, uniformly distributed, and independent, and x^i_j is a known constant (in the CPA attack). For $f \overset{\mathcal{R}}{\quad} F$, where F is a (q, t, ϵ) PRF family, we obtain

$$Pr[x_j = a] = Pr[(w \oplus w^i) \oplus x^i_j = a] = \frac{1}{2^l} + \epsilon.$$

Hence, for any forgery in case (a), $(D^{F_K} o \ g)(y') \neq Null$ and

$$\begin{aligned} &Pr[(D' \ o \ g)(y') \neq Null \text{ and } (D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is known }] \\ &\quad = Pr[(D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is known }] \quad Pr[x_j = a] = \frac{1}{2^l} + \epsilon. \end{aligned}$$

In case (b), the forgery prefix $\tilde{y}$ is itself a forgery for the given secure EF-CPA scheme $(E \ o \ g, D \ o \ g, KG)$, and hence:

$$\begin{aligned} &Pr[(D' \ o \ g)(y') \neq Null \text{ and } (D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is known }] \\ &\qquad Pr[(D' \ o \ g)(y') \neq Null] = Pr[(D^{F_K} o \ g)(\tilde{y}) \neq Null] \quad \delta, \end{aligned}$$

where δ is negligible. Hence, for any forgery,

$$Pr[(D' \ o \ g)(y') \neq Null \text{ and } (D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is known }] \quad \epsilon' \stackrel{def}{=} \max\left(\frac{1}{2^l} + \epsilon, \delta\right),$$

where ϵ' is negligible. Or, equivalently,
$Pr[(D' \ o \ g)(y') \neq Null \qquad (D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown }] \quad 1 \Leftrightarrow \epsilon' \stackrel{def}{=} \lambda$,
where $1 \Leftrightarrow \lambda$ is a negligible quantity. Hence, the derived scheme $(E' \ o \ g, D' \ o \ g, KG)$ is PU-CPA secure.

Theorem 3: EF-CPA $>$ NM-CPA
Proof
(1) EF-ATK   NM-ATK
An encryption scheme (mode) that is secure against existential forgeries (EFs) in an attack ATK (i.e., CPA or CoA) is also secure against NM forgeries in the same attack.

Part (1) of the proof follows immediately from goal definitions.

$$Pr[(D^{F_K} o\ g)(y') \neq Null \text{ and } \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y'))] \quad Pr[(D^{F_K} o\ g)(y) \neq Null].$$

However, if a scheme is EF secure, then for any forgery y $Pr[(D^{F_K} o\ g)(y) \neq Null] \quad \epsilon$, where ϵ is negligible. Thus,

$$Pr[(D^{F_K} o\ g)(y') \neq Null \text{ and } \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y'))] \quad \epsilon,$$

for any forgery $y \neq y^i, 1 \quad i \quad q_2$, which means that the scheme is NM-CPA secure.

(2) NM-CPA $\not\Rightarrow$ EF-CPA
An encryption scheme (mode) that is non-malleable in a CPA attack (i.e., NM-CPA secure) is not necessarily secure in an EF-CPA attack.

In Part (2) of the proof, we show that there is a scheme that is NM-CPA secure, but is not EF-CPA secure. In Section 5, we show that the scheme BIGE$-nzg is NM-CPA secure (Lemma 6) but not EF-CPA secure (Lemma 7).

Theorem 4: PU-CPA > CPF-CPA
Proof
(1) PU-CPA CPF-CPA
An encryption scheme (mode) that is secure against plaintext-uncertain (PU) forgeries in a CPA is also secure against chosen-plaintext forgeries (CPFs) in the same attack.

Part (1) of the proof follows immediately from goal definitions. If a scheme is PU-CPA secure, then for any forgery y

$$Pr[(D^{F_K} o\ g)(y) \neq Null \qquad (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown }] \quad \lambda,$$

where $x^i, 1 \quad i \quad q_e$, are plaintext strings used in encryption and $1 \Leftrightarrow \lambda$ is a negligible quantity. However,

$$((D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is chosen }) \qquad ((D^{F_K} o\ g)(y) = x \text{ is known}).$$

Or, equivalently,
$((D^{F_K} o\ g)(y) = x \text{ is unknown }) \qquad (D^{F_K} o\ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e, \text{ is not chosen})$.
This implies that

$$((D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown })$$
$$((D^{F_K} o\ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e, \text{ is not chosen}).$$

Hence,

$$\lambda \quad Pr[((D^{F_K} o\ g)(y) \neq Null) \qquad ((D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown})$$
$$\text{and}$$
$$((D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is unknown })$$
$$((D^{F_K} o\ g)(y) = x = x^i, \text{ for some i } i, 1 \quad i \quad q_e, \text{ is not chosen})]$$
$$Pr[((D^{F_K} o\ g)(y) \neq Null) \qquad ((D^{F_K} o\ g)(y) = x = x^i, \text{ for some i } i, 1 \quad i \quad q_e, \text{ is not chosen})]$$

or, equivalently,

$$\lambda \quad 1 \Leftrightarrow Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is chosen }]$$

or,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is chosen }] \quad 1 \Leftrightarrow \lambda \stackrel{def}{=} \epsilon,$$

which means that the scheme is CPF-CPA secure.

(2) CPF-CPA $\not\Rightarrow$ PU-CPA
An encryption scheme (mode) that is secure against chosen-plaintext forgeries (CPFs) in a CPA attack is not necessarily secure against PU forgeries in the same attack.

Part (2) of the proof follows immediately from Lemmas 4 and 5, Section 5. That is, the scheme IGE\$-z_0 is CPF-CPA secure (Lemma 5) and is not PU-CPA secure (Lemma 4).

Theorem 5: PI-CPA > CPF-CPA
Proof
(1) PI-ATK CPF-ATK
An encryption scheme (mode) that is secure against plaintext-integrity (PI) forgeries in an attack ATK (i.e., CPA or CoA) is also secure against chosen-plaintext forgeries (CPFs) in the same attack.

Part (1) of the proof follows immediately from goal definitions. For any forgery y,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e, \text{ is chosen }]$$
$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e] \quad \epsilon,$$

since the scheme supposed to be PI-ATK secure.

(2) CPF-CPA $\not\Rightarrow$ PI-CPA
An encryption scheme (mode) that is secure against chosen-plaintext forgeries (CPFs) in a CPA attack is not necessarily secure against PI forgeries in the same attack.

Part (2) of the proof follows immediately from Lemmas 4 and 5, Section 5. That is, the scheme IGE\$-z_0 is CPF-CPA secure (Lemma 5) and is not PI-CPA secure (Lemma 4).

Theorem 6: NM-CPA > CPF-CPA
Proof
(1) NM-CPA CPF-CPA
An encryption scheme (mode) that is non-malleable (NM) in a CPA is also secure against chosen-plaintext forgeries (CPFs) in the same attack.

Part (1) of the proof follows immediately from goal definitions. For any message length m and challenge ciphertexts $y^1, \quad , y^{q_2}$ of unknown plaintext messages $x^1, \quad , x^{q_2} \in \{0,1\}^m$, and for any forgery $y \neq y^i, 1$ $i \quad q_2$ and any relationship $\mathcal{R}$,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y))] \quad \epsilon,$$

where ϵ is a negligible quantity. Hence, by definition,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \qquad \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y)) \text{ does not exist}] \quad 1 \Leftrightarrow \epsilon \stackrel{def}{=} \lambda.$$

However,

$$((D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_1, \text{ is chosen}) \qquad (\mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y)) \text{ exists}),$$

is true, since the plaintext challenge in a successful CPF-CPA attack could always be $x = 111 \quad 1$ (i.e., a block of 1's), which means that $\mathcal{R} \overset{def}{=}$ " ". Equivalently,

$(\mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\, g)(y))$ does not exist) $\quad ((D^{F_K} o\, g)(y) = x = x^i$ for some $i, 1 \quad i \quad q_1$, is not chosen).

Hence,

$$Pr[((D^{F_K} o\ g)(y) \neq Null) \qquad ((D^{F_K} o\ g)(y) = x = x^i \text{ for some } i, 1 \quad i \quad q_1, \text{ is not chosen})]$$
$$Pr[((D^{F_K} o\ g)(y) \neq Null) \qquad (\mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y)) \text{ does not exist})$$
$$\text{and } (\mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y)) \text{ does not exist})$$
$$((D^{F_K} o\ g)(y) = x = x^i \text{ for some } i, 1 \quad i \quad q_1, \text{ is not chosen})]$$
$$\lambda.$$

This means that

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_1, \text{ is chosen}] \quad \epsilon,$$

which means that the scheme is CPF-CPA secure.

(2) CPF-CPA $\not\Rightarrow$ NM-CPA
An encryption scheme (mode) that is secure against chosen-plaintext forgeries (CPFs) in a CPA attack is not necessarily non-malleable in the same attack.

Part (2) of the proof follows immediately from Lemmas 4 and 5, Section 5. That is, the scheme IGE\$-z_0 is CPF-CPA secure (Lemma 5) and is not NM-CPA secure (Lemma 4).

4.2 Incomparability and Separability

Theorem 7: PU-CPA and PI-CPA are Incomparable
Proof
(1) PU-CPA $\not\Rightarrow$ PI-CPA
An encryption scheme (mode) that is PU secure in a CPA attack is not necessarily secure against PI forgeries in the same attack.

For Part (1) of the proof, we choose the same scheme as in the proof of Theorem 2, namely, $(E'\ o\ g, D'\ o\ g, KG)$, where $(E'\ o\ g)(x) = (E^{F_K} o\ g)(w \oplus x)||r, w = f(r), r \overset{\mathcal{R}}{} \{0, 1\}^l$, $w \oplus x \overset{def}{=} w \oplus x_1 || w \oplus x_2 \quad || w \oplus x_n$, and $f = F_K$. We showed in the proof of Theorem 2 that this scheme is PU-CPA secure. Here, we show that this scheme is not PI-CPA secure.

Let us choose the forgery $y' = y||r = (E\ o\ g)(x^i)||r$, where $x^i, 1 \quad i \quad q_e$ is an old plaintext string. The underlying plaintext for this forgery (which decrypts correctly) is

$$x' = (w \oplus w^i) \oplus x^i.$$

Hence, for $f \overset{\mathcal{R}}{} R$ and for any old plaintext string $p, 1 \quad p \quad q_e$ and for any block index $j, 1 \quad j$ $\min(|x'|, |x^p|)$

$$Pr[x'_j = x^p_j] = Pr[(w \oplus w^i) \oplus x^i_j = x^p_j] = \frac{1}{2^l},$$

since w, w^i are random, uniformly distributed, and independent. For $f \ \stackrel{\mathcal{R}}{} \ F$, where F is a (q, t, ϵ) PRF family,

$$Pr[x'_j = x^p_j] = \frac{1}{2^l} + \epsilon.$$

Hence, $(D' \ o \ g)(y') \neq Null$, but

$$Pr[(D' \ o \ g)(y') = x^p, \text{ for some } i, 1 \quad p \quad q_e] \quad Pr[x'_j = x^p_j] \quad \frac{1}{2^l} + \epsilon.$$

Hence,

$$\begin{aligned} &Pr[(D' \ o \ g)(y') \neq Null \text{ and } (D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e] \\ &\quad = Pr[(D' \ o \ g)(y') = x' \neq x^i, \forall i, 1 \quad i \quad q_e] \\ &\quad = 1 \Leftrightarrow Pr[(D' \ o \ g)(y') = x^i, \text{ for some } i, 1 \quad i \quad q_e] \quad 1 \Leftrightarrow \frac{1}{2^l} \Leftrightarrow \epsilon, \end{aligned}$$

and hence, it cannot be negligible.

(2) PI-ATK $\not\Rightarrow$ PU-ATK
An encryption scheme (mode) that is PI secure in an attack (i.e., CPA or CoA) is not necessarily secure against PU forgeries in the same attack.

For Part (2) of the proof, we construct the encryption scheme $(E' \ o \ g, D' \ o \ g, KG)$ from the EF secure encryption scheme $(E \ o \ g, D \ o \ g, KG)$ in the same way as in the proof of Theorem 1. The encryption scheme $(E' \ o \ g, D' \ o \ g, KG)$ is thus PI secure. We show that this scheme is not PU secure. Let us construct a forgery in the same way, namely $y' = (E \ o \ g)(x) || y'_0, y'_0 \neq y_0$, where x is a plaintext used at encryption. This forgery obviously decrypts correctly; i.e., $(D' \ o \ g)(y') = x$ is known, hence,

$$Pr[(D' \ o \ g)(y') \neq Null \text{ and } (D' \ o \ g)(y') = x' \text{ is known }] = 1.$$

Hence, the scheme is not PU secure. □

Theorem 8: NM-CPA is separable from PI-CPA, PU-CPA, and KPF-CPA
Proof

In Section 5, we show that scheme BIGE\$-nzg is NM-CPA secure (Lemma 6), but not PI-CPA and KPF-CPA secure (Lemma 7). Hence, NM-CPA $\not\Rightarrow$ PI-CPA and NM-CPA $\not\Rightarrow$ KPF-CPA.

When implemented with the CBC mode and used to encrypt messages consisting of an integer number of l-bit blocks (possibly after padding), the Variable Input Length (VIL) cipher of Bellare and Rogaway [5, 6] can be shown generate at least a random block in the plaintext outcome of any forgery produced in a CPA [11]. Hence, the composition of this scheme with the MDC function nzg(x) defined for the BIGE\$-nzg scheme (viz., Section 5), namely VIL-CBC-nzg, is a PU-CPA secure scheme. However, this scheme is not NM-CPA secure for the same reasons the scheme IGE\$-z_0 is not NM-CPA secure (viz., end of the Proof of Lemma 4). Hence, PU-CPA $\not\Rightarrow$ NM-CPA.

4.3 Extensions of the CPA Lattice

Theorems 1 - 8 show that the integrity goals defined in Section 3 form a lattice for chosen-plaintext attacks. In this section we show that, if we also consider ciphertext-only attacks, the top of the lattice remains EF-CPA, but CPF-CoA becomes the new bottom of the lattice.

