Separations among formulations of non-malleable encryption under valid ciphertext condition*

Yodai Watanabe

University of Aizu, Aizuwakamatsu, Fukushima 9658580, Japan
yodai@u-aizu.ac.jp

Abstract. Non-malleability is one of the basic security goals for encryption schemes which ensures the resistance of the scheme against ciphertext modifications in the sense that any adversary, given a ciphertext of a plaintext, cannot generate another ciphertext whose underlying plaintext is meaningfully related to the initial one. There are multiple formulations of non-malleable encryption schemes, depending on whether they are based on simulation or comparison, or whether they impose valid ciphertext condition, in which an adversary is required to generate only valid ciphertexts, or not. In addition to the simulation-based and comparison-based formulations (SNM and CNM), an indistinguishability-based characterization of non-malleability (IND), called ciphertext indistinguishability against parallel chosen-ciphertext attacks has been proposed. These three formulations, SNM, CNM and IND, have been shown equivalent if the valid ciphertext condition is not imposed; however, if that condition is imposed, then they have been shown equivalent only against the strongest type of attack models, and the relations among them against the weaker types of the attack models remain open. This work answers this open question by showing the separations $\mathrm{SNM}^* \not\Longrightarrow \mathrm{CNM}^*$ and $\mathrm{IND}^* \not\Longrightarrow \mathrm{SNM}^*$ against the weaker types of the attack models, where the asterisk attached to the short-hand notations represents that the valid ciphertext condition is imposed. Moreover, motivated by the proof of the latter separation, this paper introduces simulation-based and comparison-based formulations of semantic security (SSS^* and CSS^*) against parallel chosen-ciphertext attacks, and shows the equivalences $\mathrm{SSS}^* \Longleftrightarrow \mathrm{SNM}^*$ and $\mathrm{CSS}^* \Longleftrightarrow \mathrm{CNM}^*$ against all types of the attack models. It thus follows that $\mathrm{IND}^* \not\Longrightarrow \mathrm{SSS}^*$, that is, semantic security and ciphertext indistinguishability, which have been shown equivalent in various settings, separate against the weaker parallel chosen-ciphertext attacks under the valid ciphertext condition.

Keywords: Public key encryption · Non-malleability · Relation among security notions

* This work was supported in part by JSPS Grants-in-Aid for Scientific Research (C) No. 19K11831.

1 Introduction

The security of public key encryption schemes is commonly specified by the security goal and the attack model. Here, the security goals formulate what type of security of the scheme is intended to be protected, and the attack models formulate what type of external resources is assumed to be available to an adversary attacking the scheme. Non-malleability [8] is one of the basic security goals which ensures the resilience of the scheme against ciphertext modifications in the sense that any adversary, given a ciphertext of a plaintext called a challenge ciphertext, cannot generate another ciphertext whose underlying plaintext is meaningfully related to the initial one.

There are multiple formulations of non-malleable encryption schemes, depending on whether they are based on simulation or comparison, or whether they impose valid ciphertext condition, in which an adversary is required to generate only valid ciphertexts, or not.[1] Here, in the simulation-based formulation [8], the probability of successful ciphertext modifications by an adversary (with a challenge ciphertext) is measured against that by a simulator without a challenge ciphertext, while in the comparison-based formulation [2], the former probability is measured against that by "random guess" which corresponds to coincidence that a plaintext independently sampled according to the message distribution specified by the adversary is meaningfully related to the initial plaintext. Since a simulator is more powerful than "random guess", it can be shown that the comparison-based formulation is stronger than the simulation-based formulation regardless of whether the valid ciphertext condition is imposed or not [3]. The original work [8] employed the simulation-based formulation with the valid ciphertext condition, which would be the most natural one at least from our intuition for non-malleability.[2] Later, the comparison-based formulation of non-malleability was proposed [2] and then shown to be equivalent to the simulation-based one if the valid ciphertext condition is not imposed [3]. On the other hand, if that condition is imposed, then these two notions were known to equivalent only against the strongest attack model, namely adaptive chosen-ciphertext attack (CCA2) [18], and the relation between them against the weaker attack models, namely chosen plaintext attack (CPA) and non-adaptive chosen ciphertext attack (CCA1) [17], still remains open.[3] Moreover, an indistinguishability-based characterization of non-malleability, called ciphertext indistinguishability against parallel chosen-ciphertext attacks (PCA0, PCA1 and PCA2), was also

[1] In addition to encryption schemes, non-malleability has been formulated for various primitives (see e.g. [4, 7–9, 13, 19]).

[2] In fact, Katz and Yung [15] imposed the valid ciphertext condition to formulate non-malleability for private-key encryption based on the consideration that "the current definition more closely corresponds to our intuitive notion." It should, however, be mentioned that they employed the comparison-based formulation. Here, one advantage of the comparison-based formulation for private key encryption schemes would be that it frees us from considering the encryption oracle for a simulator.

[3] In the full version of [3], clarifying this relation was mentioned as the last open question. This work was inspired by this question.

introduced and shown to be equivalent to non-malleability if the valid ciphertext condition is not imposed [3]. On the other hand, if that condition is imposed, then the above two notions were known to be equivalent only against the strongest type of the attack models (PCA2 and CCA2), and the relation between them against the weaker types of the attack models, also remains open.

1.1 Contributions of this work

This work shows that (perhaps surprisingly) the simulation-based formulation of non-malleability (SNM*) is strictly weaker than the comparison-based one (CNM*) against CPA and CCA1, which answers the last open question mentioned in the full version of [3]. Moreover, this work also shows that ciphertext indistinguishability (IND*) against PCA0 and PCA1 is strictly weaker than the simulation-based non-malleability against CPA and CCA1, respectively. Here, the asterisk attached to the above short-hand notations represents that the valid ciphertext condition is imposed. The proofs of these results follow the standard procedure to show the separation $X \not\Rightarrow Y$ for computational security notions X and Y, in which (a) the existence of an X-secure encryption scheme Π is assumed and then (b) Π is modified to Π' so that Π' is still X-secure but not Y-secure; however, the modifications of encryption schemes and the estimation of adversaries' advantages given in this paper are specifically aimed at showing the separations.

In addition, motivated by the proof of the latter separation, this paper introduces simulation-based and comparison-based formulations of semantic security (SSS* and CSS*) against parallel chosen-ciphertext attacks, and shows that SSS* and CSS* are equivalent to SNM* and CNM*, respectively. This, together with the latter separation, shows that semantic security and ciphertext indistinguishability, which have been shown equivalent in various settings (see e.g. [1, 10–12, 16, 20]), separate against parallel chosen-ciphertext attacks under the valid ciphertext condition. We note that parallel chosen-ciphertext attacks were introduced not to show this separation, but to give an indistinguishability-based characterization of non-malleability, which is expected to facilitate the study on non-malleability.[4] Figure 1 shows the complete relations among formulations of non-malleable encryption.

2 Preliminaries

Let A be a probabilistic algorithm. The result of running A on inputs $x_1, x_2, \ldots$ and randomness r is denoted by $A(x_1, x_2, \ldots; r)$. The notation $y \leftarrow A(x_1, x_2, \ldots)$ denotes the experiment of choosing r at random and setting $y = A(x_1, x_2, \ldots; r)$.

[4] It may be of interest to note that parallel chosen-ciphertext attacks have been extended to self-destruct attacks (SDA) [6], yielding NM-SDA security, which has natural applications and can be achieved from IND-CPA security [5]. Here, in SDA, an adversary is allowed to make multiple (and so adaptive) parallel chosen-ciphertext queries up to the point when the first invalid ciphertext is submitted.

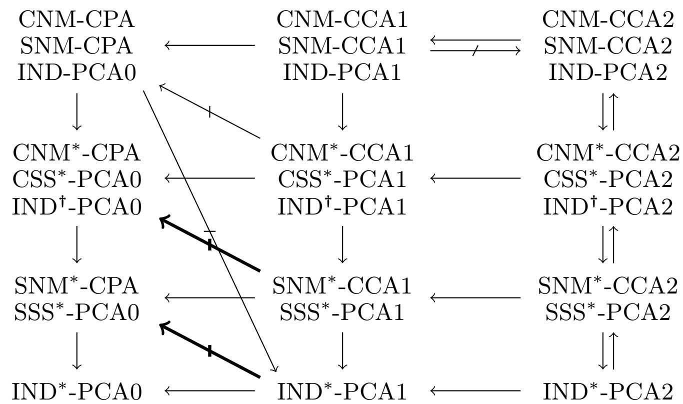

Fig. 1. Relations among formulations of non-malleability. The bold barred arrows represent the separations shown by this paper. The separation CNM*-CCA1 $\not\Rightarrow$ CNM-CPA can be shown by the result of [14], together with the idea mentioned in the full version of [14]. The other relations are consequences of [2, 3, 12].

If S is a distribution (resp. a finite set), then S in the notation $x \leftarrow S$ is considered an algorithm which returns a sample drawn according to S (resp. the uniform distribution over S). For an event E, the notation

$$\Pr[x \leftarrow A(a_1, a_2, \ldots);\ y \leftarrow B(b_1, b_2, \ldots);\ \cdots : E]$$

denotes the probability that E occurs after ordered execution of the listed experiments.

The length of a string s is denoted by $|s|$. The concatenation of strings s_1 and s_2 is denoted by $s_1 s_2$. A sequence is denoted in boldface. The length of a sequence $\boldsymbol{x}$ is denoted by $|\boldsymbol{x}|$ and its i-th component by $\boldsymbol{x}_i$, so that $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_l)$ with $l = |\boldsymbol{x}|$. The concatenation of sequences $\boldsymbol{x}$ and $\boldsymbol{y}$ is denoted by $\boldsymbol{x}||\boldsymbol{y}$. For an operation F and a sequence $\boldsymbol{x}$ of length l whose components are in the domain of F, we use the notation $F(\boldsymbol{x})$ to denote

$$F(\boldsymbol{x}) = (F(\boldsymbol{x}_1), \cdots, F(\boldsymbol{x}_l)).$$

For a sequence $\boldsymbol{x}$ of length l_1 whose components are sequences of length l_2,

$$\boldsymbol{x} = ((\boldsymbol{x}_{11}, \cdots, \boldsymbol{x}_{1l_2}), \cdots, (\boldsymbol{x}_{l_1 1}, \cdots, \boldsymbol{x}_{l_1 l_2})),$$

we define a sequence $\boldsymbol{x}_{*j}$ for $j \in [l_2]$ by

$$\boldsymbol{x}_{*j} = (\boldsymbol{x}_{1j}, \cdots, \boldsymbol{x}_{l_1 j}). \tag{1}$$

For sequences $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and $\boldsymbol{d}$ of the same length l, we introduce the notation $\langle \boldsymbol{a} : \boldsymbol{b} | \boldsymbol{c} = \boldsymbol{d} \rangle$ to denote the sequence of length l whose i-th component is given by

$$\langle \boldsymbol{a} : \boldsymbol{b} | \boldsymbol{c} = \boldsymbol{d} \rangle_i = \begin{cases} \boldsymbol{a}_i & \text{if } \boldsymbol{c}_i = \boldsymbol{d}_i, \\ \boldsymbol{b}_i & \text{otherwise,} \end{cases} \tag{2}$$

with $i \in [l]$. In this notation, a symbol x not in boldface is considered as the sequence $(x)^l$ of length l whose components are all x; e.g.,

$$\langle a : \boldsymbol{b} | \boldsymbol{c} = d \rangle = \langle (a)^l : \boldsymbol{b} | \boldsymbol{c} = (d)^l \rangle.$$

A function ϵ from $\mathbb{N}$ to $\mathbb{R}$, $\epsilon : \mathbb{N} \to \mathbb{R}$, is called *negligible* if for all $c > 0$, there exists an integer n_c such that $\epsilon(n) \leq n^{-c}$ for all $n \geq n_c$.