Theorem 9: EF-CPA > EF-CoA
Proof
(1) EF-CPA   EF-CoA An encryption scheme (mode) that is EF-CPA secure is also secure against EF-CoA attacks.

Part (1) of the proof follows directly from the following observation.

Observation:
An encryption scheme (mode) that is secure with respect to a given goal (i.e., EF, PI, PU, PA, NM, CPA) in an CPA attack is also secure with respect to the same goal in a CoA attack.

This is true because an adversary that breaks integrity with respect to a goal in a CoA attack will break security in a CPA attack, since the adversary can obviously ignore the plaintext and use only the ciphertext obtained.

(2) EF-CoA $\not\Rightarrow$ EF-CPA An encryption scheme (mode) that is EF-CoA secure is not necessarily secure against EF-CPA attacks.

Part (2) of the proof follows directly from Lemmas 2 and 4, Section 5. That is scheme IGE\$-z_0 is EF-CoA secure (Lemma 2) and is not EF-CPA secure (Lemma 4).

Theorem 10: CPF-CPA > CPF-CoA
Proof
(1) CPF-CPA   CPF-CoA An encryption scheme (mode) that is CPF-CPA secure is also secure against CPF-CoA attacks.

Part (1) of the proof follows directly from the the observation of the Proof in Theorem 9.

(2) CPF-CoA $\not\Rightarrow$ CPF-CPA An encryption scheme (mode) that is CPF-CoA secure is not necessarily secure against CPF-CPA attacks.

Part (2) of the proof is based on a counter-example. Let scheme Π o g be consist of $\Pi \equiv$ XOR\$ [2], and $g(x) \equiv$ {per-block, bitwise exclusive-or}. It is easy to see that this scheme is CPF-CoA secure since any modification of the ciphertext that causes the bitwise exclusive-or check to pass remains unknown to (and therefore cannot be a priori predicted by) the adversary. In contrast, if the adversary can encrypt plaintext of his choice, he can (1) encrypt a plaintext message that differs from the challenge plaintext by a single bit, and (2) simply flip the appropriate bit of the ciphertext obtained.

4.4 Other Relationships

Theorem 11: PI-CPA > KPF-CPA
Proof
(1) PI-CPA   KPF-CPA
An encryption scheme (mode) that is PI-CPA secure is also KPF-CPA secure.

Part (1) of the proof follows immediately from goal definitions. If a scheme that is PI secure, then for any forgery y

$$Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } (D^{F_K} o\ g)(y) = x \neq x^i, \forall i, 1 \quad i \quad q_e] \quad \epsilon,$$

where $x^i, 1 \quad i \quad q_e$ are the plaintext strings used for the encryption queries and ϵ is a negligible quantity.

Equivalently,

$$1 \Leftrightarrow Pr[(D^{F_K} o\ g)(y) \neq Null \quad (D^{F_K} o\ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e] \quad \epsilon,$$

or,

$$Pr[(D^{F_K} o\ g)(y) \neq Null \quad (D^{F_K} o\ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e] \quad 1 \Leftrightarrow \epsilon \stackrel{def}{=} \lambda.$$

However, $((D^{F_K} o\ g)(y) = x = x^i, \text{ for some } i, 1 \quad i \quad q_e) \quad ((D^{F_K} o\ g)(y) = x \text{ is known })$. Hence,

$$\begin{aligned} \lambda \quad & Pr[((D^{F_K} o\ g)(y) \neq Null \quad (D^{F_K} o\ g)(y) = x = x^i \text{ for some } i, 1 \quad i \quad q_e) \\ & \text{and } ((D^{F_K} o\ g)(y) = x = x^i \text{ for some } i, 1 \quad i \quad q_e) \quad (D^{F_K} o\ g)(y) = x \text{ is known })] \\ & Pr[(D^{F_K} o\ g)(y) \neq Null \quad (D^{F_K} o\ g)(y) = x \text{ is known }], \end{aligned}$$

which means that the scheme is KPF-CPA secure. □

(2) KPF-CPA $\not\Rightarrow$ PI-CPA
An encryption scheme (mode) that is KPF-CPA secure is not necessarily secure against PI-CPA attacks.

Part (2) of the proof follows immediately from the counter-example provided by Lemmas 3 and 4, Section 5. That is, scheme IGE$-c is KPF-CPA secure (Lemma 3) but it is not PI-CPA secure (Lemma 4).

Theorem 12: KPF-CPA is incomparable with CPF-CPA and with PU-CPA
Proof
(1) KPF-CPA $\not\Rightarrow$ CPF-CPA
An encryption scheme (mode) that is KPF-CPA secure is not necessarily CPF-CPA secure.

Part (1) of the proof follows immediately from the fact that scheme IGE$-c is KPF-CPA secure (Lemma 3) but is not CPF-CPA secure in the face of a truncation attack since function $g = c$ placed in the last block of a plaintext is a known constant.

(2) CPF-CPA $\not\Rightarrow$ KPF-CPA
An encryption scheme (mode) that is CPF-CPA secure is not necessarily KPF-CPA secure.

Part (2) of the proof follows immediately from the observation that scheme BIGE$-nzg is CPF-CPA secure, as a consequence of Theorem 6, and is not KPF-CPA secure, by Lemma 7, Section 5.

Note that the scheme BIGE$-nzg also shows that CPF-CoA $\not\Rightarrow$ KPF-CPA.

(3) KPF-CPA $\not\Rightarrow$ PU-CPA
An encryption scheme (mode) that is KPF-CPA secure is not necessarily PU-CPA secure.

Part (3) follows immediately from the same example as in Part (1).

(4) PU-CPA $\not\Rightarrow$ KPF-CPA
An encryption scheme (mode) that is PU-CPA secure is not necessarily KPF-CPA secure.

Part (4) follows immediately from the observation that the VIL-CBC-nzg mode is PU-CPA secure but not KPF-CPA secure, since it generates at least a random block in the plaintext outcome of a forgery in a CPA [11].

5 Examples of Integrity Characteristics of Practical Encryption Schemes

5.1 The Infinite Garble Extension Mode

Most of the proofs of theorems presented in the previous section are based on examples provided by Lemmas 1 – 7 of this section. These lemmas refer to schemes derived from an encryption mode that was proposed by Carl Campbell at the first National Bureau of Standards *Conference on Computer Security and the Data Encryption Standard*, in February 1977 [9]. Campbell called his mode the "Infinite Garble Extension" mode and, for this reason, we denote it by IGE below. Although Campbell's mode appears to have been proposed about the same time as the CBC mode, its integrity properties have not been explained in published literature to date.

IGE uses the family F of super-pseudorandom permutation functions (SPRPs), which is defined as follows. ([3], [18]). Let $F : \{0,1\}^k \times \{0,1\}^l \to \{0,1\}^l$ be a pseudorandom permutation family and $f = F_K$ be a permutation randomly chosen by key K (i.e., $K \stackrel{\mathcal{R}}{} \; 0,1\}^k$) and $f^{-1} = F_K^{-1}$ its inverse. Let P^l denote all { the permutations on $\{0,1\}^l$, and A be a two-oracle adversary. F is a SPRP if the advantage of function family F, $Adv_F^{sprp}(t, q, \mu)$, is

$$Adv_F^{sprp}(t,q,\mu) = \max_A\{Adv_F^{sprp}(A)\} \leq \epsilon,$$

where the maximum is taken over all the adversaries A issuing q enciphering or deciphering queries totaling $\mu = ql$ bits and taking time t, ϵ is a negligible quantity, and the advantage of an adversary A is

$$Adv_F^{sprp}(A) = |Pr[A = 1 : f, f^{-1} \stackrel{\mathcal{R}}{} \; F] \Leftrightarrow Pr[A = 1 : f, f^{-1} \stackrel{\mathcal{R}}{} \; P^l]|.$$

IGE is based on the following block chaining sequence:

$$y_i = f(x_i \oplus y_{i-1}) \oplus x_{i-1}$$

for encryption, and

$$x_i = f^{-1}(y_i \oplus x_{i-1}) \oplus y_{i-1}$$

for decryption, where $f \stackrel{\mathcal{R}}{} \; F$, or $f = F_K$. Note that chaining is symmetric in encryption/decryption, and consequently this mode propagates errors until the end of a message, thereby extending the error propagation characteristics of CBC. The initialization phase could be defined as: $r_0 \quad \{0,1\}^l, y_0 = f'(r_0), x_0 = r_0$, where $f' = F_{K'}$, K and K' being two distinct keys. (Other initialization definitions can be used.) Hence, the encryption and decryption functions for the *stateless* mode (denoted by IGE$ below) are defined by $\mathcal{E} \Leftrightarrow IGE\$^{F_K}(x)$ and $\mathcal{D} \Leftrightarrow IGE\$^{F_K}(y)$, as follows:

function $\mathcal{E} \Leftrightarrow$IGE$$^f(x)$
$r_0 \quad 0,1\}^l$
$y_0 = f'(r_0);\ x_0 = r_0$
for $i = 1, \quad , n$ **do** {
$y_i = f(x_i \oplus y_{i-1}) \oplus x_{i-1}$ }
return $y = y_0 || y_1 y_2 \quad y_n$

function $\mathcal{D} \Leftrightarrow$IGE$$^f(y)$
Parse y as $y_0 || y_1 \quad y_n$ {
$r_0 = f'^{-1}(y_0);\ x_0 = r_0$
for $i = 1, \quad , n$ **do** {
$x_i = f^{-1}(y_i \oplus x_{i-1}) \oplus y_{i-1}$ }
return $x = x_1 x_2 \quad x_n$

A stateful IGE mode can be defined in a similar manner to that used for the XCBC stateful mode.

[Note that IGE$ is based on the CBC mode in the sense that its output block i is exclusive-ored plaintext block $i \Leftrightarrow 1$. Hence, the IGE$ scheme is secure in the real-or-random (or left or right) sense against adaptive chosen plaintext attacks and the proof is very similar to that of Bellare *et al.* [2].]

Let us introduce the scheme Π o $g \equiv$ IGE\$-z_0 $= (\mathcal{E}\Leftrightarrow IGE\$ \ o \ z_0, \mathcal{D}\Leftrightarrow IGE\$ \ o \ z_0, KG)$ by using function $g(x) = z_0 = f'(r_0{+}1)$ to define $y = \mathcal{E} \ o \ z_0 = \mathcal{E}^{F_K}(x||z_0))$. Hence, the scheme IGE\$-z_0 encrypts any plaintext $x = x_1 \quad x_n$ by appending block $x_{n+1} = z_0$ to plaintext x, and then encrypting string $x_1 \quad x_n x_{n+1}$.

Let us introduce the scheme Π' o $g \equiv$ IGE\$-c $= (\mathcal{E}\Leftrightarrow IGE\$ \ o \ c, \mathcal{D}\Leftrightarrow IGE\$ \ o \ c, KG)$ by using function $g(x) = c$, where c is a known constant, to define $y = \mathcal{E} \ o \ c = \mathcal{E}^{F_K}(x||c))$. Hence, the scheme IGE\$-c encrypts any plaintext $x = x_1 \quad x_n$ by appending block $x_{n+1} = c$ to plaintext x, and then encrypting string $x_1 \quad x_n x_{n+1}$.

The integrity properties of schemes IGE\$-z_0 and IGE\$-c are formalized in Lemmas 1   5 (whose proofs can be found in the appendix).

To state Lemma 1 [Main IGE Lemma], we need to introduce two sets, namely

$$S^e = \{y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p\},$$

which consists of all inputs to f^{-1} that can be made up by taking the exclusive-or of every plaintext block of the q_e strings $x^p = x_1^p \quad x_n^p$ with every block of the q_e ciphertext strings $y^p = y_0^p y_1^p \quad y_n^p$ obtained at encryption; and set

$$S_j^d = \{y_s \oplus x_{s-1}, 1 \quad s \quad j\},$$

which consists of all the combinations $y_s \oplus x_{s-1}$ of forgery y plaintext and ciphertext blocks used at the decryption of y up to (but not including) position j.

For any $f \ ^{\mathcal{R}} \ P^l$ and S^e, we define the finite family of random functions $G_S : \{0,1\}^k \quad 0,1\}^l \to \{0,1\}^l$ whose members are $f, \overline{f}$, with $\overline{f}$ defined as:

$$\overline{f} = \begin{cases} f^{-1}(t), & t \in S^e \\ v(t), & t \in \{0,1\}^l \Leftrightarrow S^e, v \ ^{\mathcal{R}} \ R^{l,l} \end{cases},$$

where $R^{l,l}$ is the set of all functions from $\{0,1\}^l$ to $\{0,1\}^l$. We denote by $f \ ^{\mathcal{R}} \ G_S$ the random selection of f and $\overline{f}$ from G_S.

The family of functions G_S behaves exactly like P^l when the plaintext blocks input to f and ciphertext blocks input to f^{-1} are those generated during the encryption of any adversary's q_e chosen-plaintext queries, and behaves exactly like $R^{l,l}$ during the decryption of any ciphertext block *not* in S^e.

Note that the family G_S is well-defined for any message-integrity attack because, by definition (viz., Section 3.2), in any such attack, all q_e encryption queries precede the forgery verification queries. (Also note that we allow $q_e = 0$ and, in this case, $S^e = \emptyset$ and $\overline{f} = v$.)

For Lemmas 1-7 we define *Succ* the event that the forgery is successful for the chosen goal-attack combination. Then in the proofs of these Lemmas, we use the result of Fact 0 below (whose proof can be found elsewhere [13]) that provides the reduction from $f \ ^{\mathcal{R}} \ F$ to $f \ ^{\mathcal{R}} \ G_S$.

Fact 0

(a)

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[\text{Succ}] \quad \epsilon + Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\text{Succ}].$$

(b)

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\text{Succ}] \quad Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[\text{Succ}] + \frac{\mu_v(\mu_v \Leftrightarrow l)}{l^2 2^{l+1}}.$$

where $\frac{\mu_v}{l}$ is the total number of ciphertext blocks used in all verified forgeries. Unless we state otherwise, assume that $f \;\; ^{\mathcal{R}} \;\; G_S$ (and drop this subscript from $Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}$[Succ].)