A public key encryption scheme is a triple of algorithms, $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, such that

- $\mathcal{K}$, the *key generation algorithm*, is a probabilistic, polynomial-time algorithm which takes as input a security parameter $k \in \mathbb{N}$ (in unary) and returns a pair (pk, sk) of matching public and secret keys,
- $\mathcal{E}$, the *encryption algorithm*, is a probabilistic, polynomial-time algorithm which takes as input a public key pk and a plaintext $x \in \{0,1\}^*$ and returns a ciphertext y,
- $\mathcal{D}$, the *decryption algorithm*, is a deterministic, polynomial-time algorithm which takes as input a secret key sk and a ciphertext y and returns either a plaintext $x \in \{0,1\}^*$ or a special symbol $\perp$ to indicate that the ciphertext is invalid,

where the correctness condition $\Pr[\mathcal{D}_{sk}(\mathcal{E}_{pk}(x)) = x] = 1$ has to hold for all $k \in \mathbb{N}$, for all (pk, sk) which can be output by $\mathcal{K}(1^k)$ and for all $x \in \{0,1\}^*$. In this paper, we assume that all algorithms have access to the key generation algorithm $\mathcal{K}(1^k)$ given the security parameter k.[5]

2.1 Formulations of non-malleability

The simulation-based and comparison-based formulations of non-malleability can be described in the common framework [2]. In the simulator-based formulation introduced in [8] and refined in [3], an adversary A, its simulator S and a relation R are considered. Here, A is a pair of algorithms, $A = (A_1, A_2)$, corresponding to two stages of an attack. Before the execution of A, the key generation algorithm on the security parameter k generates a pair of matching public and secret keys, (pk, sk). At the first stage of the attack, A_1 takes as input the public key pk and outputs a distribution M over messages (plaintexts) such that all plaintexts in the support of M are of the same length, together with state information s_1 for A_2 and side information s_2 for R. Next, a plaintext x

[5] This is necessary in some proofs where a simulator (which is not explicitly given the security parameter k in our definition) runs $\mathcal{K}(1^k)$, and has also been assumed e.g. in [3].

is sampled according to M and then encrypted to give a challenge ciphertext y. At the second stage of the attack, A_2 takes as input the challenge ciphertext y and the state information s_1 and outputs a sequence $\boldsymbol{y}$ of ciphertexts such that $y \notin \boldsymbol{y}$. Finally, $\boldsymbol{y}$ is decrypted to give $\boldsymbol{x}$. Here, A is supposed to have access to the decryption oracle $\mathcal{D}_{sk}(\cdot)$ depending on the attack model ATK; namely, A has no access to $\mathcal{D}_{sk}(\cdot)$ for chosen plaintext attack (CPA), only A_1 has access to $\mathcal{D}_{sk}(\cdot)$ for non-adaptive chosen plaintext attack (CCA1) and both A_1 and A_2 have access to $\mathcal{D}_{sk}(\cdot)$ for non-adaptive chosen plaintext attack (CCA1), where A_2 is prohibited from asking the challenge ciphertext y to $\mathcal{D}_{sk}(\cdot)$ for the last case. The experiment for a simulator $S = (S_1, S_2)$ is the same as that for A except for that S is not given a challenge ciphertext y and S has no access to $\mathcal{D}_{sk}(\cdot)$ regardless of the attack models. Then, A and S are considered successful if $R(x, \boldsymbol{x}, M, s_2) = 1$ and $\perp \notin \boldsymbol{x}$ hold, and an encryption scheme is called secure in the sense of SNM*-ATK if for all probabilistic, polynomial-time adversary A which outputs M samplable in polynomial-time and for all relation R computable in polynomial-time, there exists a probabilistic, polynomial-time simulator S which outputs M samplable in polynomial-time, such that the advantage in success probability of A against S is negligible as a function of k. A formal definition of SNM*-ATK is described below.

Definition 1 (SNM*-ATK [3, 8]). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and R be a relation. Let $A = (A_1, A_2)$ be an adversary attacking Π and $S = (S_1, S_2)$ be its simulator. For $k \in \mathbb{N}$ and $\mathrm{ATK} \in \{\mathrm{CPA}, \mathrm{CCA1}, \mathrm{CCA2}\}$, consider the following two experiments:*

Experiment $\mathsf{Expt}^{\mathrm{SNM}^*\text{-ATK-1}}_{\Pi,R,A}(k)$	Experiment $\mathsf{Expt}^{\mathrm{SNM}^*\text{-ATK-0}}_{\Pi,R,S}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$	$(pk, sk) \leftarrow \mathcal{K}(1^k)$
$(M, s_1, s_2) \leftarrow A_1^{\mathcal{O}_1}(pk)$	$(M, s_1, s_2) \leftarrow S_1(pk)$
$x \leftarrow M$; $y \leftarrow \mathcal{E}_{pk}(x_0)$	$x \leftarrow M$
$\boldsymbol{y} \leftarrow A_2^{\mathcal{O}_2}(s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$	$\boldsymbol{y} \leftarrow S_2(s_1)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
if $R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$	if $R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$	else $w \leftarrow 0$

Here, $\mathcal{O}_1$ and $\mathcal{O}_2$ are oracles given by

$$\begin{array}{llll} \mathcal{O}_1 = \varepsilon(\cdot) & \text{and} & \mathcal{O}_2 = \varepsilon(\cdot) & \textit{if } \mathrm{ATK} = \mathrm{CPA}, \\ \mathcal{O}_1 = \mathcal{D}_{sk}(\cdot) & \text{and} & \mathcal{O}_2 = \varepsilon(\cdot) & \textit{if } \mathrm{ATK} = \mathrm{CCA1}, \\ \mathcal{O}_1 = \mathcal{D}_{sk}(\cdot) & \text{and} & \mathcal{O}_2 = \mathcal{D}_{sk}(\cdot) & \textit{if } \mathrm{ATK} = \mathrm{CCA2}, \end{array}$$

respectively, where $\varepsilon(\cdot)$ denotes the empty function which, on any input, outputs the empty string ε, and it is supposed that (i) all strings in the support of M are of the same length, (ii) $y \notin \boldsymbol{y}$ and (iii) $y \notin \boldsymbol{query}(A; \mathcal{O}_2)$ in the above experiment $\mathsf{Expt}^{SNM^*\text{-}ATK\text{-}1}_{\Pi,R,A}(k)$, where $\boldsymbol{query}(A; \mathcal{O}_2)$ denotes a sequence of queries from A to $\mathcal{O}_2$ for the case of $\mathrm{ATK} = \mathrm{CCA2}$. An adversary A is called legitimate *if its outputs and queries satisfy the above conditions (i)–(iii). For a function f of k, an adversary A (resp. a simulator S) is called* bounded by time $f(k)$ *if A (resp. S) runs in time $f(k)$ and outputs M samplable in time $f(k)$. Then, an encryption* *scheme Π is called* secure in the sense of SNM*-ATK *if for all polynomial p, all probabilistic adversary A bounded by time $p(k)$ and all relation R computable in time $p(k)$, there exist a polynomial $p'(k)$ and a simulator S bounded by time $p'(k)$ such that $\mathsf{Adv}^{SNM^*\text{-}ATK}_{\Pi,R,A,S}(k)$ is negligible, where $\mathsf{Adv}^{SNM^*\text{-}ATK}_{\Pi,R,A,S}$ denotes the advantage of A against S defined by*

$$\begin{aligned}&\mathsf{Adv}^{SNM^*\text{-}ATK}_{\Pi,R,A,S}(k)\\&=\Pr[\mathsf{Expt}^{SNM^*\text{-}ATK\text{-}1}_{\Pi,R,A}(k):w=1]-\Pr[\mathsf{Expt}^{SNM^*\text{-}ATK\text{-}0}_{\Pi,R,S}(k):w=1].\end{aligned}$$

In the comparison-based formulation introduced in [2], only an adversary A is considered. Again, A is a pair of algorithms, $A=(A_1,A_2)$, corresponding to two stages of an attack. Before the execution of A, the key generation algorithm on the security parameter k generates a pair of matching public and secret keys, (pk,sk). At the first stage of the attack, A_1 takes as input the public key pk and outputs a distribution M over messages (plaintexts) such that all plaintexts in the support of M are of the same length, together with state information s for A_2. Next, a plaintext x is sampled according to M and then encrypted to give a challenge ciphertext y. At the second stage of the attack, A_2 takes as input the challenge ciphertext y and the state information s and outputs a relation R and a sequence $\boldsymbol{y}$ of ciphertexts such that $y\notin\boldsymbol{y}$. Finally, $\boldsymbol{y}$ is decrypted to give $\boldsymbol{x}$. Here, A is supposed to have access to the decryption oracle $\mathcal{D}_{sk}(\cdot)$ depending on the attack model, as in the simulation-based formulation. Then, A is considered successful if $R(x,\boldsymbol{x})=1$ and $\perp\notin\boldsymbol{x}$ hold. In contrast to the simulation-based formulation, the success probability of A is compared with that of "random guess" which corresponds to coincidence that a plaintext x' independently sampled according to M satisfies $R(x',\boldsymbol{x})=1$ (with $\perp\notin\boldsymbol{x}$). Thus, an encryption scheme is called secure in the sense of CNM*-ATK if for all probabilistic, polynomial-time adversary A which outputs M samplable in polynomial-time and R computable in polynomial-time, the advantage in success probability of A against "random guess" is negligible as a function of k. A formal definition of CNM*-ATK is described below.

Definition 2 (CNM*-ATK [2]). *Let $\Pi=(\mathcal{K},\mathcal{E},\mathcal{D})$ be an encryption scheme. Let $A=(A_1,A_2)$ be an adversary attacking Π. For $b\in\{0,1\}$, $k\in\mathbb{N}$ and* ATK $\in\{\mathrm{CPA},\mathrm{CCA1},\mathrm{CCA2}\}$*, consider the following experiment:*

Experiment $\mathsf{Expt}^{\mathrm{CNM^*\text{-}ATK\text{-}}b}_{\Pi,A}(k)$
  $(pk,sk)\leftarrow\mathcal{K}(1^k)$; $(M,s)\leftarrow A_1^{\mathcal{O}_1}(pk)$; $x_0,x_1\leftarrow M$
  $y\leftarrow\mathcal{E}_{pk}(x_1)$; $(R,\boldsymbol{y})\leftarrow A_2^{\mathcal{O}_2}(s,y)$; $\boldsymbol{x}\leftarrow\mathcal{D}_{sk}(\boldsymbol{y})$
  if $R(x_b,\boldsymbol{x})=1\wedge\perp\notin\boldsymbol{x}$ then $w\leftarrow 1$
  else $w\leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are defined as in Definition 1. For a function f of k, an adversary A is called bounded by time $f(k)$ *if A runs in time $f(k)$ and outputs M samplable in time $f(k)$ and R computable in time $f(k)$. Then, an encryption scheme Π is called* secure in the sense of CNM*-ATK *if for all polynomial p and all probabilistic adversary*

A bounded by time $p(k)$, $\mathsf{Adv}_{\Pi,A}^{CNM^*\text{-}ATK}(k)$ is negligible, where $\mathsf{Adv}_{\Pi,A}^{CNM^*\text{-}ATK}$ denotes the advantage of A defined by

$$\begin{aligned}&\mathsf{Adv}_{\Pi,A}^{CNM^*\text{-}ATK}(k)\\&= \Pr[\mathsf{Expt}_{\Pi,A}^{CNM^*\text{-}ATK\text{-}1}(k) : w = 1] - \Pr[\mathsf{Expt}_{\Pi,A}^{CNM^*\text{-}ATK\text{-}0}(k) : w = 1].\end{aligned}$$

In the indistinguishability-based characterization of non-malleability introduced in [3], only an adversary A is considered. In this case, A is a triple of algorithms, $A = (A_1, A_2, A_3)$, corresponding to three stages of an attack. Before the execution of A, the key generation algorithm on the security parameter k generates a pair of matching public and secret keys, (pk, sk). At the first stage of the attack, A_1 takes as input the public key pk and outputs two plaintexts x_0 and x_1 such that $|x_0| = |x_1|$, together with state information s_1 for A_2. Next, one of the two plaintexts x_0 and x_1, say x_b ($b \in \{0, 1\}$), is chosen at random and then encrypted to give a challenge ciphertext y. At the second stage of the attack, A_2 takes as input the challenge ciphertext y and the state information s_1 and outputs a sequence $\boldsymbol{y}$ of ciphertexts such that $y \notin \boldsymbol{y}$, together with state information s_2 for A_3. Then, $\boldsymbol{y}$ is decrypted to give $\boldsymbol{x}$. At the third stage of the attack, A_3 takes as input the sequence $\boldsymbol{x}$ and the state information s_2 and outputs $d \in \{0, 1\}$. Here, A is supposed to have access to the decryption oracle $\mathcal{D}_{sk}(\cdot)$ depending on the attack model PCAX; namely, A has no access to $\mathcal{D}_{sk}(\cdot)$ for PCA0, only A_1 has access to $\mathcal{D}_{sk}(\cdot)$ for PCA1 and all A_1, A_2 and A_3 have access to $\mathcal{D}_{sk}(\cdot)$ for PCA2, where A_2 and A_3 are prohibited from asking the challenge ciphertext y to $\mathcal{D}_{sk}(\cdot)$ for the last case. Then, A is considered successful if $d = b$ and $\perp \notin \boldsymbol{x}$ hold, and an encryption scheme is called secure in the sense of IND*-PCAX if for all probabilistic, polynomial-time adversary A, the advantage in success probability of A against probability $\frac{1}{2}$ is negligible as a function of k. A formal definition of IND*-PCAX is described below.

Definition 3 (IND*-PCAX [3]). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $A = (A_1, A_2, A_3)$ be an adversary attacking Π. For $k \in \mathbb{N}$ and* PCAX $\in$ $\{\mathrm{PCA0}, \mathrm{PCA1}, \mathrm{PCA2}\}$*, consider the following experiment:*

Experiment $\mathsf{Expt}_{\Pi,A}^{\mathrm{IND}^*\text{-}\mathrm{PCAX}}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(x_0, x_1, s_1) \leftarrow A_1^{\mathcal{O}_1}(pk)$; $b \leftarrow \{0, 1\}_U$; $y \leftarrow \mathcal{E}_{pk}(x_b)$
$(\boldsymbol{y}, s_2) \leftarrow A_2^{\mathcal{O}_2}(x_0, x_1, s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $d \leftarrow A_3^{\mathcal{O}_2}(\boldsymbol{x}, s_2)$
if $d = b \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are oracles given by

$$\begin{array}{llll}\mathcal{O}_1 = \varepsilon(\cdot) & \text{and} & \mathcal{O}_2 = \varepsilon(\cdot) & \textit{if } \mathrm{PCAX} = \mathrm{PCA0},\\ \mathcal{O}_1 = \mathcal{D}_{sk}(\cdot) & \text{and} & \mathcal{O}_2 = \varepsilon(\cdot) & \textit{if } \mathrm{PCAX} = \mathrm{PCA1},\\ \mathcal{O}_1 = \mathcal{D}_{sk}(\cdot) & \text{and} & \mathcal{O}_2 = \mathcal{D}_{sk}(\cdot) & \textit{if } \mathrm{PCAX} = \mathrm{PCA2},\end{array}$$

respectively, where $\varepsilon(\cdot)$ denotes the empty function as before. Then, an encryption scheme Π is called secure in the sense of IND*-PCAX *if for all polynomial p*

and all probabilistic adversary A runnable in time $p(k)$, $\mathsf{Adv}_{\Pi,A}^{IND^*\text{-}PCAX}(k)$ is negligible, where $\mathsf{Adv}_{\Pi,A}^{IND^*\text{-}PCAX}$ denotes the advantage of A defined by

$$\mathsf{Adv}_{\Pi,A}^{IND^*\text{-}PCAX}(k) = 2\Pr[\mathsf{Expt}_{\Pi,A}^{IND^*\text{-}PCAX}(k) : w = 1] - 1.$$

The above definitions of SNM*-ATK, CNM*-ATK and IND*-PCAX can be modified by removing the valid ciphertext condition $\perp \notin \boldsymbol{x}$, giving stronger notions, denoted as SNM-ATK, CNM-ATK and IND-PCAX, respectively. It has been shown that these three security notions are equivalent [3]. Here, it should be stated that whether imposing the valid ciphertext condition is more appropriate or not depends on applications, as mentioned in [3, 15].

3 Separation between simulation-based and comparison-based formulations

Let X and Y be security notions for encryption schemes. In order to show the separation $X \not\Rightarrow Y$, it is necessary to show that there exists an encryption scheme which is secure in the sense of X but not secure in the sense of Y. Since the existence of computationally secure encryption schemes has not been proved, it is standard to show the separation by modifying an encryption scheme Π to another encryption scheme Π' so that if Π is X-secure, then Π' is still X-secure but not Y-secure.[6] The proofs in this paper follow this standard.

To prove the separation between SNM* and CNM*, we modify an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ to $\Pi' = (\mathcal{K}', \mathcal{E}', \mathcal{D}')$ so that the modified decryption algorithm $\mathcal{D}'$ has an additional "option" which gives no (absolute) advantage to an adversary and a simulator. More precisely, $\mathcal{E}'$ takes a plaintext x and outputs a ciphertext $(0, \mathcal{E}_{pk}(x))$, and $\mathcal{D}'$ takes a ciphertext (a, y) and outputs $\mathcal{D}_{sk}(y)$ if $a = 0$ or x equals a specific string (say, 0), otherwise $\perp$. It can be seen from this definition of $\mathcal{D}'$ that there is no advantage to choose the option $a \neq 0$. An SNM* simulator may not choose this option, while a CNM* adversary can force the "random guess" to choose the option so as to take relative advantage against it.

From a technical point of view, it is convenient to slightly modify SNM* to SNM$^{\circledast}$ so that the output of the side information s_2 for the relation R is delayed from the first stage to the second one. This modification does not change the strength of security, and below we consider SNM$^{\circledast}$ instead of SNM*, which could make the intuition behind the proof clearer. A formal definition of SNM$^{\circledast}$ and the proof of its equivalence to SNM* (which may seem more complicated than expected) are given in Appendix A. We are now ready to show the separation between SNM$^{\circledast}$ and CNM*.

Theorem 1. *SNM$^{\circledast}$-CCA1 $\not\Rightarrow$ CNM*-CPA.*

[6] Since the existence of computationally secure private key encryption schemes is equivalent to that of one-way functions, we may show separations for private key encryption schemes by assuming the latter (see e.g. [15]).

Proof. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme. By using Π, let us construct another encryption scheme $\Pi' = (\mathcal{K}', \mathcal{E}', \mathcal{D}')$ as

Algorithm $\mathcal{K}'(1^k)$	Algorithm $\mathcal{E}'_{pk}(x)$	Algorithm $\mathcal{D}'_{sk}((a,y))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$ return (pk, sk)	$y \leftarrow \mathcal{E}_{pk}(x)$ return $(0, y)$	$x \leftarrow \mathcal{D}_{sk}(y)$ if $a = 0$ then return x else if $x = 0$ then return x else return $\perp$

It can be seen from this definition that for a and x such that $x \leftarrow \mathcal{D}'_{sk}((a, y))$ for some y,

$$x \neq \perp \wedge a \neq 0 \iff x \neq \perp \wedge a \neq 0 \wedge x = 0. \tag{3}$$

Then, the theorem follows from the following two lemmas. □

The following two lemmas claim that (a) If Π is SNM$^{⊛}$-CCA1, then so is Π' and (b) Π' is not CNM*-CPA, respectively. To show the first lemma (a), we construct an adversary A attacking Π and a relation R for A by using an adversary A' attacking Π' and a relation R' for A', respectively. The construction of A from A' is straightforward except for the case where A'_2 outputs a sequence $\boldsymbol{y}'$ of ciphertexts which contains a component (a', y) such that $a' \neq 0$ and y is a challenge ciphertext for A (we note that challenge ciphertexts for A' have a form $(0, y)$); in fact, A_2 can generate a sequence $\boldsymbol{y}$ of ciphertexts by simply ignoring the first component a' of each component (a', y') of $\boldsymbol{y}'$. On the other hand, in the exceptional case, the sequence $\boldsymbol{y}$ generated as above contains the challenge ciphertext y, which violates the legitimate condition (ii) (see Definition 1 for this condition). This violation can be avoided as follows. If $y \in \boldsymbol{y}$, then A_2 replaces y in $\boldsymbol{y}$ by e such that $e \neq y$, and then includes the position of this replacement in the side information s_2 for the relation R.[7] Then, the relation R can replace $\mathcal{D}_{sk}(e)$ at the position of the above replacement by the specific string 0. Here, note that (3) implies that a ciphertext (a, y) with $a \neq 0$ is valid only if $x = \mathcal{D}_{sk}(y) = 0$. It thus follows that the advantage of A is no less than that of A'. Moreover, we can show the second lemma (b) by considering a CNM*-CPA adversary which simply transforms a challenge ciphertext $(0, y)$ to $(1, y)$ and outputs it as a component of $\boldsymbol{y}$. Detailed proofs of these lemmas are described below.

Lemma 1. *If Π is SNM$^{⊛}$-CCA1, then so is Π'.*

Proof. Let p be a polynomial of k. Let R' be a relation computable in time $p(k)$ and $A' = (A'_1, A'_2)$ be a legitimate SNM$^{⊛}$-CCA1 adversary attacking Π', bounded by time $p(k)$.[8] By using A' and R', let us construct an SNM$^{⊛}$-CCA1 adversary $A = (A_1, A_2)$ attacking Π and a relation R as

[7] Since SNM$^{⊛}$ allows an adversary to output s_2 at the second stage, SNM$^{⊛}$ is convenient for simplifying the proof based on this construction.

[8] See Definition 1 for an adversary *bounded by time* $p(k)$.