Let i denote the position of the first ciphertext block in the forgery $y = y_0 y_1 \quad y_n$ such that $y_i \oplus x_{i-1}$ does not collide with any of the $y_k^p \oplus x_{k-1}^p$ values generated during the encryption of the q_e queries. Formally, i is the index of the first block such that $y_i \oplus x_{i-1} \notin S^e$ and $S_i^d \subseteq S^e$.

Lemma 1 [Main IGE Lemma]

Let $y = y_0 y_1 \quad y_n$ be a forged ciphertext and $x = x_0 x_1 \quad x_n$ be its decryption by the function $\mathcal{D} \Leftrightarrow$IGE$\$^f(y)$.
Let a be an arbitrary constant value.
(a) If $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$, then

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[x_n = a] \qquad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x_n = a] \qquad \epsilon' \overset{def}{=} \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}},$$

where q_e is the maximum number of encryption queries, totaling at most μ_e bits.
(b) If $i, 1 \quad i \quad n$, is the first block for which $y_i \oplus x_{i-1} \notin S^e$, then

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[x_n = a] \qquad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x_n = a] \qquad \epsilon' \overset{def}{=} \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}},$$

where the total number of bits for the q_e encryption queries is at most μ_e.

One can also show that the conclusions of Main IGE$ Lemma remain valid if the constant a is replaced with the random, uniformly distributed, and independent $z_0 = f'(r_0 + 1)$. This is formalized in the following corollary.

Corollary
Let $y = y_0 y_1 \quad y_n$ be a forged ciphertext and $x = x_0 x_1 \quad x_n$ be its decryption by the function $\mathcal{D} \Leftrightarrow$IGE$\$^f(y)$.
(a) If $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$, then

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[x_n = z_0] \qquad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x_n = z_0] \qquad \epsilon' \overset{def}{=} \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}},$$

where where the total number of bits for the q_e encryption queries is at most μ_e bits.
(b) If $i, 1 \quad i \quad n$, is the first block for which $y_i \oplus x_{i-1} \notin S^e$, then

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[x_n = z_0] \qquad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x_n = z_0] \qquad \epsilon' \overset{def}{=} \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}},$$

where q_e is the maximum number of encryption queries, totaling at most μ_e bits.

Lemma 2. The scheme IGE\$-z_0 is EF-CoA secure.

Lemma 3. The scheme IGE\$-c is KPF-CPA secure.

Lemma 4. The schemes IGE\$-z_0 and IGE\$-c are not EF-CPA, PU-CPA, PI-CPA, and NM-CPA secure.

Lemma 5. The scheme IGE\$-z_0 is CPF-CPA secure.

5.2 The Bidirectional Infinite Garble Extension Mode

In this section, we define a variant of the IGE modes that is intended to illustrate, among other things, the separation between NM-CPA and several other integrity notions such as EF, PI, and PA in chosen-plaintext attacks.

The bidirectional IGE (BIGE) scheme consists of the application of the IGE scheme to the input plaintext to obtain an intermediate "hidden" ciphertext, followed by the application of the IGE chaining to the hidden ciphertext in opposite direction to obtain the ciphertext that is output to the user. This general description allows for several actual variants of the bidirectional IGE scheme, namely the stateless or stateful schemes, or schemes that use different keys per pass in each direction. In our example here, the scheme that uses three keys, one per pass in one direction, and one for we have the initialization phase. That is, during initialization we set: $r_0 \quad \{0,1\}^l, y_0 = f'(r_0), x_0 = r_0$, where $f' = F_{K'}$, K and K' are the two distinct keys, and F is the SPRP family. Then, the first pass generates the hidden ciphertext as $z_i = f(x_i \oplus z_{i-1}) \oplus x_{i-1}, 1 \quad i \quad n = |x|$. The second pass consists of $y_0 = f'(z_n)$, where $f' = F_{K'}$, and $y_i = f''(z_{n-i} \oplus y_{i-1}) \oplus z_{n-i+1}, 1 \quad i \quad n$, where $f'' = F_{K''}$, K, K' and K'' are distinct keys.

In the BIGE\$ scheme defined below, the actual encryption and decryption functions for the stateless bidirectional IGE scheme that uses two keys, one for each pass, are defined by $\mathcal{E}\Leftrightarrow BIGE\$^{F_K,F'_K,F''_K}(x)$ and $\mathcal{D}\Leftrightarrow BIGE\$^{F_K,F'_K,F''_K}(y)$, as follows:

function $\mathcal{E}\Leftrightarrow$BIGE\$$^{f,f',f''}(x)$
$r_0 \quad 0,1\}^l$
$z_0 = f'(r_0);\ x_0 = r_0$
for $i = 1, \quad , n$ **do** $\{$
$z_i = f(x_i \oplus z_{i-1}) \oplus x_{i-1}$ $\}$
$y_0 = f'(z_n)$
for $i = 1, \quad , n$ **do** $\{$
$y_i = f''(z_{n-i} \oplus y_{i-1}) \oplus z_{n-i+1}$
$\}$
return $y = y_0||y_1y_2 \quad y_n$

function $\mathcal{D}\Leftrightarrow$BIGE\$$^{f,f',f''}(y)$
Parse y as $y_0||y_1 \quad y_n \{$
$z_n = f'^{-1}(y_0);$
for $i = 1, \quad , n$ **do** $\{$
$z_{n-i} = f''^{-1}(y_i \oplus z_{n-i+1}) \oplus y_{i-1}$
$\}$
$r_0 = f'^{-1}(z_0);\ x_0 = r_0$
for $i = 1, \quad , n$ **do** $\{$
$x_i = f^{-1}(z_i \oplus x_{i-1}) \oplus z_{i-1}$ $\}$
return $x = x_1x_2 \quad x_n$

Let us introduce the scheme $\Pi\ o\ g \equiv$ BIGE\$-nzg $= (\mathcal{E}\Leftrightarrow BIGE\$\ o\ nzg, \mathcal{D}\Leftrightarrow BIGE\$\ o\ nzg, KG)$ by using function $g(x) = nzg(x) = r \overset{\mathcal{R}}{\quad} 0,1\}^l, r \neq 0$ to define $y = \mathcal{E}\ o\ g = \mathcal{E}^{F_K,F'_K,F''_K}(x||g))$. Hence, the scheme BIGE\$-nzg encrypts any plaintext $x = x_1 \quad x_n$ by appending block $x_{n+1} = r$ to plaintext x, and then encrypting string $x_1 \quad x_n x_{n+1}$. The integrity check performed upon the decryption of a forgery y' is simply $x'_{n+1} \neq 0$.

The intuition behind the BIGE$ scheme is as follows. Any modification of ciphertext would cause a modification of the hidden ciphertext, which acts as the input to the second pass of encryption. The resulting hidden ciphertext's modification is unpredictable and propagates from the block position where it occurs until the block z_0 of the hidden ciphertext. The propagation cannot be stopped by the adversary with more than negligible probability, since the adversary does know the values of the hidden ciphertext input to the second pass of encryption with non-negligible probability. (To stop the propagation of any modification to ciphertext output to the user, the adversary would have to know both the input and the output to second encryption pass, as illustrated by the proof of Lemma 4.) Furthermore, any unpredictable modification of the hidden ciphertext starting with block z_0, ends up propagating throughout the message plaintext during the second decryption pass. Hence, the entire plaintext output of BIGE$ will contain blocks whose content is unpredictable.

The integrity properties of the scheme BIGE$-$nzg$ are formalized in the following lemmas (whose the proofs can be found in the appendix).

Lemma 6. The scheme BIGE$-nzg is NM-CPA secure.

Lemma 7. The scheme BIGE$-nzg is not EF-CPA, PI-CPA and KPF-CPA secure.

5.3 Other Examples

Example 1. Let Π be one of the modes {CBC, PCBC}, and function $g(x)$ be the per-block, bitwise exclusive-or function, which we denote by XOR. The schemes Π *o* XOR are CPF-CoA secure, but not CPF-CPA secure [19], or secure with respect to any other goals.

Example 2. Let Π be one of the modes {CBC, PCBC}, and function $g(x)$ be the "confounded CRC-32" function used by Kerberos V [22] and DCE [21]. The schemes Π *o* XOR are CPF-CPA secure, and not secure with respect to any of the other goals CPAs [23].

Example 3. The scheme Π *o* $g \equiv$ BIGE$-c $= (\mathcal{E}\Leftrightarrow BIGE\$ \ o \ c, \mathcal{D}\Leftrightarrow BIGE\$ \ o \ c, KG)$ by using function $g(x) = c$ where c is a constant is EF-CPA secure. (The proof is very similar to the proof of Lemma 6.)

Acknowledgements

We than thank Michaela Iorga for her help with the proof of Lemma 4, and to Jonathan Katz for his comments and suggestions on earlier drafts.

References

[1] M. Bellare and P. Rogaway, "Optimal Asymmetric Encryption How to encrypt with RSA," *Advances in Cryptology - Eurocrypt '94* (LNCS 950) (A. De Santis ed.), Springer-Verlag, 1994.

[2] M. Bellare, A. Desai, E. Jokipii, and P. Rogaway, "A Concrete Security Treatment of Symmetric Encryption," Proceedings of the 38th Symposium on Foundations of Computer Science, IEEE, 1997, (394-403). A full version of this paper is available at *http://www-cse.ucsd.edu/users/mihir.*

[3] M. Bellare, J. Killian, and P. Rogaway, "The security of cipher block chaining", *Advances in Cryptology-CRYPTO '94* (LNCS 839), 341-358, 1995.

[4] M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway, "Relations Among Notions of Security for Public-key Encryption Schemes," *Advances in Cryptology - CRYPTO '98* (LNCS 1462), 26-45, 1998.

[5] M. Bellare and P. Rogaway, "Block Cipher Mode of Operation for Secure, Length-Preserving Encryption," *U.S Patent No. 5,673,319,* September, 1997.

[6] M. Bellare and P. Rogaway, "On the construction of variable-input-length ciphers," Proceedings of the 6th Workshop on Fast Software Encryption, L. Knudsen (Ed), Springer-Verlag, 1999.

[7] M. Bellare and C. Namprempre, "Authenticated Encryption: Relations among notions and analysis of the generic composition paradigm," manuscript, May 2000, *http://eprint.iacr.org/2000.025.ps.*

[8] S.M. Bellovin, "Cryptography and the Internet," *Advances in Cryptology - CRYPTO '98* (LNCS 1462), 46-55, 1998.

[9] C.M. Campbell, "Design and Specification of Cryptographic Capabilities," in *Computer Security and the Data Encryption Standard*, (D.K. Brandstad (ed.)) National Bureau of Standards Special Publications 500-27, U.S. Department of Commerce, February 1978, pp. 54-66.

[10] D. Dolev, C. Dwork, and M.Naor, "Non-malleable Cryptography," Proc. of the 23rd ACM Symp. on Theory of Computing, pp. 542-552, 1991.

[11] V.D. Gligor and P. Donescu, "Integrity Conditions for Symmetric Encryption," University of Maryland, Computer Science Technical Report, CS-TR-3958, December 1998.

[12] V.D. Gligor, S.G. Stubblebine, and P. Donescu, "New Integrity-Aware CBC Encryption Schemes," University of Maryland, Computer Science Technical Report, CS-TR-3999, March 1999.

[13] V.D. Gligor and P. Donescu, "Fast Encryption and Authentication: XCBC Encryption and XECB Authentication Modes," manuscript August, 2000. *http://www.glue.umd.edu/~gligor.*

[14] F.H. Hinsley and A. Stripp, *Codebreakers: the inside story of Bletchley Park*, Oxford University Press, 1993.

[15] J. Katz and M. Yung, "Complete characterization of security notions for probabilistic private-key encryption," Proc. of the 32nd Annual Symp. on the Theory of Computing, ACM 2000.

[16] J. Katz and M. Yung, "Unforgeable Encryption and Adaptively Secure Modes of Operation," Proc. Fast Software Encryption 2000, B. Schneier (ed.) (to appear in Springer-Verlag, LNCS).

[17] C.S. Jutla, "Encryption Modes with Almost Free Message Integrity," IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, manuscript, August, 2000. *http://eprint.iacr.org/2000/039.*

[18] M Luby and C. Rackoff, "How to construct pseudorandom permutations from pseudorandom functions", *SIAM J. Computing*, Vol. 17, No. 2, April 1988.

[19] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, 1997.

[20] M. Naor and O. Reingold, "From Unpredictability to Indistinguishability: A Simple Construction of Pseudo-Random Functions from MACs," *Advances in Cryptology - CRYPTO '98* (LNCS 1462), 267-282, 1998.

[21] Open Software Foundation, "OSF - Distributed Computing Environment (DCE), Remote Procedure Call Mechanisms," Code Snapshot 3, Release, 1.0, March 17, 1991.

[22] RFC 1510, "The Kerberos network authentication service (V5)", Internet Request for Comments 1510, J. Kohl and B.C. Neuman, September 1993.

[23] S. G. Stubblebine and V. D. Gligor, "On message integrity in cryptographic protocols", Proceedings of the 1992 IEEE Symposium on Research in Security and Privacy, 85-104, 1992.

[24] S. G. Stubblebine and C. Meadows, "On Searching for Known and Chosen Pairs Using the NRL Protocol Analyzer," DIMACS Workshop on Design and Formal Verification of Security Protocols, September 1997 (also in *IEEE Journal on Selected Areas in Communications, 1999*).

A Proofs

Proof of Lemma 1 [Main IGE Lemma]

By using Fact 0, we reduce the proof from $f \stackrel{\mathcal{R}}{\leftarrow} F$ to $f \stackrel{\mathcal{R}}{\leftarrow} G_S$. In the proof of this lemma we use the notation $Pr[.] = Pr_{f \stackrel{\mathcal{R}}{\leftarrow} G_S}[.]$. We first present the part of the proof that is common for both parts (a) and (b) of Lemma 1, and then we complete the proof for parts (a) and (b) separately.