Algorithm $A_1^{\mathcal{D}_{sk}}(pk)$	Algorithm $A_2((s_1,e),y)$	Relation $R(x,\boldsymbol{x},M,\boldsymbol{s})$
$(M,s_1) \leftarrow A_1'^{\mathcal{D}'_{sk}}(pk)$ $x' \leftarrow M$ $L \leftarrow \lvert x'\rvert+1$ $e \leftarrow \mathcal{E}_{pk}(0^L)$ return $(M,(s_1,e))$	$(\boldsymbol{y}',s_2) \leftarrow A_2'(s_1,(0,y))$ $\boldsymbol{a}' \leftarrow \boldsymbol{y}'_{*1}$ $\boldsymbol{y} \leftarrow \langle \boldsymbol{y}'_{*2} : e\lvert \boldsymbol{a}'=0\rangle$ $\boldsymbol{s} \leftarrow (s_2)\lvert\rvert\boldsymbol{a}'$ return $(\boldsymbol{y},\boldsymbol{s})$	if $\lvert\boldsymbol{s}\rvert=0$ then return 0 parse $\boldsymbol{s}$ as $(s_2)\lvert\rvert\boldsymbol{a}'$ with $\lvert(s_2)\rvert=1$ $\tilde{\boldsymbol{x}} \leftarrow \langle \boldsymbol{x} : 0\lvert\boldsymbol{a}'=0\rangle$ return $R'(x,\tilde{\boldsymbol{x}},M,s_2)$

(see (1) and (2) for the notations $\boldsymbol{x}_{*j}$ and $\langle \boldsymbol{a} : \boldsymbol{b}|\boldsymbol{c}=\boldsymbol{d}\rangle$, respectively), where the length L is chosen so that $|0^L s_2| > |x|$ for any $s_2 \in \{0,1\}^*$ and any output x of M (note that M outputs messages of a fixed length), which ensures that $\mathcal{E}_{pk}(0^L s_2) \neq y$ with probability 1. Since A' is bounded by time $p(k)$ (and so every string output by A' has a length bounded by $p(k)$), R' is computable in time $p(k)$ and $\mathcal{E}_{pk}$ is polynomial-time, it follows that M is samplable in time $p(k)$ and A and R are polynomial-time. Moreover, A can be seen legitimate as follows: the condition (i) follows from that A' is legitimate and the condition (iii) from that A_2 has no oracle access to $\mathcal{D}_{sk}$; since $(0,y) \notin \boldsymbol{y}'$ and so $\forall i((\boldsymbol{y}'_{*2})_i = y \implies \boldsymbol{a}'_i \neq 0)$, we have $y \notin \boldsymbol{y} = \langle \boldsymbol{y}'_{*2} : e|\boldsymbol{a}'=0\rangle$, from which the condition (ii) follows. We note that A_1 can answer queries from A_1' by using her own oracle $\mathcal{D}_{sk}$ to compute $\mathcal{D}'_{sk}$.

It is now convenient to consider the experiment $\mathsf{Expt}_1(k)$ defined by

Experiment $\mathsf{Expt}_1(k)$
$(pk,sk) \leftarrow \mathcal{K}(1^k)$; $(M,s_1) \leftarrow A_1'^{\mathcal{D}'_{sk}}(pk)$; $x,x' \leftarrow M$; $y \leftarrow \mathcal{E}_{pk}(x)$
$L \leftarrow |x'|+1$; $e \leftarrow \mathcal{E}_{pk}(0^L)$; $(\boldsymbol{y}',s_2) \leftarrow A_2'(s_1,(0,y))$
$\boldsymbol{a}' \leftarrow \boldsymbol{y}'_{*1}$; $\boldsymbol{y} \leftarrow \langle \boldsymbol{y}'_{*2} : e|\boldsymbol{a}'=0\rangle$; $\boldsymbol{x}' \leftarrow \mathcal{D}'_{sk}(\boldsymbol{y}')$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $\tilde{\boldsymbol{x}} \leftarrow \langle \boldsymbol{x} : 0|\boldsymbol{a}'=0\rangle$

and to introduce, for an event E, the short-hand notation $p_1(E) = \Pr[\mathsf{Expt}_1(k) : E]$. Since $\boldsymbol{x} = \langle \boldsymbol{x}' : 0^L|\boldsymbol{a}'=0\rangle$ and $\tilde{\boldsymbol{x}} = \langle \boldsymbol{x} : 0|\boldsymbol{a}'=0\rangle$, we have

$$\perp \in \boldsymbol{x} \iff \perp \in \tilde{\boldsymbol{x}} \implies \perp \in \boldsymbol{x}'. \tag{4}$$

Moreover, since $\tilde{\boldsymbol{x}}$ can be written as $\tilde{\boldsymbol{x}} = \langle \boldsymbol{x}' : 0|\boldsymbol{a}'=0\rangle$, we have $\boldsymbol{x}' = \tilde{\boldsymbol{x}} \iff \forall i(\boldsymbol{a}'_i = 0 \vee \boldsymbol{x}'_i = 0)$. It thus follows from (3) that

$$\begin{aligned}
\perp \notin \boldsymbol{x}' &\iff \forall i(\boldsymbol{x}'_i \neq \perp \wedge (\boldsymbol{a}'_i = 0 \vee \boldsymbol{a}'_i \neq 0)) \\
&\iff \forall i(\boldsymbol{x}'_i \neq \perp \wedge (\boldsymbol{a}'_i = 0 \vee (\boldsymbol{a}'_i \neq 0 \wedge \boldsymbol{x}'_i = 0))) \\
&\iff \forall i(\boldsymbol{x}'_i \neq \perp \wedge (\boldsymbol{a}'_i = 0 \vee \boldsymbol{x}'_i = 0)) \\
&\iff \perp \notin \boldsymbol{x}' \wedge \boldsymbol{x}' = \tilde{\boldsymbol{x}}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-CCA1-1}}_{\Pi,R,A}(k) : w=1] \\
&= p_1(R(x,\boldsymbol{x},M,(s_2)||\boldsymbol{a}') = 1 \wedge \perp \notin \boldsymbol{x}) \\
&= p_1(R'(x,\tilde{\boldsymbol{x}},M,s_2) = 1 \wedge \perp \notin \tilde{\boldsymbol{x}}) \\
&\geq p_1(R'(x,\tilde{\boldsymbol{x}},M,s_2) = 1 \wedge \perp \notin \boldsymbol{x}')
\end{aligned}$$

$$
\begin{aligned}
&= p_1(R'(x, \tilde{\boldsymbol{x}}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}' \wedge \boldsymbol{x}' = \tilde{\boldsymbol{x}}) \\
&= p_1(R'(x, \boldsymbol{x}', M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}' \wedge \boldsymbol{x}' = \tilde{\boldsymbol{x}}) \\
&= p_1(R'(x, \boldsymbol{x}', M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}') \\
&= \Pr[\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}\text{-}1}_{\Pi',R',A'}(k) : w = 1],
\end{aligned}
$$

where the inequality follows from (4).

It follows from Definition 4 that if Π is secure in the sense of SNM$^{\circledast}$-CCA1, then there exist a polynomial p' and a simulator $S = (S_1, S_2)$ of the above adversary A, bounded by time $p'(k)$, such that $\mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}}_{\Pi,R,A,S}(k)$ is negligible. By using such S, let us next construct a simulator $S' = (S'_1, S'_2)$ of A' as

Algorithm $S'_1(pk')$	Algorithm $S'_2((\boldsymbol{y}', s_2))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$ $(\boldsymbol{y}, \boldsymbol{s}) \leftarrow S_2(s_1)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$ if $\lvert\boldsymbol{s}\rvert = 0$ then return $(M, ((), \varepsilon))$ parse $\boldsymbol{s}$ as $(s_2)\vert\vert\boldsymbol{a}'$ with $\lvert(s_2)\rvert = 1$ $\tilde{\boldsymbol{x}} \leftarrow \langle \boldsymbol{x} : 0 \vert \boldsymbol{a}' = 0 \rangle$; $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\tilde{\boldsymbol{x}})$ return $(M, (\boldsymbol{y}', s_2))$	return $(\boldsymbol{y}', s_2)$

Since S is bounded by time $p'(k)$ and $\mathcal{K}$, $\mathcal{E}_{pk'}$ and $\mathcal{D}_{sk}$ are polynomial-time, it follows that M is samplable in time $p'(k)$ and S' is polynomial-time. It can also be seen from the above construction of S' and R that

$$
\Pr[\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}\text{-}0}_{\Pi',R',S'}(k) : w = 1] \geq \Pr[\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}\text{-}0}_{\Pi,R,S}(k) : w = 1]
$$

(where equality holds if and only if S' always fails when $|\boldsymbol{s}| = 0$), and so

$$
\mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}}_{\Pi',R',A',S'}(k) \leq \mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}}_{\Pi,R,A,S}(k).
$$

Consequently, if Π is secure in the sense of SNM-CCA1, then $\mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-}\mathrm{CCA1}}_{\Pi,R,A,S}(k)$ is negligible, and so is $\mathsf{Adv}^{\mathrm{SNM}\text{-}\mathrm{CCA1}}_{\Pi',R',A',S'}(k)$. This completes the proof. □

Lemma 2. *Π' is not CNM*-CPA.*

Proof. Let $A = (A_1, A_2)$ be a CNM*-CPA adversary attacking Π' defined by

Algorithm $A_1(pk)$	Algorithm $A_2(s, (0, y))$
return $(\{0,1\}_U, \varepsilon)$	return $(R, ((1, y)))$

where the relation R output by A_2 is given by

Relation $R(x, \boldsymbol{x})$
  if $\boldsymbol{x} = (x)$ then return 1
  else return 0

It can be seen from this definition that M is samplable in time $O(1)$ and A is polynomial-time; moreover, since $|0| = |1|$ and $(0, y) \neq (1, y)$, A is legitimate. Since $\perp \notin \boldsymbol{x} \iff x_1 = 0$, it also follows that

$$\Pr[\mathsf{Expt}^{\text{CNM}^*\text{-ATK-1}}_{\Pi,A}(k) : w = 1] = \Pr[\mathsf{Expt}^{\text{CNM}^*\text{-ATK-1}}_{\Pi,A}(k) : x_1 = x_1 \wedge x_1 = 0] = \frac{1}{2},$$
$$\Pr[\mathsf{Expt}^{\text{CNM}^*\text{-ATK-0}}_{\Pi,A}(k) : w = 1] = \Pr[\mathsf{Expt}^{\text{CNM}^*\text{-ATK-0}}_{\Pi,A}(k) : x_0 = x_1 \wedge x_1 = 0] = \frac{1}{4},$$

and so

$$\mathsf{Adv}^{\text{CNM}^*\text{-ATK}}_{\Pi,A}(k) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4},$$

which is not negligible. This completes the proof. □

4 Relation to indistinguishability-based characterizations

To prove the separation between SNM* and IND*, we modify an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ to $\Pi' = (\mathcal{K}', \mathcal{E}', \mathcal{D}')$ so that the modified decryption algorithm $\mathcal{D}'$ has an additional "option" which makes an adversary and a simulator fail with probability at least $\frac{1}{2}$. More precisely, $\mathcal{E}'$ takes a plaintext x and outputs a ciphertext $(0, \mathcal{E}_{pk}(ux))$ with u being a random bit, and $\mathcal{D}'$ takes a ciphertext (a, y) and outputs $\hat{x}$ if $a = 0$ or $\hat{u} = 0$, otherwise $\perp$, where we have introduced $\hat{u}$ and $\hat{x}$ to denote the first bit and the remaining bits of $\mathcal{D}_{sk}(y)$, respectively (i.e. $\mathcal{D}_{sk}(y) = \hat{u}\hat{x}$ with $|\hat{u}| = 1$). It can be seen from this definition of $\mathcal{D}'$ that an adversary and a simulator fail with probability at least $\Pr[\hat{u} = 1] = \frac{1}{2}$ if they choose the option $a \neq 0$. Hence, there is no advantage for an IND* adversary with a message distribution whose support consists of two elements x_0 and x_1 to choose this option, while an SNM* adversary may take advantage from this option by choosing a message distribution M whose support consists of more than two elements. We are now ready to show the separation between SNM* and IND*.