Block $x_n = \overline{f}(y_n \oplus x_{n-1}) \oplus y_{n-1}$ of the decrypted forgery y is random, uniformly distributed, and independent of anything else, including value a, whenever $\overline{f}(y_n \oplus x_{n-1})$ is random, uniformly distributed, and independent of anything else. For this to happen, $y_n \oplus x_{n-1}$ must not collide with any element of either S^e or S^d_n (since, in this case, $\overline{f}(y_n \oplus x_{n-1}) = v(y_n \oplus x_{n-1}), v \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}$ and $y_n \oplus x_{n-1}$ has never been encountered before). Let event $\overline{C_n}$ be defined as:

$$C_n : y_n \oplus x_{n-1} \in S^e \cup S^d_n.$$

In this case, i.e., when there are no collisions, we have

$$Pr[x_n = a|\overline{C_n}] \leq \frac{1}{2^l}.$$

By standard conditioning,

$$Pr[x_n = a] \leq Pr[C_n] + Pr[x_n = a|\overline{C_n}] \leq Pr[C_n] + \frac{1}{2^l}.$$

To determine $Pr[C_n]$, we use standard conditioning again, and obtain

$$Pr[C_n] \leq \cdots \leq Pr[C_{n-1}] + Pr[C_n|\overline{C_{n-1}}] \leq Pr[C_1] + \sum_{j=1}^{n-1} Pr[C_{j+1}|\overline{C_j}],$$

where event C_j is defined in a similar manner to that of C_n, namely

$$C_j : y_j \oplus x_{j-1} \in S^e \cup S^d_j.$$

We now determine an upper bound for $Pr[C_{j+1}|\overline{C_j}]$. Event $\overline{C_j}$ is true for $1 \leq j \leq n \Leftrightarrow 1$ means that $y_j \oplus x_{j-1}$ does not collide with any element of either S^e or S^d_n. In this case, $x_j = \overline{f}(y_j \oplus x_{j-1}) \oplus y_{j-1} = v(y_j \oplus x_{j-1}) \oplus y_{j-1}$ is random, uniformly distributed and independent of anything else, since y_{j-1} is a constant, $v \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}$ and $y_i \oplus x_{i-1}$ has never been encountered before. Hence, since y_{j+1} is a chosen constant, $y_{j+1} \oplus x_j$ is also random, uniformly distributed, and independent of anything else. This means that

$$\begin{aligned}
Pr[y_{j+1} \oplus x_j = y^p_k \oplus x^p_{k-1}|\overline{C_j}] &\leq \frac{1}{2^l}, \quad \forall p, k, 1 \leq p \leq q_e, 1 \leq k \leq n_p, \\
Pr[y_{j+1} \oplus x_j = y_s \oplus x_{s-1}|\overline{C_j}] &\leq \frac{1}{2^l}, \quad \forall s, 1 \leq s \leq j.
\end{aligned}$$

But, by standard conditioning and union bound,

$$\begin{aligned}
Pr[C_{j+1}|\overline{C_j}] &= Pr[y_{j+1} \oplus x_j \in S^e \cup S^d_{j+1}|\overline{C_j}] \\
&\leq Pr[y_{j+1} \oplus x_j \in S^e|\overline{C_j}] + Pr[y_{j+1} \oplus x_j \in S^d_{j+1}|\overline{C_j}] \\
&\leq \sum_{p=1}^{q_e} \sum_{k=1}^{n_p} Pr[y_{j+1} \oplus x_j = y^p_k \oplus x^p_{k-1}|\overline{C_j}] \\
&+ \sum_{s=1}^{j} Pr[y_{j+1} \oplus x_j = y_s \oplus x_{s-1}|\overline{C_j}].
\end{aligned}$$

Thus,

$$Pr[C_{j+1}|\overline{C_j}] \quad \frac{\frac{e}{l}+j}{2^l}$$

because there are at most $\frac{e}{l}$ elements in S^e (the $\frac{e}{l}$ ciphertext blocks include $y_0^p, 1 \quad p \quad q_e$ and $y_k^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p$), and j elements in S_{j+1}^d ($y^1 \oplus x_0, \quad , y^j \oplus x^{j-1}$).

Now we consider event C_1 of part (a) of the Lemma separately from event C_i of part (b) of the Lemma.

(a) Since $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$, it follows that $r_0 = \overline{f'}(y_0) = v'(y_0)$ is random, uniformly distributed, and independent of anything else. Here v' is the corresponding function for f' constructed in the same way as v, namely, $v' \stackrel{\mathcal{R}}{} R^{l,l}$. Hence $x_0 = r_0$ is random, uniformly distributed, and independent of anything else. Hence $y_1 \oplus x_0 \in S^e$ happens with probability at most $\frac{|S^e|}{2^l} = \frac{\frac{\mu_e}{l}}{2^l}$. From here on, we apply the same idea (viz., also part (b) below), namely:

$$\begin{aligned} Pr[x_n = a] \quad & Pr[x_n = a | y_1 \oplus x_0 \notin S^e] + Pr[y_1 \oplus x_0 \in S^e] \quad \frac{\mu_e}{l2^l} + \frac{(n \Leftrightarrow i)\mu_e}{l2^l} + \frac{n^2 \Leftrightarrow i^2}{2^{l+1}} \\ & \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}. \end{aligned}$$

Hence, by Fact 0 with $\frac{v}{l} = n$,

$$\begin{aligned} Pr_{f \stackrel{\mathcal{R}}{\leftarrow} P^l}[x_n = a] \quad & Pr_{f \stackrel{\mathcal{R}}{\leftarrow} P^l}[x_n = a] + \frac{n(n \Leftrightarrow 1)}{2^{l+1}} \quad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}} + \frac{n(n \Leftrightarrow 1)}{2^{l+1}} \\ = \; & \epsilon' \stackrel{def}{=} \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}}. \end{aligned}$$

(b) However, by the Lemma hypothesis, event C_j is true for $j< \; i$ and event C_i is false. Hence,

$$\begin{aligned} Pr[C_n] \quad & Pr[C_i] + \sum_{j=i}^{n-1} Pr[C_{j+1}|\overline{C_j}] \\ = \; & \sum_{j=i}^{n-1} Pr[C_{j+1}|\overline{C_j}]. \end{aligned}$$

Using the formula for $Pr[C_{j+1}|\overline{C_j}]$ we obtain

$$\begin{aligned} Pr[C_n] \quad & \sum_{j=i}^{n-1} Pr[C_{j+1}|\overline{C_j}] \quad \sum_{j=i}^{n-1} \frac{\frac{e}{l}+j}{2^l} = \frac{(n \Leftrightarrow i)\mu_e}{l2^l} + \frac{(n \Leftrightarrow i)(i + n \Leftrightarrow 1)}{2^{l+1}} \\ & \frac{(n \Leftrightarrow i)\mu_e}{l2^l} + \frac{n^2 \Leftrightarrow i^2}{2^{l+1}}. \end{aligned}$$

Finally,

$$Pr[x_n = a] \quad Pr[C_n] + \frac{1}{2^l} \quad \frac{(n \Leftrightarrow i)\mu_e}{l2^l} + \frac{n^2 \Leftrightarrow i^2}{2^{l+1}} + \frac{1}{2^l} \quad \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}}.$$

Hence, by Fact 0 with $\frac{v}{l} = n$,

$$Pr_{f \stackrel{\mathcal{R}}{\leftarrow} P^l}[x_n = a] \quad \epsilon' \stackrel{def}{=} \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}} + \frac{n(n \Leftrightarrow 1)}{2^{l+1}} = \frac{n\mu_e}{l2^l} + \frac{n(2n \Leftrightarrow 1)}{2^{l+1}}.$$

$\square$

Proof of Lemma 2

We have to show that for the IGE\$-z_0 encryption mode, whenever the adversary knows only valid ciphertext strings (by the definition of EF-CoA), any forgery y passes the integrity check with negligible probability. In CoA, the plaintext strings used for generating the valid ciphertexts the adversary sees are random strings.

The forged ciphertext that the adversary generates can fall into one of the following complementary classes:
(a) the forgery is a truncation of a known valid ciphertext string;
(b) the forgery is an extension of a known valid ciphertext string;
(c) the forgery is neither a truncation nor an extension of a known ciphertext string.

In case (c), the forged ciphertext y can be such that either (c1) $y_0 = y_0^p$ for some $p, 1 \quad p \quad q_e$, or (c2) $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$; in the former case, the forged ciphertext and ciphertext string y^i will differ from each other in at least one block $y_k, 1 \quad k \quad \min(n_i + 1, n + 1)$. Hence, case (c) can be further divided into two complementary subcases:
(c1) the forged ciphertext string has a common prefix with an existent ciphertext;
(c2) the forged ciphertext is different from any existent ciphertext starting with its first block (y_0).

We summarize these classes of forgeries and define them formally. The forged ciphertext y belongs to one of the following complementary classes defined as follows:
(a) $\exists i, 1 \quad i \quad q_e : n \quad n_i$ and $\forall j, 1 \quad j \quad n + 1 : y_k = y_k^i$; i.e., the forged ciphertext is a truncation of ciphertext y^i;
(b) $\exists i, 1 \quad i \quad q_e : n > n_i$ and $\forall j, 1 \quad j \quad n_i + 1 : y_k = y_k^i$; i.e., the forged ciphertext is a extension of ciphertext y^i;
(c1) $\exists i, 1 \quad i \quad q_e, \exists j, 1 \quad j \quad \min(n_i + 1, n + 1) : \forall k, 1 \quad k \quad j : y_k = y_k^i$ and $y_j \neq y_j^i$; i.e., the forged ciphertext and ciphertext y^i have a common prefix;
(c2) $y_0 \neq y_0^i, \forall i, 1 \quad i \quad q_e$; i.e., there is no previous ciphertext that has a common prefix with the forged ciphertext.

Now, we show that, for an arbitrary forgery in each of the complementary cases defined above, the probability of adversary's success is negligible. We determine upper bounds on $Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow z_0)(y) \neq Null]$ and the maximum of these bounds is an upper bound for $Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow z_0)(y) \neq Null]$ for any forgery type.

By using Fact 0, we have

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[Succ] \quad \epsilon + Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[Succ],$$

where $Succ \equiv (x_{n+1} = z_0)$. Hence, for the balance of this proof, we use the notation $Pr[.] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[.]$, unless otherwise specified.

Upper bound for forgeries of type (a).
In this case, the forgery is a truncation of ciphertext i, and hence, the decrypted plaintext blocks are: $x_k = x_k^i, \forall k, 0 \quad k \quad n + 1 \quad n_i + 1$. Thus, the integrity condition $x_{n+1} = z_0$ becomes $x_{n+1}^i = z_0^i$ and hence, it happens with probability $1/2^l$; i.e.,

$$Pr[x_{n+1} = z_0] = \frac{1}{2^l},$$

since x_{n+1}^i is random, uniformly distributed, and independent of anything else, by the definition of CoAs.

Upper bound for forgeries of type (b).
In this case, the forgery is a extension of ciphertext i, and hence, the decrypted plaintext blocks are: $x_k = x_k^i, \forall k, 0 \quad k \quad n_i + 1 \qquad + 1.$ $\qquad <n$

Since $n + 1 > n_i + 1$, there must exist a ciphertext block y_{n_i+2}. To compute an upper bound on the probability of successful forgery, we condition on the event of collisions between $y_{n_i+2} \oplus x_{n_i+1}$ with $y_k^p \oplus x_{k-1}^p, \forall p, k, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$. Let D be the event defining the collisions $y_{n_i+2} \oplus x_{n_i+1} = y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$, or using the definition for set S^e, $y_{n_i+2} \oplus x_{n_i+1} \in S^e$; we obtain

$$D : y_{n_i+2} \oplus x_{n_i+1} \in S^e.$$

By union bound,

$$Pr[D] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[y_{n_i+2} \oplus x_{n_i+1} = y_k^p \oplus x_{k-1}^p].$$

Since event D implies $y_{n_i+2} \oplus x_{n_i+1} = y_k^p \oplus x_{k-1}^p$, and since $x_{n_i+1} = x_{n_i+1}^i = z_0^i$, it follows that $y_{n_i+2} \oplus z_0^i = y_k^p \oplus x_{k-1}^p$. In this equality, x_{k-1}^p is random and uniformly distributed because either $x_{k-1}^p = x_0^p$ when $k \Leftrightarrow 1 = 0$ or x_{k-1}^p is a random block due to the CoA attack when $k \Leftrightarrow 1 \quad 1$. Furthermore, since z_0^i is encrypted with key K', it follows that x_{k-1}^p and z_0^i are independent. Since y_{n_i+2} and $y_k^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$ are chosen constants,

$$Pr[y_{n_i+2} \oplus z_0^i = y_k^p \oplus x_{k-1}^p] = \frac{1}{2^l}.$$

Thus, since $\frac{e}{l}$ includes all the ciphertext blocks $(y_0^p, y_1^p, \quad , y_{n+p+1}^p, 1 \quad p \quad q_e)$,

$$Pr[D] \quad \frac{\mu_e}{l2^l}.$$

If D is false, then we can choose $n_i + 2$ as the position of the first block that does *not* yield a collision with any element in S^e. Furthermore, by the Corollary to Lemma 1 [Main IGE Lemma] with $a = z_0$, we obtain,

$$Pr[x_{n+1} = z_0 \mid \overline{D}] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^{l+1}}.$$

Hence, by standard conditioning,

$$\begin{aligned} Pr[x_{n+1} = z_0] \quad & Pr[x_{n+1} = z_0 \mid \overline{D}] + Pr[D] \\ & \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^{l+1}} + \frac{\mu_e}{l2^l} = \frac{(n+2)\mu_e}{l2^l} + \frac{(n+1)^2}{2^{l+1}}. \end{aligned}$$

Upper bound for forgeries of type (c1).
In a similar manner to the proof for the forgeries of type (b), we condition the probability of successful forgery on the event of collisions between $y_j \oplus x_{j-1}$ and $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$. Let D_j the event defining these collisions. Formally,

$$D_j : y_j \oplus x_{j-1} \in S^e.$$

By union bound,

$$Pr[D_j] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[y_j \oplus x_{j-1} = y_k^p \oplus x_{k-1}^p].$$

Consider the collisions $y_j \oplus x_{j-1} = y_k^p \oplus x_{k-1}^p$. Since j is the first index such that $y_j \neq y_j^i$, it follows that $x_{j-1} = x_{j-1}^i$. Hence, these collisions can be expressed as $y_j \oplus x_{j-1}^i = y_k^p \oplus x_{k-1}^p$. In this equality, x_{j-1}^i

is random, uniformly distributed and independent of any $x_k^p, y_k^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$, with the exception of x_{j-1}^i, by the definition of CoA. For $p = i, k = j$, we have $x_{j-1}^i = x_{k-1}^p$, but by the definition of j ($y_j \neq y_j^i$), $y_j \oplus x_{j-1}^i \neq y_j^i \oplus x_{j-1}^i$. Since $y_j, y_k^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$ are constants, then

$$Pr[y_j \oplus x_{j-1}^i = y_k^p \oplus x_{k-1}^p] \quad \frac{1}{2^l}.$$

Note that $Pr[y_j \oplus x_{j-1}^i = y_k^p \oplus x_{k-1}^p] = 0$ for $i = p, j = k$ from the definition of y_j. Hence, by the same arguments as for the case of forgeries of type (b),

$$Pr[D_j] \quad \frac{\mu_e}{l2^l}.$$

Furthermore, in a manner similar to that for the case of forgeries of type (b),

$$Pr[x_{n+1} = z_0 \mid \overline{D_j}] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}},$$

by the Corollary to Lemma 1 [Main IGE Lemma] with $a = z_0$.