Theorem 2. *IND*-PCA1 $\not\Rightarrow$ SNM*-CPA.*

Proof. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme. By using Π, let us construct another encryption scheme $\Pi' = (\mathcal{K}', \mathcal{E}', \mathcal{D}')$ as

Algorithm $\mathcal{K}'(1^k)$	Algorithm $\mathcal{E}'_{pk}(x)$	Algorithm $\mathcal{D}'_{sk}((a, y))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$	$u \leftarrow \{0,1\}_U$	$x' \leftarrow \mathcal{D}_{sk}(y)$
return (pk, sk)	$y \leftarrow \mathcal{E}_{pk}(ux)$	if $\lvert x'\rvert = 0$ then return $\perp$
	return $(0, y)$	else parse x' as ux with $\lvert u\rvert = 1$
		if $a = 0$ then return x
		else if $u = 0$ then return x
		else return $\perp$

Then, the theorem follows from the following two lemmas. □

The following two lemmas claim that (a) If Π is IND*-PCA1, then so is Π' and (b) Π' is not SNM*-CPA, respectively. To show the first lemma (a), we construct an adversary A attacking Π by using an adversary A' attacking Π' as before. Again, the construction of A from A' is straightforward except for the case where A'_2 outputs a sequence $\boldsymbol{y}'$ of ciphertexts which contains a component (a', y) such that $a' \neq 0$ and y is a challenge ciphertext for A. Let us thus consider the following construction for the exceptional case. Receiving two plaintexts x_0 and x_1 from A'_1, A_1 generates two plaintexts vx_0 and $\bar{v}x_1$ with v being a random bit, where $\bar{v}$ denotes the bit-wise complement of v. Note that vx_0 and $\bar{v}x_1$ can be expressed as $vx_0 = (v \oplus 0)x_0$ and $\bar{v}x_1 = (v \oplus 1)x_1$, respectively, and the distributions of $(v \oplus b)x_b$ and ux_b are identical if b, v and u are independent random bits. If we consider v as a guess of b, then $(v \oplus b)x_b$ has a form $0x_b$ if the guess is correct (i.e. $v = b$), otherwise it has a form of $1x_b$ and so $\mathcal{D}'_{sk}$ returns $\perp$. Now, A_2 generates a sequence $\boldsymbol{y}$ of ciphertexts by simply ignoring the first component a' of each component (a', y') of $\boldsymbol{y}'$. Next, A_2 replaces y in $\boldsymbol{y}$ by e such that $e \neq y$, and then includes the position of this replacement in the state information s_2 for A_3. Finally, A_3 replaces $\mathcal{D}_{sk}(e)$ at the position of the above replacement by x_v. It can be seen from this construction that A can completely simulate the view of A' if the guess is correct, while if the guess is not correct, then A' always fails because $\perp \in \mathcal{D}_{sk}(\boldsymbol{y}')$. It thus follows that the advantage of A' is upper-bounded by that of A. Moreover, we can show the second lemma (b) by considering a SNM*-CPA adversary which simply transforms a challenge ciphertext $(0, y)$ to $(1, y)$ and outputs it as a component of $\boldsymbol{y}$. Detailed proofs of these lemmas are described below.

Lemma 3. *If Π is IND*-PCA1, then so is Π'.*

Proof. Let p be a polynomial of k. Let $A' = (A'_1, A'_2, A'_3)$ be a legitimate IND*-PCA1 adversary attacking Π', bounded by time $p(k)$. By using A', let us construct an IND*-PCA1 adversary $A = (A_1, A_2, A_3)$ attacking Π as

Algorithm $A_1^{\mathcal{D}_{sk}}(pk)$
$(x_0, x_1, s_1) \leftarrow A_1'^{\mathcal{D}'_{sk}}(pk)$
$v \leftarrow \{0,1\}_U$; $L \leftarrow |x_0|+1$; $e \leftarrow \mathcal{E}_{pk}(0^L)$
return $(vx_0, \bar{v}x_1, (s_1, e, x_v))$

Algorithm $A_2(vx_0, \bar{v}x_1, (s_1, e, x_v), y)$
$(\boldsymbol{y}', s_2) \leftarrow A'_2(s_1, (0, y))$
$\boldsymbol{y} \leftarrow \langle e : \boldsymbol{y}'_{*2} | \boldsymbol{y}'_{*2} = y \rangle$
return $(\boldsymbol{y}, (s_2, \boldsymbol{y}'_{*2}, y, x_v))$

Algorithm $A_3(\boldsymbol{x}, (s_2, \boldsymbol{y}'_{*2}, y, x_v))$
$\tilde{\boldsymbol{x}} \leftarrow \langle x_v : \boldsymbol{x} | \boldsymbol{y}'_{*2} = y \rangle$; $d \leftarrow A'_3(\tilde{\boldsymbol{x}}, s_2)$
return d

where $\bar{v}$ denotes the bit-wise complement of v. Since A' is bounded by time $p(k)$ and $\mathcal{E}_{pk}$ is polynomial-time, it follows that M is samplable in time $p(k)$ and A and R are also polynomial-time. Moreover, A can be seen legitimate as follows: the condition (i) follows from that A' is legitimate, the condition (ii) from that $\boldsymbol{y} = \langle e : \boldsymbol{y}'_{*2} | \boldsymbol{y}'_{*2} = y \rangle$, where every component y has been replaced by e, and the condition (iii) from that A_2 has no oracle access to $\mathcal{D}_{sk}$. We note that A_1 can answer queries from A'_1 by using her own oracle $\mathcal{D}_{sk}$ to compute $\mathcal{D}'_{sk}$.

It is now convenient to consider the experiment $\mathsf{Expt}_2(k)$ defined by

$$
\begin{array}{l}
\mathsf{Experiment}\ \ \mathsf{Expt}_2(k) \\
\quad (pk, sk) \leftarrow \mathcal{K}(1^k);\ (x_0, x_1, s_1) \leftarrow A_1'^{\mathcal{D}'_{sk}}(pk);\ L \leftarrow |x_0| + 1;\ e \leftarrow \mathcal{E}_{pk}(0^L) \\
\quad b, u, v \leftarrow \{0,1\}_U;\ y \leftarrow \mathcal{E}_{pk}((v \oplus b)x_b);\ \hat{y} \leftarrow \mathcal{E}_{pk}(ux_b) \\
\quad \boldsymbol{y}' \leftarrow A_2'(s_1, (0, y));\ \hat{\boldsymbol{y}} \leftarrow A_2'(s_1, (0, \hat{y}));\ \boldsymbol{y} \leftarrow \langle e : \boldsymbol{y}'_{*2} | \boldsymbol{y}'_{*2} = y \rangle \\
\quad \boldsymbol{x}' \leftarrow \mathcal{D}'_{sk}(\boldsymbol{y}');\ \boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y});\ \hat{\boldsymbol{x}} \leftarrow \mathcal{D}_{sk}(\hat{\boldsymbol{y}});\ \tilde{\boldsymbol{x}} \leftarrow \langle x_v : \boldsymbol{x} | \boldsymbol{y}'_{*2} = y \rangle \\
\quad d' \leftarrow A_3'(\boldsymbol{x}', s_2);\ d \leftarrow A_3'(\tilde{\boldsymbol{x}}, s_2);\ \hat{d} \leftarrow A_3'(\hat{\boldsymbol{x}}, s_2)
\end{array}
$$

where we have used that vx_0 and $\bar{v}x_1$ can be expressed as $vx_0 = (v \oplus 0)x_0$ and $\bar{v}x_1 = (v \oplus 1)x_1$, respectively, and to introduce the short-hand notation $p_2(E) = \Pr[\mathsf{Expt}_2(k) : E]$, as before. We first note that the distributions of $(v \oplus b)x_b$ and ux_b are identical. Since $\boldsymbol{x} = \langle 0^L : \boldsymbol{x}' | \boldsymbol{y}'_{*2} = y \rangle$, we have

$$\perp \in \boldsymbol{x} \implies \perp \in \boldsymbol{x}'.$$

It thus follows that

$$
\begin{aligned}
\Pr[\mathsf{Expt}^{\mathrm{IND}^*\text{-PCA1}}_{\Pi,A}(k) : w = 1] &= p_2(d = b \wedge \perp \notin \boldsymbol{x}) \\
&\geq p_2(d = b \wedge \perp \notin \boldsymbol{x}') \\
&= p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v = b) \\
&\quad + p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v \neq b \wedge y \in \boldsymbol{y}'_{*2}) \\
&\quad + p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v \neq b \wedge y \notin \boldsymbol{y}'_{*2}).
\end{aligned}
$$

We now estimate the above three terms in the right-hand side. To consider the first term, we begin with expressing $\boldsymbol{x}$ as $\boldsymbol{x} = \langle 0^L : \boldsymbol{x}' | \boldsymbol{y}'_{*2} = y \rangle$. It can be seen from this expression that $\tilde{\boldsymbol{x}} = \langle x_v : \boldsymbol{x} | \boldsymbol{y}'_{*2} = y \rangle = \langle x_v : \boldsymbol{x}' | \boldsymbol{y}'_{*2} = y \rangle$, and so

$$v = b \implies \tilde{\boldsymbol{x}} = \langle x_b : \boldsymbol{x}' | \boldsymbol{y}'_{*2} = y \rangle = \boldsymbol{x}' \implies d = d'.$$

Hence, on noting that $v = b \iff v \oplus b = 0$ and $p_2(v \oplus b = 0) = p_2(u = 0) = \frac{1}{2}$, we have

$$
\begin{aligned}
p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v = b) &= p_2(d' = b \wedge \perp \notin \boldsymbol{x}' \wedge v = b) \\
&= p_2(d' = b \wedge \perp \notin \boldsymbol{x}' | v \oplus b = 0) p_2(v \oplus b = 0) \\
&= p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}} | u = 0) p_2(u = 0) \\
&= p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}} \wedge u = 0).
\end{aligned}
$$

To consider the second term, suppose that $y \in \boldsymbol{y}'_{*2}$, and let i be an index such that $\boldsymbol{y}'_i = (a, y)$. Then, since A' is legitimate, we have $(0, y) \notin \boldsymbol{y}'$ and so $a \neq 0$. Note here that $v \neq b \iff v \oplus b = 1$, and hence $\mathcal{D}_{sk}(y) = 1x_b$. It thus follows from the definition of $\mathcal{D}'_{sk}$ that $\boldsymbol{x}'_i = \mathcal{D}'_{sk}((a, y)) = \perp$. Similarly, if $\hat{y} \in \hat{\boldsymbol{y}}_{*2}$ and $u \neq 0$, then $\perp \in \hat{\boldsymbol{x}}$. Therefore,

$$p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v \neq b \wedge y \in \boldsymbol{y}'_{*2}) = p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}} \wedge u \neq 0 \wedge \hat{y} \in \hat{\boldsymbol{y}}_{*2}) = 0.$$