Upper bound for forgeries of type (c2).
In a similar manner to the proof for the forgeries of type (c1), we condition the probability of successful forgery on the event of collisions between $y_1 \oplus x_0$ and $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$. Hence, we define event D as for the case of forgeries of type (c1)

$$D : y_1 \oplus x_0 \in S^e.$$

By union bound,

$$Pr[D] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[y_1 \oplus x_0 = y_k^p \oplus x_{k-1}^p].$$

Thus, we consider the collision $y_1 \oplus x_0 = y_k^p \oplus x_{k-1}^p$. In this equality x_{k-1}^p is random and uniformly distributed since it is either r_0^p for $k \Leftrightarrow 1 = 0$ or is a random and uniformly distributed plaintext block in a CoA for $k \Leftrightarrow 1 \quad 1$. Furthermore, $x_0 = r_0 = f'^{-1}(y_0)$ is the decryption of block $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$ with a different key, hence x_0 is independent of anything else, and hence, it si independent of x_{k-1}^p. Therefore,

$$Pr[y_1 \oplus x_0 = y_k^p \oplus x_{k-1}^p] \quad \frac{1}{2^l},$$

and

$$Pr[D] \quad \frac{\mu_e}{l2^l}.$$

From here on, the computation of the upper bound for forgeries of type (c2) is similar with the computation for the upper bound for forgeries of type (c1) in which one chooses $j = 1$.

Hence, for any forgery type

$$Pr[x_{n+1} = z_0] \quad \frac{(n+2)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}},$$

i.e., the probability that the integrity check passes, or equivalently that the EF-CoA adversary is successful, is

$$Pr[(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow z_0)(y) \neq Null] \quad \frac{(n+2)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}} + \frac{1}{2^l},$$

and, by Fact 0

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow z_0)(y) \neq Null] \quad \epsilon + \frac{(n+2)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}} + \frac{1}{2^l};$$

i.e., this probability is negligible, and scheme IGE\$-z_0 is EF-CoA secure. □

Proof of Lemma 3

We prove that scheme IGE\$-c is KPF-CPA secure. Hence, we must show that, for any forgery y,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null \qquad (\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }] \quad \lambda,$$

where $1 \Leftrightarrow \lambda$ is negligible. By definition, $((\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null \qquad (\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }) \equiv ((\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null \text{ or } (\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known })$. Hence, we must show that

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null \text{ or } (\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }] \quad \lambda,$$

where $1 \Leftrightarrow \lambda$ is negligible.

For the balance of this proof, we use the notation $Pr[.] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[.]$, unless otherwise specified.

To prove this lemma, we divide the space of all possible forgeries into two complementary classes: (a) forgeries that have at least a ciphertext block y_i such that $y_i \oplus x_{i-1}$ does not collide with any element of S^e, $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$, and (b) forgeries for which any block leads to a collision with some element of S^e, $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$.

Let y be an arbitrary forgery in class (a), and i the index of the first block such that $y_i \oplus x_{i-1}$ does not collide with any elements of S^e, $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$. Then, since c is a constant, by Lemma 1 [Main IGE Lemma] with $a = c$,

$$Pr[x_{n+1} = c] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}.$$

Hence, by the definition of event $(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null$:

$$Pr[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null] = Pr[x_{n+1} = c] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}.$$

and, by Fact 0

$$\begin{aligned} Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null] &= Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[x_{n+1} = c] \quad 1 \Leftrightarrow \lambda \\ &= \epsilon + \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}. \end{aligned}$$

Thus,

$$\begin{aligned} &Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null \text{ or } (\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }] \\ &Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null] = 1 \Leftrightarrow Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[(\mathcal{D}\Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null] \quad \lambda \end{aligned}$$

where $1 \Leftrightarrow \lambda = \epsilon + \frac{(n+1) \quad _e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}$ is negligible.

Let y be an arbitrary forgery in class (b); i.e., for any i a block, $y_i \oplus x_{i-1}$ collides with any element of S^e, $y_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1$. Hence, $x_i = f^{-1}(y_i \oplus x_{i-1}) \oplus y_{i-1} = x_k^p \oplus y_{k-1}^p \oplus y_{i-1}$ is known. If the last decrypted plaintext block leads to $x_{n+1} = c$, then the ciphertext decrypts correctly and the adversary knows the entire plaintext outcome of forgery. Hence event $((\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null$ or $(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x$ is known) is true. If the last decrypted plaintext block leads to $x_{n+1} \neq c$, then the ciphertext does not decrypt correctly. Hence $(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null$, and thus event $((\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null$ or $(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x$ is known) is still true. Thus, for any forgery in class (b)

$$Pr_{f \stackrel{\mathcal{R}}{\leftarrow} F}[(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = Null \text{ or } (\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }] = 1.$$

Hence, for any forgery (either of class (a) or (b)),

$$Pr_{f \stackrel{\mathcal{R}}{\leftarrow} F}[(\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) \neq Null \qquad (\mathcal{D} \Leftrightarrow IGE\$ \Leftrightarrow c)(y) = x \text{ is known }] \quad \lambda,$$

where $1 \Leftrightarrow \lambda$ is negligible, and by the definition of security against known-plaintext forgeries, the IGE\$-c scheme is KPF-CPA secure.

Proof of Lemma 4

First, we prove that the IGE\$-z_0 and IGE\$-c encryption modes are not secure against EF-CPA attacks, and then we prove that these schemes are not PI-CPA, PU-CPA, and NM-CPA secure. To prove the first part of the lemma, it is sufficient to provide counter-examples that show that an adversary can construct a forgery y that passes the integrity check provided by $x_{n+1} = z_0$ for IGE\$-z_0 and the integrity check $x_{n+1} = c$ for IGE\$-c (and whose plaintext x is known to the adversary.)

We show that the adversary can choose a plaintext with certain properties, obtain the ciphertext, then, he can construct a forgery that yields some changes plaintext blocks only in the middle of the plaintext, and thus, the beginning and the ending of the plaintext are unmodified and, hence, the decrypted plaintext passes the integrity checks $x_{n+1} = z_0$ or $x_{n+1} = c$.

Let an adversary submit for encryption the chosen plaintext
$x = x_1 \quad x_{i-2}x_{i-1}x_ix_{i+1} \quad x_m$, where $x_{i-2} = x_i$ and $x_{i-1} = x_{i+1}$. That is, the adversary simply constructs a plaintext that replicates two consecutive blocks in the two positions that follow those blocks. The adversary obtains ciphertext $y = y_1 \quad y_{i-2}y_{i-1}y_iy_{i+1} \quad y_m$, and constructs forgery (of equal length, m) as follows:

$$y' = y_1' \quad y_{i-2}'y_{i-1}'y_i'y_{i+1}' \quad y_m',$$

where

$$\begin{aligned}
&y_1' \quad y_{i-2}' = y_1 \quad y_{i-2} \\
&y_{i-1}' = y_{i+1} \\
&y_i' = y_{i-2} \\
&y_{i+1}' \quad y_m' = y_{i+1} \quad y_m.
\end{aligned}$$

In other words, the forgery $y' \neq y$ is

$$y' = y_1 \quad y_{i-2}y_{i+1}y_{i-2}y_{i+1} \quad y_m.$$

Next, we describe the attack outcome. The decryption of forgery y', namely x', will contain (1) the same plaintext blocks as those of the chosen plaintext x up to position $i \Leftrightarrow 2$; i.e., $x_j' = x_j, \forall j, 1 \quad j \quad i \Leftrightarrow 2$; (2)

the same plaintext blocks as those of the chosen plaintext x from position $i+1$ to the end of the message; i.e., $x'_j = x_j, \forall j, i+1 \quad j \quad m$; and (3) two modified plaintext blocks (both with a known/predictable modification) at position $i \Leftrightarrow 1$, i.e., $x'_{i-1} = x_{i+1} \oplus y_i \oplus y_{i-2}$, and at position i; i.e., $x'_i = x_i \oplus y_{i-1} \oplus y_{i+1}$.

To verify the outcome of this attack, we compute x'_{i-1}, x'_i, and x'_{i-1}. That is,

$$\begin{aligned} x'_{i-1} &= f^{-1}(y'_{i-1} \oplus x'_{i-2}) \oplus y'_{i-2} = f^{-1}(y_{i+1} \oplus x_i) \oplus y_{i-2} \\ &= x_{i+1} \oplus y_i \oplus y_{i-2} \end{aligned}$$

which is *known* to the adversary.

$$\begin{aligned} x'_i &= f^{-1}(y'_i \oplus x'_{i-1}) \oplus y'_{i-1} = f^{-1}(y_{i-2} \oplus x_{i+1} \oplus y_i \oplus y_{i-2}) \oplus y_{i+1} \\ &= f^{-1}(x_{i-1} \oplus y_i) \oplus y_{i+1} = x_i \oplus y_{i-1} \oplus y_{i+1} \end{aligned}$$

which is *known* to the adversary.

$$\begin{aligned} x'_{i+1} &= f^{-1}(y'_{i+1} \oplus x'_i) \oplus y'_i = f^{-1}(y_{i+1} \oplus x_i \oplus y_{i-1} \oplus y_{i+1}) \oplus y_{i-2} \\ &= f^{-1}(x_i \oplus y_{i-1}) \oplus y_{i-2} = f^{-1}(x_{i-2} \oplus y_{i-1}) \oplus y_{i-2} = x_{i-1} = x_{i+1}. \end{aligned}$$

which means that the plaintext at position $i+1$ remains unmodified.

$$x'_{i+2} = f^{-1}(y'_{i+2} \oplus x'_{i+1}) \oplus y'_{i+1} = f^{-1}(y_{i+2} \oplus x_{i+1}) \oplus y_{i+1} = x_{i+2}.$$

which means that the plaintext at position $i+2$ also remains unmodified. From this point on, all remaining plaintext blocks remain unmodified to the end of the message.

Hence, the integrity conditions $x'_{n+1} = z_0$ for the IGE\$-z_0 or $x'_{n+1} = c$ for the IGE\$-c are verified with probability 1 (one), i.e., neither scheme is secure against EF-CPA.
The same counter-example as that given above is sufficient to show that the IGE\$-z_0 and IGE\$-c are not PU-CPA, and PI-CPA secure. (The actual proof for PI-CPA security involves the event that there are no collisions in the inputs to function f; i.e., includes the bound δ_R defined in the proof of Lemma 6, Fact 1 below.) A similar example can be used to prove that these schemes are not NM-CPA secure, also. For instance, construct a forgery in which all but the last two blocks of the plaintext outcome contain all 1's, and the last two blocks contain the known but garbled data produced by the exclusive-or operations with ciphertext blocks obtained at encryption. Modify the plaintext outcome of the forgery as follows: divide (i.e., by integer division) the plaintext outcome of the forgery by 2^{2l}, where l is the block size, thereby shifting the garbled blocks out of the message and zero-filling its first two blocks. The relationship $\mathcal{R} \stackrel{def}{=}$ holds among the modified plaintext outcome and the similarly modified (but unknown) plaintext of the challenge ciphertexts. □

Proof of Lemma 5

To prove this lemma, we partition all possible forgeries into successively smaller classes, and demonstrate that, for each class of forgery, either the integrity check fails or the plaintext outcome of forgery includes an unknown block.

We note that all forgeries can be created in the following three complementary ways. That is, a forgery $y' = y'_0 y'_1 \quad y'_n y'_{n+1}$ can be:
(1) a truncation of a ciphertext message y^p_k of length n_p+1 obtained at encryption, namely, $y'_j = y^p_j, \forall j, 0$ $j \quad n+1 \quad {}_p + 1;$ $<n$

(2) an extensions of a ciphertext message y^p of length n_p+1 obtained at encryption, namely, $y'_j = y^p_j, \forall j, 0$ $j \quad n_p + 1 \quad + 1$; and $<n$
(3) in neither class (1) nor (2). That is, the forgery is a ciphertext message such that there exists index $s \quad \min\{n+1, n_p+1\} : y'_s \neq y^p_s$ whose ciphertext block differs from block s of a ciphertext message y^p of length $n_p + 1$ obtained at encryption. We denote by j be the minimum of these indices s.

It is easy to see that for forgeries of types (1) and (2) the lemma is proved, since for case (1) the integrity check passes with only negligible probability whereas for case (2) plaintext block x'_{n_p+1}, which contains random block z^p_0, is unknown, and hence could not be chosen by the adversary. That is, for any forgery of type (1), $x'_{n+1} = x^p_{n+1}$ is a constant since $n \quad n_p$, and z'_0 is a random variable. Thus

$$\begin{aligned} Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[z'_0 = x'_{n+1}] &= Pr_{f,f' \overset{\mathcal{R}}{\leftarrow} P^l}[z'_0 = x'_{n+1}] \\ &= Pr_{f,f' \overset{\mathcal{R}}{\leftarrow} P^l}[z'_0 = x'_{n+1}] \Leftrightarrow Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[z'_0 = x'_{n+1}] + Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[z'_0 = x'_{n+1}] \\ &= Adv_{\mathcal{D}}(P^l, R^{l,l}) + \frac{1}{2^l}, \end{aligned}$$

where $Adv_{\mathcal{D}}(P^l, R^{l,l})$ is the advantage of an adversary $\mathcal{D}$ in distinguishing between $f' \overset{\mathcal{R}}{} P^l$ from $f' \overset{\mathcal{R}}{} R^{l,l}$ using an encryption oracle for f' in the process of implementing the IGE$ scheme. Also, since random variable z'_0 is uniformly distributed when $f' \overset{\mathcal{R}}{} R^{l,l}$ and since x'_{n+1} is a constant, it follows that $Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[z'_0 = x'_{n+1}] = 1/2^l$. However, by the bound of the birthday attack, $Adv_{\mathcal{D}}(P^l, R^{l,l}) \quad \frac{q_e(q_e-1)}{2^{l+1}}$ since $z'_0 = z^p_0 = f'(r^p_0 + 1)$ and $1 \quad p \quad q_e$. Hence,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null \text{ and } ((D^{F_K} o\ g)(y') = x \neq x^i, 1 \quad i \quad q_e, \text{ is chosen}]$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[z'_0 = x'_{n+1}] = \frac{1}{2^l} + \frac{q_e(q_e \Leftrightarrow 1)}{2^{l+1}}.$$

For any forgery of type (2), $x'_{n_p+1} = x^p_{n_p+1} = z^p_0$, which is random. Hence, event $x'_{n_p+1} = x_{n_p+1}$ has the same distribution as z^p_0 and happens with probability $\frac{1}{2^l} + \frac{q_e(q_e-1)}{2^{l+1}}$ whenever $f' \overset{\mathcal{R}}{} P^l$ (by the same argument as in case (1)). Hence,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null \text{ and } ((D^{F_K} o\ g)(y') = x \neq x^i, 1 \quad i \quad q_e, \text{ is chosen}]$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_{n_p+1} = x_{n_p+1}] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x_{n_p+1} = z^p_0] = Pr_{f' \overset{\mathcal{R}}{\leftarrow} P}[x_{n_p+1} = z^p_0]$$
$$= \frac{1}{2^l} + \frac{q_e(q_e \Leftrightarrow 1)}{2^{l+1}}.$$

To complete the proof of the lemma, we partition forgeries of type (3) further. We first distinguish the case whereby there exists a ciphertext block position $j, 0 \quad j \quad n+1$, such that the input to f^{-1} at that block position does not collide with any of possible inputs to f used during encryption. That is, $y'_0 \neq y^p_0, \forall p, 1 \quad p \quad q_e$ or $y'_j \oplus x'_{j-1} \notin S^e$. Then, by the Corollary to the Main IGE Lemma (Lemma 1),

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_{n+1} = z'_0] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}.$$

Hence,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null \text{ and } ((D^{F_K} o\ g)(y') = x \neq x^i, 1 \quad i \quad q_e, \text{ is chosen}]$$
$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_{n+1} = z'_0] \quad \frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}.$$

The lemma is proven for this case also.

In all remaining type (3) cases, all inputs to f^{-1} during decryption collide with some inputs to f used during encryption. That is, $y'_0 = y^p_0$, for some $p, 1 \quad p \quad q_e$ or all $y'_j \oplus x'_{j-1} \in S^e, \forall j, 1 \quad j \quad n$.

Let a type (3) forgery y' differ from any of the q_e encrypted messages at block position j, $1 \quad j \quad n+1$; i.e., the adversary chooses $x'_i = x^p_i, \forall i, 1 \quad i \quad j \Leftrightarrow 1$. We show that plaintext obtained at position j during the decryption of the forgery y', namely x'_j, can be chosen only with negligible probability, or that the integrity condition happens with negligible probability. This completes the proof since the maximum of all the probabilities of passing the integrity check and choosing all the plaintext of the forgery decryption is negligible.

If the adversary chooses $x'_i = x^p_i$, the chosen plaintext blocks could be obtained up to position j of the forgery decryption. Now, we show that the chosen plaintext can be obtained at position j with only negligible probability. We have two complementary cases to analyze: (a) $j \quad n$ and (b) $j = n+1$.

(a) For $j \quad n$, we compute an upper bound on the probability of the integrity condition $x'_j = x_j$, where x_j is the chosen value. However, by definition, $x'_j = f^{-1}(y'_j \oplus x'_{j-1}) \oplus y'_{j-1}$, and by hypothesis, $y'_j \oplus x'_{j-1} \in S^e, \forall j, 1 \quad j \quad n$. Thus, a collision $y'_j \oplus x'_{j-1} = y^s_t \oplus x^s_{t-1}$ must take place for some s, t, $1 \quad t \quad n_s + 1, 1 \quad s \quad q_e$.

If $y'_j \oplus x'_{j-1} = y^s_t \oplus x^s_{t-1}$, $1 \quad s \quad q_e, 1 \quad t \quad n_s + 1$, then, since $x'_j = f^{-1}(y'_j \oplus x'_{j-1}) \oplus y'_{j-1}$ and $y'_{j-1} = y^p_{j-1}$ by the definition of block position j, we obtain $x'_j = x^s_t \oplus y^s_{t-1} \oplus y'_{j-1} = x^s_t \oplus y^s_{t-1} \oplus y^p_{j-1}$. Now note that $(s,t) \neq (p,j) \;\Leftrightarrow\; (s, t \Leftrightarrow 1) \neq (p, j \Leftrightarrow 1)$ by the definition of block position j. This means that $x'_j = x_j \;\Leftrightarrow\; x^s_t \oplus y^s_{t-1} = x_j \oplus y^p_{j-1}$. However, the two sides of the equation $x^s_t \oplus y^s_{t-1} = x_j \oplus y^p_{j-1}$ are random because x^s_t, x_j are chosen constants and y^s_{t-1}, y^p_{j-1} are random since $f \overset{\mathcal{R}}{\ } P^l$. The two sides of the equation are also independent of each other whenever y^s_{t-1} and y^p_{j-1} are distinct (i.e., do not collide with each other). To compute the probability that y^s_{t-1} and y^p_{j-1} are distinct, we define D (*Distinct*) to be the event that all inputs to function $f = F_K$ used during the q_e encryptions are distinct. Fact 1 provides a bound for the probability of $\overline{D}$.

Fact 1
Let D *Distinct* denote the event at all inputs to function $f = F_K$ used during the q_e encryptions $y^p_k = f(x^p_k \oplus y^p_{k-1}) \oplus x^p_{k-1}, 1 \quad p \quad q_e, 1 \quad k \quad n_p$, are distinct. Then,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[\overline{D}] \quad \delta_R \overset{def}{=} \frac{1}{2^{l+1}} \left(\frac{\mu_e^2}{l^2} \Leftrightarrow \frac{\mu_e}{l} \right)$$

and

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\overline{D}] \quad \delta_P \overset{def}{=} \delta_R + \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} = \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^l}.$$

Then, the probability of event $x'_j = x_j \;\Leftrightarrow\; x^s_t \oplus y^s_{t-1} = x_j \oplus y^p_{j-1}$ can be bound by using standard conditioning and Fact 1.

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j] \quad Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j \mid D] + Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\overline{D}] \quad Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j \mid D] + \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^l}.$$

However, by the same argument as that used in (1), we obtain

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j \mid D] = Adv_{\mathcal{D}}(P^l, R^{l,l}) + Pr_{f \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[x'_j = x_j \mid D],$$

or

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j \mid D] \quad \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} + \frac{1}{2^l},$$

since, when $f \;\; ^{\mathcal{R}} \; R^{l,l}$ and event D is true, y^s_{t-1} and y^p_{j-1} where $(s, t \Leftrightarrow 1) \neq (p, j \Leftrightarrow 1)$, are random, uniformly distributed, and independent, and thus $Pr_{f \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[x'_j = x_j | D] = 1/2^l$. Hence,

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j] \quad \frac{3\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} + \frac{1}{2^l},$$

which shows that $Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = x_j]$ is negligible.

(b) For $j = n+1$, we compute an upper bound for the probability of the integrity condition $x'_j = z'_0$. (The proof of the negligible upper bound for this case is almost identical to that for case $j \quad n$. We repeat it here for completeness.) However, by definition $x'_j = f^{-1}(y'_j \oplus x'_{j-1}) \oplus y'_{j-1}$, and by hypothesis $y'_j \oplus x'_{j-1} \in S^e$. Hence, a collision $y'_j \oplus x'_{j-1} = y^s_t \oplus x^s_{t-1}$ must take place for some s, t, $1 \quad t \quad n_s + 1, 1 \quad s \quad q_e$.

If $y'_j \oplus x'_{j-1} = y^s_t \oplus x^s_{t-1}$, $1 \quad s \quad q_e, 1 \quad t \quad n_s + 1$, then, since $x'_j = f^{-1}(y'_j \oplus x'_{j-1}) \oplus y'_{j-1}$ and $y'_{j-1} = y^p_{j-1}$ by the definition of of block position j, we obtain $x'_j = x^s_t \oplus y^s_{t-1} \oplus y'_{j-1} = x^s_t \oplus y^s_{t-1} \oplus y^p_{j-1}$. Note that $(s,t) \neq (p,j) \;\Leftrightarrow\; (s, t \Leftrightarrow 1) \neq (p, j \Leftrightarrow 1)$ by the definition of block position j.

The integrity condition $x'_j = z'_0 \;\Leftrightarrow\; x^s_t \oplus y^s_{t-1} = z^p_0 \oplus y^p_{j-1}$, where the right hand side is random and independent of the left hand side. This is the case because z^p_0 is random and independent of y^s_{t-1} and y^p_{j-1}, since it is generated using function f' with key $K' \neq K$, and $x^s_t \neq x^p_j = z^p_0$, since block position $(s,t) \neq (p,j), j = n+1$, and x^s_t is a chosen constant. Using the same arguments as in case (a), we obtain an upper bound for the probability of $x'_j = z'_0$, as follows:

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[x'_j = z'_0] = Pr_{f' \overset{\mathcal{R}}{\leftarrow} P}[x'_j = z'_0] \quad \frac{1}{2^l} + \frac{q_e(q_e \Leftrightarrow 1)}{2^{l+1}}.$$

Finally, for any possible forgery, the probability of success is bounded by the maximum of the probabilities obtained for cases (1) - 3(a)(b); i.e.,

$$\begin{aligned} &Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[((D^{F_K} o\ g)(y') \neq Null \text{ and } ((D^{F_K} o\ g)(y') = x \neq x^i \text{ is chosen}, 1 \quad i \quad q_e] \\ &\max\left\{\frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}, \frac{3\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} + \frac{1}{2^l}, \frac{1}{2^l} + \frac{q_e(q_e \Leftrightarrow 1)}{2^{l+1}}\right\}. \end{aligned}$$

Hence, when the scheme is implemented with the SPRP family F,

$$\begin{aligned} &Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[((D^{F_K} o\ g)(y') \neq Null \text{ and } ((D^{F_K} o\ g)(y') = x \neq x^i, 1 \quad i \quad q_e \text{ is chosen}] \quad \epsilon' \stackrel{def}{=} \\ &\max\left\{\frac{(n+1)\mu_e}{l2^l} + \frac{(n+1)(2n+1)}{2^{l+1}}, \frac{3\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} + \frac{1}{2^l}, \frac{1}{2^l} + \frac{q_e(q_e \Leftrightarrow 1)}{2^{l+1}}\right\} + \epsilon, \end{aligned}$$

and ϵ' is negligible. □

Proof of Lemma 6

This proof is based first on replacing SPRP family F with the family of random functions G_S, i.e., $f, \overline{f} \;\; ^{\mathcal{R}} \; G_S$. Next, we use the idea that if the inputs to function $\overline{f}$ in the reverse pass of the decryption are different from all the quantities obtained at encryption (either from the unknown plaintext of the challenges or the plaintext the adversary chooses to encrypt), and if they are different (i.e., do not collide among themselves), the plaintext outcome of the forgery is random, uniformly distributed, and independent of anything else (since, for these input, $\overline{f} = v$ and $v \ \overset{\mathcal{R}}{\ } \ R^{l,l}$). Hence, for the most part, the proof focuses on determining upper bounds for these events.

Let $q_e = q_1 + q_2$ with q_1, q_2 defined in the NM-CPA, and define the following sets (encompassing both the unknown plaintexts corresponding to the ciphertext challenges, and the plaintexts chosen by the adversary):

$$\begin{aligned} S^e &= \{z_k^p \oplus x_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1\} \\ S^d &= \{z_s \oplus x_{s-1}, 1 \quad s \quad n+1\}. \\ T^e &= \{x_k^p \oplus z_{k-1}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1\}. \end{aligned}$$

If the elements of the set S^d do not collide with each other (i.e., the set S^e is collision-free) and $S^e \cap S^d = \phi$ (i.e., the empty set), then the inputs to the functions $\overline{f}$ at decryption are new, and hence the quantities $\overline{f}(z_s \oplus x_{s-1}) = v(z_s \oplus x_{s-1})$ are random, uniformly distributed, and mutually independent and independent of anything else. Furthermore, all plaintexts $x_s = \overline{f}(z_s \oplus x_{s-1}) \oplus z_{s-1}$ are random, uniformly distributed, and mutually independent and independent of anything else. Hence, there is no relationship among the decrypted plaintext and the challenge plaintexts.