To consider the third term, we begin with

$$y \notin \boldsymbol{y}'_{*2} \implies \tilde{\boldsymbol{x}} = \langle x_v : \boldsymbol{x}' | \boldsymbol{y}'_{*2} = y \rangle = \boldsymbol{x}' \implies d = d'.$$

Hence, on noting that $v \neq b \iff v \oplus b = 1$, $u \neq 0 \iff u = 1$ and $p_2(v \oplus b = 1) = p_2(u = 1) = \frac{1}{2}$, we have

$$\begin{aligned}
&p_2(d = b \wedge \perp \notin \boldsymbol{x}' \wedge v \neq b \wedge y \notin \boldsymbol{y}'_{*2}) \\
&= p_2(d' = b \wedge \perp \notin \boldsymbol{x}' \wedge v \neq b \wedge y \notin \boldsymbol{y}'_{*2}) \\
&= p_2(d' = b \wedge \perp \notin \boldsymbol{x}' \wedge y \notin \boldsymbol{y}'_{*2} | v \oplus b = 1) p_2(v \oplus b = 1) \\
&= p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}} \wedge \hat{y} \notin \hat{\boldsymbol{y}}_{*2} | u = 1) p_2(u = 1) \\
&= p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}} \wedge u \neq 0 \wedge \hat{y} \notin \hat{\boldsymbol{y}}_{*2}).
\end{aligned}$$

Having estimated the three terms, we now combine these terms to give

$$\begin{aligned}
\Pr[\mathsf{Expt}^{\text{IND}^*\text{-PCA1}}_{\Pi,A}(k) : w = 1] &\geq p_2(d = b \wedge \perp \notin \boldsymbol{x}') \\
&= p_2(\hat{d} = b \wedge \perp \notin \hat{\boldsymbol{x}}) \\
&= \Pr[\mathsf{Expt}^{\text{IND}^*\text{-PCA1}}_{\Pi',A'}(k) : w = 1],
\end{aligned}$$

and hence

$$\mathsf{Adv}^{\text{IND}^*\text{-PCA1}}_{\Pi',A'}(k) \leq \mathsf{Adv}^{\text{IND}^*\text{-PCA1}}_{\Pi,A}(k).$$

Consequently, if Π is secure in the sense of IND*-PCA1, then $\mathsf{Adv}^{\text{IND}^*\text{-PCA1}}_{\Pi,A}(k)$ is negligible, and so is $\mathsf{Adv}^{\text{IND}^*\text{-PCA1}}_{\Pi',A'}(k)$. This completes the proof. □

Lemma 4. *Π' is not SNM*-CPA.*

Proof. Let $A = (A_1, A_2)$ be an SNM*-CPA adversary attacking Π' defined by

Algorithm $A_1(pk)$	Algorithm $A_2(s_1, (0, y))$
return $(\{0,1\}^2_U, \varepsilon, \varepsilon)$	return $((1, y))$

and let R be a relation defined by

```
Relation R(x, x, M, s_2)
  if M = {0,1}^2_U ∧ x = x_1 then return 1
  else return 0
```

It can be seen from the above definition of A that M is samplable in time $O(1)$ and A is polynomial-time; it also follows from $|00| = |01| = |10| = |11|$ and $(0, y) \neq (1, y)$ that A is legitimate. Now, it follows from the construction of A that

$$\Pr[\mathsf{Expt}^{\text{SNM}^*\text{-ATK-1}}_{\Pi',R,A}(k) : w = 1] = \Pr[\mathsf{Expt}^{\text{SNM}^*\text{-ATK-1}}_{\Pi',R,A}(k) : u = 0] = \frac{1}{2},$$

where u is the random variable introduced in the definition of $\mathcal{E}'_{pk}$. On the other hand, S is given no information about the plaintext x, and hence the outputs from S are statistically independent of x. Consequently, since x is uniformly distributed on $\{0,1\}^2$, we have

$$\begin{aligned}\Pr[\mathsf{Expt}^{\mathrm{SNM}^*\text{-ATK-0}}_{\Pi',R,S}(k) : w = 1] &= \Pr[\mathsf{Expt}^{\mathrm{SNM}^*\text{-ATK-0}}_{\Pi',R,S}(k) : x = \boldsymbol{x}_1] \\ &\leq \frac{1}{|\{0,1\}^2|} = \frac{1}{4}\end{aligned}$$

(where equality holds if and only if S outputs $M = \{0,1\}^2$ and $\boldsymbol{y}$ such that $\mathcal{D}_{sk'}(\boldsymbol{y}_1) \in \{0,1\}^2$), and so

$$\mathsf{Adv}^{\mathrm{SNM}^*\text{-ATK}}_{\Pi',R,A,S}(k) \geq \frac{1}{4},$$

which is not negligible. This completes the proof. □

We have examined the relation between the simulator-based non-malleability and the indistinguishability-based characterization of non-malleability under the valid ciphertext condition. Hence, it may be natural to next consider the relation between the comparison-based non-malleability. and an indistinguishability-based characterization of non-malleability. In the private-key setting, Katz and Yung [15] introduced a slightly modified indistinguishability-based characterization of non-malleability, denoted as IND†-PCAX in this paper, and proved that IND†-PCAX are equivalent to CNM*-ATK. The proof of this equivalence for the private-key setting straightforwardly applies to the public-key setting. A formal definition of IND†-PCAX is given in Appendix C.

5 Concluding remarks

We have shown two separations SNM* $\not\Rightarrow$ CNM* and IND* $\not\Rightarrow$ SNM* against the weaker types of the attack models. As long as we consider the attack by the CNM* adversary in the proof of Lemma 2, CNM* may seem stronger than expected from our intuition for non-malleability, and so SNM*-CPA and SNM*-CCA1 may receive independent interest as (intuitively) natural formulations of non-malleability. Here, one motivation to consider a weaker security notion would be to provide a better construction of cryptosystems secure in the sense of the weaker notion, and thus it may be of interest to consider the possibility of an indistinguishability-based characterization of SNM*.

In the proof of the latter separation, it is essential that an IND* adversary has to output a message distribution whose support consists of exactly two elements, but an SNM* adversary is free of such restriction on a message distribution. This may motivate us to consider simulation-based and comparison-based formulations of semantic security against parallel chosen-ciphertext attacks, SSS*-PCAX and CSS*-PCAX, because semantic security is commonly formulated without such restriction on a message distribution, and so may be (potentially) stronger than IND*-PCAX. In fact, it turns out that SSS*-PCAX and CSS*-PCAX are equivalent to SNM*-ATK and CNM*-ATK, respectively (see Appendix B). Hence, it follows from this equivalence, together with Theorem 2, that semantic security and ciphertext indistinguishability separate against the weaker parallel chosen-ciphertext attacks under the valid ciphertext condition.

The pairs of notions appearing in the above two separations have in common that one is a natural formulation of security and the other is its simpler characterization shown equivalent to the original one in some standard settings. Therefore, these separations demonstrate that even such security notions may separate if we additionally impose some natural condition (e.g. the valid ciphertext condition) or we introduce some technically useful setting (e.g. parallel chosen-ciphertext attacks).

References

1. Bellare, M., Desai, A., Jokipii, E., Rogaway, P.: A concrete security treatment of symmet-ric encryption: Analysis of the des modes of operation. In: FOCS. pp. 394–403 (1997)
2. Bellare, M., Desai, A., Pointcheval, D., Rogaway, P.: Relations among notions of security for public-key encryption schemes. In: CRYPTO. pp. 26–45 (1998), https://eprint.iacr.org/1998/021
3. Bellare, M., Sahai, A.: Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In: CRYPTO. pp. 519–536 (1999), Full version available at https://eprint.iacr.org/2006/228
4. Canetti, R., Varia, M.: Non-malleable obfuscation. In: TCC. pp. 73–90 (2009)
5. Coretti, S., Dodis, Y., Tackmann, B., Venturi, D.: Non-malleable encryption: Simpler, shorter, stronger. In: TCC. pp. 306–335 (2016)
6. Coretti, S., Maurer, U., Tackmann, B., Venturi, D.: From single-bit to multi-bit public-key encryption via non-malleable codes. In: TCC. pp. 532–560 (2015)
7. Di Crescenzo, G., Ishai, Y., Ostrovsky, R.: Non-interactive and non-malleable commitment. In: STOC. pp. 141–150 (1998)
8. Dolev, D., Dwork, C., Naor, M.: Non-malleable cryptography. SIAM Journal on Computing **30**(2), 391–437 (2000)
9. Dziembowski, S., Pietrzak, K., Wichs, D.: Non-malleable codes. In: TCC. pp. 434–452 (2010), https://eprint.iacr.org/2009/608
10. Goldreich, O.: A uniform-complexity treatment of encryption and zero-knowledge. Journal of Cryptology **6**(1), 21–53 (1993)
11. Goldreich, O., Lustig, Y., Naor, M.: On chosen ciphertext security of multiple encryptions. Cryptology ePrint Archive, Paper 2002/089 (2002), https://eprint.iacr.org/2002/089
12. Goldwasser, S., Micali, S.: Probabilistic encryption. Journal of Computer and System Sciences **28**(2), 270–299 (1984)
13. Goyal, V., Kumar, A.: Non-malleable secret sharing. In: STOC. pp. 685–698 (2018), https://eprint.iacr.org/2018/316
14. Herranz, J., Hofheinz, D., Kiltz, E.: Kem/dem: Necessary and sufficient conditions for secure hybrid encryption (2006), http://eprint.iacr.org/2006/265

15. Katz, J., Yung, M.: Characterization of security notions for probabilistic private-key encryption. Journal of Cryptology **19**(1), 67–95 (2006)
16. Micali, S., Rackoff, C., Sloan, B.: The notion of security for probabilistic cryptosystems. SIAM Journal on Computing **17**(2), 412–426 (1988)
17. Naor, M., Yung, M.: Public-key cryptosystems provably secure against chosen ciphertext attacks. In: STOC. pp. 427–437 (1990)
18. Rackoff, C., Simon, D.R.: Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In: CRYPTO. pp. 433–444 (1991)
19. Sahai, A.: Non-malleable non-interactive zero knowledge and adaptive chosen-ciphertext security. In: FOCS. pp. 543–553 (1999)
20. Watanabe, Y., Shikata, J., Imai, H.: Equivalence between semantic security and indistinguishability against chosen ciphertext attacks. In: PKC. pp. 71–84 (2002)

A Equivalence between SNM* and SNM$^{\circledast}$

In this appendix, we first provide a formal definition of SNM$^{\circledast}$, a slight modification of SNM* in which the output of the side information s_2 for the relation R is delayed from the first stage to the second one.