Let us define the following events:

$$\begin{aligned} D &: T^e \text{ is collision-free} \\ A &: S^e \cap S^d = \phi \\ B &: S^d \text{ is collision-free.} \end{aligned}$$

Event D is the event *Distinct* from Fact 1, hence

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\overline{D}] \quad \delta_P \overset{def}{=} \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^l}.$$

In the following we consider $Pr[.] = Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[.]$ and drop the subscript.

If both the events A and B are true, then the event $\mathcal{R}(x^1, \quad, x^{q_2}, (D^{F_K} o\ g)(y))$ is false, i.e., there does not exist any relationship between the decrypted plaintext and the challenge plaintexts. Hence, the following implication is true: $\mathcal{R}(x^1, \quad, x^{q_2}, (D^{F_K} o\ g)(y)) \qquad \overline{A \text{ and } B} \equiv \overline{A}$ or $\overline{B}$. Hence,

$$(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \quad, x^{q_2}, (D^{F_K} o\ g)(y)) \qquad (D^{F_K} o\ g)(y) \neq Null \text{ and } (\overline{A} \text{ or } \overline{B})$$

Hence,

$$\begin{aligned} &Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \quad, x^{q_2}, (D^{F_K} o\ g)(y))] \\ &\quad Pr[((D^{F_K} o\ g)(y) \neq Null \text{ and } (\overline{A} \text{ or } \overline{B})] \quad Pr[\overline{A} \text{ or } \overline{B}]. \end{aligned}$$

Now, we compute an upper bound for the probability of event $\overline{A}$ or $\overline{B}$.

Let us define the following set:

$$S_i^d = \{z_s \oplus x_{s-1}, 1 \quad s \quad i\}$$

and events:

$$\begin{aligned} A_i &: S^e \cap S_i^d = \phi \\ B_i &: S_i^d \text{ is collision-free.} \end{aligned}$$

Hence, event $A = A_{n+1}$ and event $B = B_{n+1}$. For any index i, we obtain, by standard conditioning,

$$Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}}] \quad Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i] + Pr[\overline{A_i} \text{ or } \overline{B_i}],$$

and, using standard conditioning repeatedly, we obtain

$$\begin{aligned} Pr[\overline{A} \text{ or } \overline{B}] &= Pr[\overline{A_{n+1}} \text{ or } \overline{B_{n+1}}] \\ & \quad Pr[\overline{A_{n+1}} \text{ or } \overline{B_{n+1}} \mid A_n \text{ and } B_n] + Pr[\overline{A_n} \text{ or } \overline{B_n}]. \\ & \quad Pr[\overline{A_1} \text{ or } \overline{B_1}] + \sum_{i=1}^{n} Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i]. \end{aligned}$$

First, we determine an upper bound for $Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i]$. By union bound,

$$\begin{aligned} Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i] & \quad Pr[\overline{A_{i+1}} \mid A_i \text{ and } B_i] + Pr[\overline{B_{i+1}} \mid A_i \text{ and } B_i] \\ &= Pr[z_{i+1} \oplus x_i \in S^e \mid A_i \text{ and } B_i] + Pr[z_{i+1} \oplus x_i \in S_i^d \mid A_i \text{ and } B_i]. \end{aligned}$$

To see this, note that if event A_i is true, then $S^e \cap S_i^d = \phi$. Hence, since $S_{i+1}^d = S_i^d \cup \{z_{i+1} \oplus x_i\}$, then, for $S^e \cap S_{i+1}^d \neq \phi$, $z_{i+1} \oplus x_i$ must be in S^e. Hence, $Pr[\overline{A_{i+1}} \mid A_i \text{ and } B_i] = Pr[z_{i+1} \oplus x_i \in S^e \mid A_i \text{ and } B_i]$. Similarly, if event B_i is true, i.e., S_i^d is collision-free, then for B_{i+1} to be false, $z_{i+1} \oplus x_i$ must be in S_i^d. Hence, $Pr[\overline{B_{i+1}} \mid A_i \text{ and } B_i] = Pr[z_{i+1} \oplus x_i \in S_i^d \mid A_i \text{ and } B_i]$.

Furthermore, by union bound,

$$\begin{aligned} Pr[z_{i+1} \oplus x_i \in S^e \mid A_i \text{ and } B_i] & \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[z_{i+1} \oplus x_i = z_k^p \oplus x_{k-1}^p \mid A_i \text{ and } B_i] \\ Pr[z_{i+1} \oplus x_i \in S_i^d \mid A_i \text{ and } B_i] & \quad \sum_{j=1}^{i} Pr[z_{i+1} \oplus x_i = z_j \oplus x_{j-1} \mid A_i \text{ and } B_i]. \end{aligned}$$

Whenever A_i and B_i are true, element $z_i \oplus x_{i-1}$ has never been seen before, and hence $\overline{f}(z_i \oplus x_{i-1}) = v(z_i \oplus x_{i-1})$ is random, uniformly distributed and independent of anything else. Thus, $x_i = \overline{f}(z_i \oplus x_{i-1}) \oplus z_{i-1}$ is random, uniformly distributed, and independent of anything else, and each of the events $z_{i+1} \oplus x_i = z_k^p \oplus x_{k-1}^p$ and $z_{i+1} \oplus x_i = z_j \oplus x_{j-1}$ happens with probability $1/2^l$. Hence,

$$\begin{aligned} Pr[z_{i+1} \oplus x_i \in S^e \mid A_i \text{ and } B_i] & \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[z_{i+1} \oplus x_i = z_k^p \oplus x_{k-1}^p \mid A_i \text{ and } B_i] = \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} \frac{1}{2^l} \quad \frac{\mu_e}{l2^l} \\ Pr[z_{i+1} \oplus x_i \in S_i^d \mid A_i \text{ and } B_i] & \quad \sum_{j=1}^{i} Pr[z_{i+1} \oplus x_i = z_j \oplus x_{j-1} \mid A_i \text{ and } B_i] = \sum_{j=1}^{i} \frac{1}{2^l} = \frac{i}{2^l}. \end{aligned}$$

Then

$$Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i] \quad \frac{\mu_e}{l2^l} + \frac{i}{2^l}.$$

Second, we find an upper bound for $Pr[\overline{A_1} \text{ or } \overline{B_1}]$. $S_1^d = \{z_1 \oplus x_0\}$ has only one element, and hence it is collision free. Therefore, event B_1 is always true. Hence, we find an upper bound for $Pr[\overline{A_1}]$. We introduce event

$$C : z_0 \neq z_0^p \text{ and } z_0 \neq y_0^p, \quad \forall p, 1 \quad p \quad q_e.$$

By standard conditioning,

$$Pr[\overline{A_1}] \quad Pr[\overline{A_1} \mid C] + Pr[\overline{C}].$$

By union bound,

$$Pr[\overline{A_1} \mid C] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[z_1 \oplus x_0 = z_k^p \oplus x_{k-1}^p \mid C].$$

Now, let us assume event C is true. In this case, z_0 has never been the input to $\overline{f'}$, and hence $x_0 = \overline{f'}(z_0) = v'(z_0), v' \ {}^{\mathcal{R}} \ R^{l,l}$ is random, uniformly distributed, and independent of anything else, hence

$$Pr[z_1 \oplus x_0 = z_k^p \oplus x_{k-1}^p \mid C] = \frac{1}{2^l}.$$

Thus,

$$\sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr[z_1 \oplus x_0 = z_k^p \oplus x_{k-1}^p \mid C] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} \frac{1}{2^l} \quad \frac{\mu_e}{l2^l},$$

and

$$Pr[\overline{A_1}] \quad \frac{\mu_e}{l2^l} + Pr[\overline{C}].$$

Now, we find an upper bound for $Pr[\overline{C}]$. Using the conditioning on the event D (*Distinct*) and standard conditioning, we obtain:

$$Pr[\overline{C}] \quad Pr[\overline{C} \mid D] + Pr[\overline{D}],$$

where, by Fact 1 and the fact that $f \ {}^{\mathcal{R}} \ G_S$ means that $f, f', f'' \ {}^{\mathcal{R}} \ P^l$, we have

$$Pr[\overline{D}] = Pr_{f \overset{\mathcal{R}}{\leftarrow} P^l}[\overline{D}] \quad \delta_P \overset{def}{=} \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^l}.$$

Next, we use the following claim (whose proof is at the end of this Lemma).

Claim

$$Pr[\overline{C} \mid D] \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

Thus,

$$Pr[\overline{C}] \quad \delta_P + \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

Hence,

$$Pr[\overline{A_1} \text{ or } \overline{B_1}] \quad \frac{\mu_e}{l2^l} + \delta_P + \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

Furthermore,

$$\begin{aligned}
Pr[\overline{A} \text{ or } \overline{B}] \quad & Pr[\overline{A_1} \text{ or } \overline{B_1}] + \sum_{i=1}^{n} Pr[\overline{A_{i+1}} \text{ or } \overline{B_{i+1}} \mid A_i \text{ and } B_i] \\
& \delta_P + \frac{\mu_e}{l2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \sum_{i=1}^{n} \left(\frac{\mu_e}{l2^l} + \frac{i}{2^l} \right) \\
& \delta_P + \frac{\mu_e}{l2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{n\mu_e}{l2^l} + \frac{n^2}{2^{l+1}} \\
& \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l2^l} + \frac{3(n+1)^2}{2^{l+1}}.
\end{aligned}$$

Finally,

$$\begin{aligned}
&Pr[(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y))] \\
&\quad Pr[((D^{F_K} o\ g)(y) \neq Null \text{ and } (\overline{A} \text{ or } \overline{B})] \quad Pr[\overline{A} \text{ or } \overline{B}] \\
&\quad \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l 2^l} + \frac{3(n+1)^2}{2^{l+1}}.
\end{aligned}$$

Hence, when the scheme is implemented with the pseudo-random family F, by Fact 0 (with $\mu_v/l = 2(n+1)$), we have

$$\begin{aligned}
&Pr_{f \stackrel{\mathcal{R}}{\leftarrow} F}[(D^{F_K} o\ g)(y) \neq Null \text{ and } \mathcal{R}(x^1, \quad , x^{q_2}, (D^{F_K} o\ g)(y))] \\
&\quad \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l 2^l} + \frac{3(n+1)^2}{2^{l+1}} + \frac{2(n+1)(2n+1)}{2^{l+1}} + \epsilon;
\end{aligned}$$

i.e., the scheme is NM-CPA secure. □

Proof of Claim

We introduce the set of all inputs to function $\overline{f''}$ at decryption in the reversed direction, namely

$$R^e = \{y_k^p \oplus z_{n_p-k+2}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p + 1\}.$$

Note that z_0^p does not appear in the definition of set R^e.

To compute $Pr[\overline{C} \mid D]$ we divide the choice of ciphertext forgeries into several complementary classes, then compute the probability for each class of forgeries. The forged ciphertext that the adversary generates can fall into one of the following complementary classes:
(a) the forgery is a truncation of a known valid ciphertext string;
(b) the forgery is an extension of a known valid ciphertext string;
(c) the forgery is neither a truncation nor an extension of a known ciphertext string. Case (c) can be further divided into two complementary subcases:
(c1) the forged ciphertext string has a common prefix with an existent ciphertext;
(c2) the forged ciphertext is different from any existent ciphertext starting with its first block (y_0).

For each classes of forgery we find an upper bound the probability that z_0 collides with some z_0^p or y_0^p.

(a) If the forgery is a truncation of a valid ciphertext, then there exists $s, 1 \quad s \quad q_e : y = y_0 y_1 \quad y_{n+1}, y_k = y_k^s, \forall k, 0 \quad k \quad n+1 \quad n_s + 1$. Then $z_0 = z_{n_s-n}^s$ by the definition of the BIGE$ decryption. [4] Then, we have the collision between $z_{n_s-n}^s$ and z_0^p, or between $z_{n_s-n}^s$ and $y_0^p, 1 \quad p \quad q_e$, w here z_0^p and y_0^p are computed by enciphering with a different key. Furthermore, $Pr[\overline{C} \mid D] = Pr_{f \stackrel{\mathcal{R}}{\leftarrow} G_S}[\overline{C} \mid D]$, (based on our notation), then $f, f', f'' \stackrel{\mathcal{R}}{} P^l$. Hence, an adversary can distinguish between $f' \stackrel{\mathcal{R}}{} P^l$ and $f' \stackrel{\mathcal{R}}{} R^{l,l}$ in the computation of z_0^p or $y_0^p, 1 \quad p \quad q_e$. Hence,

$$\begin{aligned}
Pr[\overline{C} \mid D] \quad &= \quad Pr_{f \stackrel{\mathcal{R}}{\leftarrow} G_S}[\overline{C} \mid D] \stackrel{def}{=} Pr_{\overline{f}, \overline{f'}, \overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G_S, f, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l}[\overline{C} \mid D] \\
&= \quad Pr_{\overline{f}, \overline{f'}, \overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G_S, f, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l}[\overline{C} \mid D] \Leftrightarrow Pr_{\overline{f}, \overline{f'}, \overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G_S, f, f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f' \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[\overline{C} \mid D] \\
&+ \quad Pr_{\overline{f}, \overline{f'}, \overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G_S, f, f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f' \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[\overline{C} \mid D] \quad Adv_{\mathcal{D}}(P^l, R^{l,l}) + Pr_{\overline{f}, \overline{f'}, \overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G_S, f, f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f' \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[\overline{C} \mid D],
\end{aligned}$$

[4] Since $y_0 = y_0^s$, then $z_{n+1} = z_{n_s+1}^s$; furthermore, if $y_1 = y_1^s$ then $z_n = f'^{-1}(y_1 \oplus z_{n+1}) \oplus y_0 = f'^{-1}(y_1^s \oplus z_{n_s+1}^s) \oplus y_0^s = z_{n_s}^s$; etc.