Definition 4 (SNM$^{\circledast}$-ATK). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and R be a relation. Let $A = (A_1, A_2)$ be an adversary attacking Π and $S = (S_1, S_2)$ be its simulator. For $k \in \mathbb{N}$ and $\mathrm{ATK} \in \{\mathrm{CPA}, \mathrm{CCA1}, \mathrm{CCA2}\}$, consider the following two experiments:*

Experiment $\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-ATK-1}}_{\Pi,R,A}(k)$	Experiment $\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-ATK-0}}_{\Pi,R,S}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$	$(pk, sk) \leftarrow \mathcal{K}(1^k)$
$(M, s_1) \leftarrow A_1^{\mathcal{O}_1}(pk)$	$(M, s_1) \leftarrow S_1(pk)$
$x \leftarrow M$; $y \leftarrow \mathcal{E}_{pk}(x_0)$	$x \leftarrow M$
$(\boldsymbol{y}, s_2) \leftarrow A_2^{\mathcal{O}_2}(s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$	$(\boldsymbol{y}, s_2) \leftarrow S_2(s_1)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
if $R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$	if $R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$	else $w \leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are defined as in Definition 1. For a function f of k, an adversary A (resp. a simulator S) is called bounded by time $f(k)$ *analogously to Definition 1. Then, an encryption scheme Π is called* secure in the sense of SNM$^{\circledast}$-ATK *if for all polynomial p, all probabilistic adversary A bounded by time $p(k)$ and all relation R computable in time $p(k)$, there exist a polynomial $p'(k)$ and a simulator S bounded by time $p'(k)$ such that $\mathsf{Adv}^{SNM^{\circledast}\text{-}ATK}_{\Pi,R,A,S}(k)$ is negligible, where $\mathsf{Adv}^{SNM^{\circledast}\text{-}ATK}_{\Pi,R,A,S}$ denotes the advantage of A against S defined by*

$$\begin{aligned}&\mathsf{Adv}^{SNM^{\circledast}\text{-}ATK}_{\Pi,R,A,S}(k)\\&= \Pr[\mathsf{Expt}^{SNM^{\circledast}\text{-}ATK\text{-}1}_{\Pi,R,A}(k) : w = 1] - \Pr[\mathsf{Expt}^{SNM^{\circledast}\text{-}ATK\text{-}0}_{\Pi,R,S}(k) : w = 1].\end{aligned}$$

Having provided a formal definition of SNM$^{\circledast}$, we next show that SNM$^{\circledast}$-ATK is equivalent to SNM*-ATK. The proof is rather straightforward, but it may help to note how to construct an SNM* adversary, which outputs side information for a relation at the first stage, from an SNM$^{\circledast}$ adversary, which outputs it at the second stage. For this construction, we may employ the technique used in the proof of SNM-ATK $\Longrightarrow$ IND-PCAX in [3]; namely, we make an SNM* adversary to concatenate an encryption of side information received from an SNM$^{\circledast}$ adversary to the sequence $\boldsymbol{y}$ of ciphertexts, so that the relation can take the side information as a part of $\boldsymbol{x} = \mathcal{D}_{sk}(\boldsymbol{y})$. A detailed proof of the equivalence is described below.

Proposition 1. *SNM*-ATK $\Longleftrightarrow$ SNM$^{\circledast}$-ATK for* ATK $\in$ {CPA, CCA1, CCA2}.

Proof. (I) SNM*-ATK $\Longrightarrow$ SNM$^{\circledast}$-ATK: Let p be a polynomial of k. Let R' be a relation computable in time $p(k)$ and $A' = (A'_1, A'_2)$ be a legitimate SNM$^{\circledast}$-ATK adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A' and R', let us construct an SNM*-ATK adversary $A = (A_1, A_2)$ attacking Π and a relation R as

Algorithm $A_1^{\mathcal{O}_1}(pk)$	Algorithm $A_2^{\mathcal{O}_2}((s_1, pk, 0^L), y)$
$(M, s_1) \leftarrow A_1'^{\mathcal{O}_1}(pk)$	$(\boldsymbol{y}', s_2) \leftarrow A_2'^{\mathcal{O}_2}(s_1, y)$
$x' \leftarrow M$; $L \leftarrow \lvert x'\rvert + 1$	$\boldsymbol{y} \leftarrow \boldsymbol{y}' \| (\mathcal{E}_{pk}(0^L s_2))$
return $(M, (s_1, pk, 0^L), 0^L)$	return $\boldsymbol{y}$

Relation $R(x, \boldsymbol{x}, M, s_2)$
  $L \leftarrow |s_2|$
  if $|\boldsymbol{x}| = 0$ then return 0
  else parse $\boldsymbol{x}$ as $\tilde{\boldsymbol{x}} \| (s')$ with $|(s')| = 1$
  parse s' as $s'_1 s'_2$ with $|s'_1| = \min(L, |s'|)$
  return $R'(x, \tilde{\boldsymbol{x}}, M, s'_2)$

where L is chosen so that $\mathcal{E}_{pk}(0^L s_2) \neq y$ with probability 1 as before. Since A' is bounded by time $p(k)$, R' is computable in time $p(k)$ and $\mathcal{E}_{pk}$ is polynomial-time, it follows that M is samplable in time $p(k)$ and A and R are also polynomial-time. Moreover, since A' is legitimate and $\mathcal{E}_{pk}(0^L s_2) \neq y$, A is also legitimate. We note that A can answer queries from A' by using her own oracle. It is now straightforward to see from the above construction of A and R that $\tilde{\boldsymbol{x}} = \mathcal{D}_{sk}(\boldsymbol{y}')$ and $s'_2 = s_2$, and so

$$\Pr[\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-ATK-1}}_{\Pi,R',A'}(k) : w = 1] = \Pr[\mathsf{Expt}^{\mathrm{SNM}^{*}\text{-ATK-1}}_{\Pi,R,A}(k) : w = 1].$$

It follows from Definition 1 that if Π is secure in the sense of SNM*-ATK, then there exist a polynomial p' and a simulator $S = (S_1, S_2)$ of the above adversary A, bounded by time $p'(k)$, such that $\mathsf{Adv}^{\mathrm{SNM}^{*}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible. By using such S, let us next construct a simulator $S' = (S'_1, S'_2)$ of A' as

Algorithm $S'_1(pk')$ $(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1, s_2) \leftarrow S_1(pk)$ $\boldsymbol{y} \leftarrow S_2(s_1)$; $L \leftarrow \lvert s_2 \rvert$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$ if $\lvert\boldsymbol{x}\rvert = 0$ then return $(M, ((), \varepsilon))$ else parse $\boldsymbol{x}$ as $\tilde{\boldsymbol{x}} \| (s')$ with $\lvert(s')\rvert = 1$ parse s' as $s'_1 s'_2$ with $\lvert s'_1\rvert = \min(L, \lvert s'\rvert)$ $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\tilde{\boldsymbol{x}})$ return $(M, (\boldsymbol{y}', s'_2))$	Algorithm $S'_2((\boldsymbol{y}', s'_2))$ return $(\boldsymbol{y}', s'_2)$

Since S is bounded by time $p'(k)$ and $\mathcal{K}$, $\mathcal{E}_{pk'}$ and $\mathcal{D}_{sk}$ are polynomial-time, it follows that M is samplable in time $p'(k)$ and S' is also polynomial-time. Hence, the above construction of S' and R gives that

$$\Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-ATK-0}}_{\Pi,R',S'}(k) : w = 1] \geq \Pr[\mathsf{Expt}^{\text{SNM}^{*}\text{-ATK-0}}_{\Pi,R,S}(k) : w = 1]$$

(where equality holds if and only if S' always fails when $|\boldsymbol{x}| = 0$), and so

$$\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R',A',S'}(k) \leq \mathsf{Adv}^{\text{SNM}^{*}\text{-ATK}}_{\Pi,R,A,S}(k).$$

Consequently, if Π is secure in the sense of SNM-ATK, then $\mathsf{Adv}^{\text{SNM}^{*}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible, and so is $\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi',R',A',S'}(k)$. This completes the proof of (I).

(II) SNM$^{\circledast}$-ATK $\Longrightarrow$ SNM*-ATK: Let p be a polynomial of k. Let R' be a relation computable in time $p(k)$ and $A' = (A'_1, A'_2)$ be a legitimate SNM*-ATK adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A' and R', let us construct an SNM$^{\circledast}$-ATK adversary $A = (A_1, A_2)$ attacking Π and a relation R as

Algorithm $A_1^{\mathcal{O}_1}(pk)$ $(M, s_1, s_2) \leftarrow A_1'^{\mathcal{O}_1}(pk)$ return $(M, (s_1, s_2))$	Algorithm $A_2^{\mathcal{O}_2}((s_1, s_2), y)$ $\boldsymbol{y}' \leftarrow A_2'^{\mathcal{O}_2}(s_1, y)$ return $(\boldsymbol{y}', s_2)$	Relation $R(x, \boldsymbol{x}, M, s_2)$ return $R'(x, \boldsymbol{x}, M, s_2)$

Since A' is bounded by time $p(k)$ and R' is computable in time $p(k)$, it follows that M is samplable in time $p(k)$ and A and R are polynomial-time. Moreover, since A' is legitimate, A is also legitimate. We note that A can answer queries from A' by using her own oracle. The above construction of A and R at once gives that

$$\Pr[\mathsf{Expt}^{\text{SNM}^{*}\text{-ATK-1}}_{\Pi,R',A'}(k) : w = 1] = \Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-ATK-1}}_{\Pi,R,A}(k) : w = 1].$$

It follows from Definition 4 that if Π is secure in the sense of SNM$^{\circledast}$-ATK, then there exist a polynomial p' and a simulator $S = (S_1, S_2)$ of the above adversary A, bounded by time $p'(k)$, such that $\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible. By using such S, let us next construct a simulator $S' = (S'_1, S'_2)$ of A' as[9]

[9] The notation $\langle 0 : \boldsymbol{x} | \boldsymbol{x} = \perp \rangle$ represents a copy of sequence $\boldsymbol{x}$ such that every component $\perp$ in $\boldsymbol{x}$ has been replaced by 0. If the symbol $\perp$ is in the domain of $\mathcal{E}_{pk'}$, then we may omit this replacement, which makes the proof simpler.

Algorithm $S'_1(pk')$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$; $(\boldsymbol{y}, s_2) \leftarrow S_2(s_1)$
$\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $\tilde{\boldsymbol{x}} \leftarrow \langle 0 : \boldsymbol{x} | \boldsymbol{x} = \perp \rangle$; $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\tilde{\boldsymbol{x}})$
return $(M, (\boldsymbol{y}', s_2))$

Algorithm $S'_2((\boldsymbol{y}', s_2))$
return $(\boldsymbol{y}', s_2)$

Since S is bounded by time $p'(k)$ and $\mathcal{K}$, $\mathcal{E}_{pk'}$ and $\mathcal{D}_{sk}$ are polynomial-time, it follows that M is samplable in time $p'(k)$ and S' is also polynomial-time. It is now convenient to consider the experiment $\mathsf{Expt}_0(k)$ defined by

Experiment $\mathsf{Expt}_0(k)$
$(pk, sk), (pk', sk') \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$; $x \leftarrow M$; $(\boldsymbol{y}, s_2) \leftarrow S_2(s_1)$
$\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $\tilde{\boldsymbol{x}} \leftarrow \langle 0 : \boldsymbol{x} | \boldsymbol{x} = \perp \rangle$; $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\tilde{\boldsymbol{x}})$; $\boldsymbol{x}' \leftarrow \mathcal{D}_{sk'}(\boldsymbol{y}')$

and to introduce the short-hand notation $p_0(E) = \Pr[\mathsf{Expt}_0(k) : E]$, as before. Since $\boldsymbol{x}' = \tilde{\boldsymbol{x}} = \langle 0 : \boldsymbol{x} | \boldsymbol{x} = \perp \rangle$, we have

$$\perp \notin \boldsymbol{x}' \quad \text{and} \quad \perp \notin \boldsymbol{x} \implies \boldsymbol{x} = \boldsymbol{x}'.$$