where the advantage refers to distinguishing between $f' \overset{\mathcal{R}}{} P^l$ and $f' \overset{\mathcal{R}}{} R^{l,l}$. Since there are $2q_e$ queries to f', it follows that

$$Adv_{\mathcal{D}}(P^l, R^{l,l}) \quad \frac{2q_e(2q_e \Leftrightarrow 1)}{2^{l+1}} = \frac{q_e(2q_e \Leftrightarrow 1)}{2^l}.$$

We introduce the notation $Pr'[\overline{C} \mid D] \overset{def}{=} Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f'' \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[\overline{C} \mid D]$, and we compute an upper bound for $Pr'[\overline{C} \mid D]$. By union bound,

$$Pr'[\overline{C} \mid D] \quad \sum_{p=1}^{q_e} \left(Pr'[z_0 = z_0^p \mid D] + Pr'[z_0 = y_0^p \mid D]\right).$$

Since $z_0^p = f'(r_0^p)$ and $y_0^p = f'(z_{n_p+1}^p)$ are encrypted with a different key than the one used to obtain $z_{n_s-n}^s, n_s \Leftrightarrow n \quad 1$, then z_0^p and y_0^p are random, uniformly distributed, and independent of $z_{n_s-n}^s$ since $f' \overset{\mathcal{R}}{} R^{l,l}$. Hence,

$$\begin{aligned} Pr'[z_0 = z_0^p \mid D] &= Pr'[z_{n_s-n}^s = z_0^p \mid D] = \frac{1}{2^l} \\ Pr'[z_0 = y_0^p \mid D] &= Pr'[z_{n_s-n}^s = y_0^p \mid D] = \frac{1}{2^l}. \end{aligned}$$

Hence, by union bound,

$$Pr'[\overline{C} \mid D] \quad \frac{2q_e}{2^l}.$$

Hence,

$$Pr[\overline{C} \mid D] \quad \frac{q_e(2q_e \Leftrightarrow 1)}{2^l} + \frac{2q_e}{2^l} = \frac{q_e(2q_e + 1)}{2^l}.$$

(b) If $y = y_0^i y_1^i \quad y_{n_i+1}^i \quad y_{n+1}$ where $n > n_i$, then we show that $y_{n_i+2} \oplus z_{n-n_i} \in R^e$ with low probability, and this enables us to show that events $z_0 = z_0^p$ or $z_0 = y_0^p$ occur with low probability in a manner similar to the Main IGE Lemma. Hence, by standard conditioning we have

$$Pr[\overline{C} \mid D] \quad Pr[\overline{C} \mid D \text{ and } y_{n_i+2} \oplus z_{n-n_i} \notin R^e] + Pr[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D].$$

If $y_{n_i+2} \oplus z_{n-n_i} \notin R^e$ and $f \overset{\mathcal{R}}{} G_S$, then $z_0 = z_0^p$ happens with probability $\frac{(n+1)\ \ e}{l2^l} + \frac{(n+1)^2}{2^{l+1}}$ in a manner similar to the Corollary to the Main IGE Lemma, since z_0^p is obtained by encrypting with a different key. The same conclusion is reached for the collisions $z_0 = y_0^p$. Hence,

$$Pr[\overline{C} \mid D \text{ and } y_{n_i+2} \oplus z_{n-n_i} \notin R^e] \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{2(n+1)^2}{2^{l+1}} = \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l}.$$

Now, we compute an upper bound for $Pr[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D]$. For the extension forgery, we have $z_{n-n_i} = z_0^i = f'(r_0^i)$ by the definition of the decryption of the BIGE\$ scheme. (The argument is similar to the one used in case (a).) Hence, we use the same argument as in case (a) for the computing an upper bound for the probability when $f \overset{\mathcal{R}}{} G_S$. We use the advantage of an adversary in making the distinction between $f' \overset{\mathcal{R}}{} P^l$ and $f' \overset{\mathcal{R}}{} R^{l,l}$ in computing z_0^i

$$\begin{aligned} Pr[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] &= Pr_{f \overset{\mathcal{R}}{\leftarrow} G_S}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ &= Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f', f'' \overset{\mathcal{R}}{\leftarrow} P^l}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ &= Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f'' \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ &\Leftrightarrow Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f'' \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ &+ Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f'' \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ &\quad Adv_{\mathcal{D}}(P^l, R^{l,l}) + Pr_{\overline{f},\overline{f'},\overline{f''} \overset{\mathcal{R}}{\leftarrow} G_S, f, f'' \overset{\mathcal{R}}{\leftarrow} P^l, f' \overset{\mathcal{R}}{\leftarrow} R^{l,l}}[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D]. \end{aligned}$$

In a manner similar to case (a),

$$Adv_{\mathcal{D}}(P^l, R^{l,l}) \quad \frac{q_e(2q_e \Leftrightarrow 1)}{2^l}.$$

Now, we compute an upper bound for the second term $Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f, f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f' \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[.]$, which we denote by $Pr'[.]$; i.e., we compute an upper bound for $Pr'[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D]$. Since $z_0^i = f'(r_0^i), f' \stackrel{\mathcal{R}}{\quad} R^{l,l}$ is computed with a different key, it follows that z_0^i is random and uniformly distributed, and since it does not appear in R^e, then z_0^i is independent of any terms in R^e. Hence, since y_{n_i+2} is a constant, it follows that $y_{n_i+2} \oplus z_{n-n_i} = y_{n_i+2} \oplus z_0^i$ is random uniformly distributed, and independent of any element of R^e. Hence

$$Pr'[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \quad \frac{|R^e|}{2^l} \quad \frac{\mu_e}{l2^l}.$$

Hence,

$$Pr[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \quad \frac{q_e(2q_e \Leftrightarrow 1)}{2^l} + \frac{\mu_e}{l2^l},$$

and, by standard conditioning,

$$\begin{aligned} Pr[\overline{C} \mid D] \quad & Pr[\overline{C} \mid D \text{ and } y_{n_i+2} \oplus z_{n-n_i} \notin R^e] + Pr[y_{n_i+2} \oplus z_{n-n_i} \in R^e \mid D] \\ & \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{q_e(2q_e \Leftrightarrow 1)}{2^l} + \frac{\mu_e}{l2^l}. \end{aligned}$$

(c1) Let j be the index of the first block where $y_j \neq y_j^i, 1 \quad j \quad \min\{n+1, n_i+1\}$. By standard conditioning,

$$Pr[\overline{C} \mid D] \quad Pr[\overline{C} \mid D \text{ and } y_j \oplus z_{n-j+2} \notin R^e] + Pr[y_j \oplus z_{n-j+2} \in R^e \mid D].$$

In a similar manner to the proof for the forgeries of type (b) (using the Corollary to the Main IGE Lemma), we have

$$Pr[\overline{C} \mid D \text{ and } y_j \oplus z_{n-j+2} \notin R^e] \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l}.$$

Now, we find an upper bound for collisions between $y_j \oplus z_{n-j+2}$ and $y_k^p \oplus z_{n_p-k+2}^p, 1 \quad p \quad q_e, 1 \quad k \quad n_p+1$. Let D_j the event defining these collisions. Formally,

$$D_j : y_j \oplus z_{n-j+2} \in R^e.$$

Since j is the first index such that $y_j \neq y_j^i$, it follows that $z_{n-j+2} = z_{n_i-j+2}^i = f(x_{n_i-j+2}^i \oplus z_{n_i-j+1}^i) \oplus x_{n_i-j+1}^i$, i.e., they are the image through $f \stackrel{\mathcal{R}}{\quad} P^l$. Hence, as in case (b) an adversary can distinguish between $f \stackrel{\mathcal{R}}{\quad} P^l$ and $f \stackrel{\mathcal{R}}{\quad} R^{l,l}$ and

$$\begin{aligned} Pr[D_j \mid D] &= Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l}[D_j \mid D] \\ &= Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l}[D_j \mid D] \Leftrightarrow Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[D_j \mid D] \\ &+ Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[D_j \mid D] \\ & \quad Adv_{\mathcal{D}}(P^l, R^{l,l}) + Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[D_j \mid D] \end{aligned}$$

where the advantage of the distinguisher takes into account that f sees $\frac{e}{l}$ blocks, i.e.,

$$Adv_{\mathcal{D}}(P^l, R^{l,l}) \quad \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}}.$$

Hence, we compute an upper bound for $Pr'[D_j|D] \stackrel{def}{=} Pr_{\overline{f},\overline{f'},\overline{f''} \stackrel{\mathcal{R}}{\leftarrow} G, f', f'' \stackrel{\mathcal{R}}{\leftarrow} P^l, f \stackrel{\mathcal{R}}{\leftarrow} R^{l,l}}[D_j \mid D]$. By union bound we have

$$Pr'[D_j \mid D] \quad \sum_{p=1}^{q_e} \sum_{k=1}^{n_p+1} Pr'[y_j \oplus z_{n-j+2} = y_k^p \oplus z_{n-k+2}^p \mid D].$$

Since j is the first index such that $y_j \neq y_j^i$, it follows that $z_{n-j+2} = z_{n_i-j+2}^i$. Hence, these collisions can be expressed as $y_j \oplus z_{n_i-j+2}^i = y_k^p \oplus z_{n_p-k+2}^p$. For $i \neq p$ or $j \neq k$, since D is true, $z_{n_i-j+2}^i$ and $z_{n_p-k+2}^p$ are random, uniformly distributed, and mutually independent (since $f \;\; ^{\mathcal{R}} \;\; R^{l,l}$); hence, the collision happens with probability $1/2^l$. If $i = p, j = k$, the collision would reduce to $y_j = y_j^i$, which would be impossible by the definition of index j. Thus,

$$Pr'[D_j \mid D] \quad \frac{|R^e|}{2^l} = \frac{\mu_e}{l2^l}.$$

Hence,

$$Pr[D_j \mid D] \quad \frac{\mu_e(\mu_e \Leftrightarrow l)}{l^2 2^{l+1}} + \frac{\mu_e}{l2^l} = \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

Thus, by standard conditioning,

$$\begin{aligned} Pr[\overline{C} \mid D] \quad & \quad Pr[\overline{C} \mid D \text{ and } \overline{D_j}] + Pr[D_j \mid D] \\ & \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}. \end{aligned}$$

(c2) If $y_0 \neq y_0^p, \forall p, 1 \quad p \quad q_e$, then z_{n+1} is random, uniformly distributed, and independent of any z_k^p since it is encrypted with a different key. The same argument as in case (c1) is applied to $y_1 \oplus z_{n+1}$. Hence,

$$Pr[\overline{C} \mid D] \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

Thus, for any forgery type,

$$Pr[\overline{C} \mid D] \quad \frac{2(n+1)\mu_e}{l2^l} + \frac{(n+1)^2}{2^l} + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}}.$$

□

Proof of Lemma 7

This proof is similar to the Proof of Lemma 6. Let $Pr[.] = Pr_{f \stackrel{\mathcal{R}}{\leftarrow} G_S}[.]$. Let y be any forgery, $y \neq y^p, 1 \quad p \quad q_e$. If the events A and B that are defined in the proof of Lemma 6 are true, then the resulting plaintext is random and uniformly distributed (since $\overline{f}, \overline{f'}, \overline{f''} \;\; ^{\mathcal{R}} \;\; G_S$ and we have inputs to $\overline{f}$ that have not been seen before). Thus, the condition $x_{n+1} = 0$ happens with probability $1/2^l$. Hence, by standard conditioning,

$$\begin{aligned} Pr[x_{n+1} = 0] \quad & \quad Pr[x_{n+1} = 0 \mid A \text{ and } B] + Pr[(\overline{A} \text{ or } \overline{B})] \\ & \quad \frac{1}{2^l} + \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l2^l} + \frac{3(n+1)^2}{2^{l+1}}. \end{aligned}$$

Hence, when the scheme is implemented with the SPRP family F, we have by Fact 0,

$$Pr_{f \stackrel{\mathcal{R}}{\leftarrow} F}[x_{n+1} = 0] \quad \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l2^l} + \frac{3(n+1)^2}{2^{l+1}} + \frac{2(n+1)(2n+1)}{2^{l+1}} + \epsilon.$$

Finally, the integrity condition passes with probability

$$Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[x_{n+1} \neq 0] = 1 \Leftrightarrow Pr_{f \overset{\mathcal{R}}{\leftarrow} F}[x_{n+1} = 0]$$

$$1 \Leftrightarrow \quad \delta_P + \frac{\mu_e(\mu_e + l)}{l^2 2^{l+1}} + \frac{3(n+1)\mu_e}{l 2^l} + \frac{3(n+1)^2}{2^{l+1}} + \frac{2(n+1)(2n+1)}{2^{l+1}} + \epsilon \Bigg) ,$$

i.e., this probability is not negligible, and hence the scheme is not EF-CPA secure.

Since any forgery that passes the integrity check of scheme BIGE$-nzg includes at least a random block with non-negligible probability, the scheme BIGE$, which is not EF-CPA secure, cannot be KPF-CPA and PI-CPA secure. □

Proof of Fact 1

It is clear that if all inputs to $f = F_K$ are distinct, then the ciphertext blocks obtained at encryption are random, uniformly distributed, and mutually independent. Let $y_k^p = f(x_k^p \oplus y^{p_{k-1}}) \oplus y_{k-1}^p$ with all distinct inputs to f. It follows that $f(x_k^p \oplus y^{p_{k-1}})$ is random, uniformly distributed, and independent of anything else, and hence y_k^p is random, uniformly distributed, and independent of anything else.

To bound the probability of the event defining collisions in the input to f, namely $\overline{D}$, we use the same proof idea used by Bellare *et al.* [2] in their proof of the Main CBC Lemma. The only difference is that, in this case, the collisions include only the given plaintext strings and there is no notion of left or right plaintext strings. Hence, following the proof of the Main CBC Lemma, the size of the prohibited set in this case is half of the size obtained by Bellare *et al.*; viz., their Claim 4 [2].

Up to now, we have considered $f = F_K$ a random function. When f is a random permutation, the bound changes by adding the term $\frac{1}{2^{l+1}} \frac{e}{l} \left(\frac{e}{l} \Leftrightarrow 1 \right)$. □