Therefore,

$$\begin{aligned}
&\Pr[\mathsf{Expt}^{\text{SNM}^*\text{-CCA1-0}}_{\Pi,R',S'}(k) : w = 1] \\
&= p_0(R'(x, \boldsymbol{x}', M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}') \\
&= p_0(R(x, \boldsymbol{x}', M, s_2) = 1) \\
&\geq p_0(R(x, \boldsymbol{x}', M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}) \\
&= p_0(R(x, \boldsymbol{x}', M, s_2) = 1 \wedge \perp \notin \boldsymbol{x} \wedge \boldsymbol{x} = \boldsymbol{x}') \\
&= p_0(R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x} \wedge \boldsymbol{x} = \boldsymbol{x}') \\
&= p_0(R(x, \boldsymbol{x}, M, s_2) = 1 \wedge \perp \notin \boldsymbol{x}) \\
&= \Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-CCA1-0}}_{\Pi,R,S}(k) : w = 1],
\end{aligned}$$

and so

$$\mathsf{Adv}^{\text{SNM}^*\text{-ATK}}_{\Pi,R',A',S'}(k) \leq \mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k).$$

Consequently, if Π is secure in the sense of SNM$^{\circledast}$-ATK, then $\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible, and so is $\mathsf{Adv}^{\text{SNM}^*\text{-ATK}}_{\Pi',R',A',S'}(k)$. This completes the proof of (II), and the proposition follows. □

B Semantic security against parallel chosen-ciphertext attacks

In this appendix, we first provide formal definitions of simulation-based and comparison-based semantic security against parallel chosen-ciphertext attacks under the valid ciphertext condition, SSS*-PCAX and CSS*-PCAX, each of which is a straightforward combination of the definitions of semantic security [12] and parallel chosen-ciphertext attacks [3].

Definition 5 (SSS*-PCAX). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and F be a function. Let $A = (A_1, A_2, A_3)$ be an adversary attacking Π and $S = (S_1, S_2, S_3)$ be its simulator. For $k \in \mathbb{N}$ and* PCAX $\in$ {PCA0, PCA1, PCA2}, *consider the following two experiments:*

Experiment $\mathsf{Expt}^{\mathrm{SSS^*\text{-}PCAX\text{-}1}}_{\Pi,F,A}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$
$(M, s_1) \leftarrow A_1^{\mathcal{O}_1}(pk)$
$x \leftarrow M$; $y \leftarrow \mathcal{E}_{pk}(x_0)$
$(\boldsymbol{y}, s_2) \leftarrow A_2^{\mathcal{O}_2}(s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
$(v, s_3) \leftarrow A_3^{\mathcal{O}_2}(\boldsymbol{x}, s_2)$
if $F(x, M, s_3) = v \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$

Experiment $\mathsf{Expt}^{\mathrm{SSS^*\text{-}PCAX\text{-}0}}_{\Pi,F,S}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$
$(M, s_1) \leftarrow S_1(pk)$
$x \leftarrow M$
$(\boldsymbol{y}, s_2) \leftarrow S_2(s_1)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
$(v, s_3) \leftarrow S_3(\boldsymbol{x}, s_2)$
if $F(x, M, s_3) = v \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are defined as in Definition 3. Then, an encryption scheme Π is called secure in the sense of SSS*-PCAX *if for all polynomial p, all probabilistic adversary A bounded by time $p(k)$ and all function F computable in time $p(k)$, there exist a polynomial $p'(k)$ and a simulator S bounded by time $p'(k)$ such that $\mathsf{Adv}^{SSS^*\text{-}ATK}_{\Pi,R,A,S}(k)$ is negligible, where $\mathsf{Adv}^{SSS^*\text{-}ATK}_{\Pi,R,A,S}$ denotes the advantage of A against S defined by*

$$\mathsf{Adv}^{SSS^*\text{-}PCAX}_{\Pi,F,A,S}(k) = \Pr[\mathsf{Expt}^{SSS^*\text{-}PCAX\text{-}1}_{\Pi,F,A}(k) : w = 1] - \Pr[\mathsf{Expt}^{SSS^*\text{-}PCAX\text{-}0}_{\Pi,F,S}(k) : w = 1].$$

Definition 6 (CSS*-PCAX). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $A = (A_1, A_2, A_3)$ be an adversary attacking Π. For $k \in \mathbb{N}$ and* PCAX $\in$ {PCA0, PCA1, PCA2}, *consider the following experiment:*

Experiment $\mathsf{Expt}^{\mathrm{CSS^*\text{-}PCAX\text{-}}b}_{\Pi,A}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow A_1^{\mathcal{O}_1}(pk)$; $x_0, x_1 \leftarrow M$; $y \leftarrow \mathcal{E}_{pk}(x_1)$
$(\boldsymbol{y}, s_2) \leftarrow A_2^{\mathcal{O}_2}(s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $(f, v) \leftarrow A_3^{\mathcal{O}_2}(\boldsymbol{x}, s_2)$
if $f(x_b) = v \wedge \perp \notin \boldsymbol{x}$ then $w \leftarrow 1$
else $w \leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are defined as in Definition 3. Then, an encryption scheme Π is called secure in the sense of CSS*-PCAX *if for all polynomial p and all probabilistic adversary A runnable in time $p(k)$, $\mathsf{Adv}^{CSS^*\text{-}PCAX}_{\Pi,A}(k)$ is negligible, where $\mathsf{Adv}^{CSS^*\text{-}PCAX}_{\Pi,A}$ denotes the advantage of A defined by*

$$\mathsf{Adv}^{CSS^*\text{-}PCAX}_{\Pi,A}(k) = \Pr[\mathsf{Expt}^{CSS^*\text{-}PCAX\text{-}1}_{\Pi,A}(k) : w = 1] - \Pr[\mathsf{Expt}^{CSS^*\text{-}PCAX\text{-}0}_{\Pi,A}(k) : w = 1].$$

Having provided formal definitions, we next show that SSS*-PCAX and CSS*-PCAX are equivalent to SNM*-ATK and CNM*-ATK, respectively. We note that the techniques used in the proofs of equivalence between non-malleability and its indistinguishability-based characterization [3] can also be used to show the above equivalence; for example, in order for relation R to run a probabilistic algorithm, one can include randomness for the algorithm in side information for R. A detailed proof of the equivalence is described below. For simplicity of the proof, we consider SNM$^{\circledast}$ instead of SNM* as before.

Proposition 2. *SSS***-PCAX* $\Longleftrightarrow$ *SNM*$^{\circledast}$*-ATK and CSS***-PCAX* $\Longleftrightarrow$ *CNM***-ATK for* (PCAX, ATK)$\in$ {(PCA0, CPA), (PCA1, CCA1), (PCA2, CCA2)}.

Proof. (I) SSS*-PCAX $\Longrightarrow$ SNM$^{\circledast}$-ATK: Let p be a polynomial of k. Let R' be a relation computable in time $p(k)$ and $A' = (A'_1, A'_2)$ be a legitimate SNM$^{\circledast}$-ATK adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A' and R', let us construct an SSS*-PCAX adversary $A = (A_1, A_2, A_3)$ attacking Π and a function F as

Algorithm $A_1^{\mathcal{O}_1}(pk)$	Algorithm $A_2^{\mathcal{O}_2}(s_1, y)$	Algorithm $A_3^{\mathcal{O}_2}(s_2, \boldsymbol{x})$
$(M, s_1) \leftarrow A_1'^{\mathcal{O}_1}(pk)$	$(\boldsymbol{y}, s_2) \leftarrow A_2'^{\mathcal{O}_2}(s_1, y)$	$\boldsymbol{s} \leftarrow \boldsymbol{x}\|\|(s_2)$
return (M, s_1)	return $(\boldsymbol{y}, s_2)$	return $(1, \boldsymbol{s})$

Function $F(x, M, \boldsymbol{s})$
if $|\boldsymbol{s}| = 0$ then return 0
else parse $\boldsymbol{s}$ as $\boldsymbol{x}||(s)$ with $|(s)| = 1$
return $R'(x, \boldsymbol{x}, M, s)$

Since A' is bounded by time $p(k)$ and R' is computable in time $p(k)$, it follows that M is samplable in time $p(k)$ and A and F are also polynomial-time. Moreover, since A' is legitimate, A is also legitimate. We note that A can answer queries from A' by using her own oracle. It is now straightforward to see from the above construction of A and F that

$$\Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-ATK-1}}_{\Pi,R',A'}(k) : w = 1] = \Pr[\mathsf{Expt}^{\text{SSS}^{*}\text{-ATK-1}}_{\Pi,F,A}(k) : w = 1].$$

It follows from Definition 5 that if Π is secure in the sense of SSS*-PCAX, then there exist a polynomial p' and a simulator $S = (S_1, S_2, S_3)$ of the above adversary A, bounded by time $p'(k)$, such that $\mathsf{Adv}^{\text{SSS}^{*}\text{-PCAX}}_{\Pi,F,A,S}(k)$ is negligible. By using such S, let us next construct a simulator $S' = (S'_1, S'_2)$ of A' as

Algorithm $S'_1(pk')$	Algorithm $S'_2((\boldsymbol{y}', s))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$	return $(\boldsymbol{y}', s)$
$(\boldsymbol{y}, s_2) \leftarrow S_2(s_1)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $(v, \boldsymbol{s}) \leftarrow S_3(\boldsymbol{x}, s_2)$	
if $\|\boldsymbol{s}\| = 0$ then return $(M, ((), \varepsilon))$	
else parse $\boldsymbol{s}$ as $\boldsymbol{x}'\|\|(s)$ with $\|(s)\| = 1$	
$\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\boldsymbol{x}')$	
return $(M, (\boldsymbol{y}', s))$	

Since S is bounded by time $p'(k)$ and $\mathcal{K}$, $\mathcal{E}_{pk'}$ and $\mathcal{D}_{sk}$ are polynomial-time, it follows that M is samplable in time $p'(k)$ and S' is also polynomial-time. Then, the above construction of S' and F gives that

$$\Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-ATK-0}}_{\Pi,R',S'}(k) : w = 1] \geq \Pr[\mathsf{Expt}^{\text{SSS}^{*}\text{-PCAX-0}}_{\Pi,F,S}(k) : w = 1]$$

(where equality holds if and only if S' always fails when $|\boldsymbol{s}| = 0$), and so

$$\mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-ATK}}_{\Pi,R',A',S'}(k) \le \mathsf{Adv}^{\mathrm{SSS}^{*}\text{-PCAX}}_{\Pi,F,A,S}(k).$$

Consequently, if Π is secure in the sense of SSS*-PCAX, then $\mathsf{Adv}^{\mathrm{SSS}^{*}\text{-PCAX}}_{\Pi,F,A,S}(k)$ is negligible, and so is $\mathsf{Adv}^{\mathrm{SNM}^{\circledast}\text{-ATK}}_{\Pi,R',A',S'}(k)$. This completes the proof of (I).

(II) SNM$^{\circledast}$-ATK $\Longrightarrow$ SSS*-PCAX: Let p be a polynomial of k. Let F' be a function computable in time $p(k)$ and $A' = (A'_1, A'_2, A'_3)$ be a legitimate SSS*-PCAX adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A' and F', let us construct an SSS*-ATK adversary $A = (A_1, A_2)$ attacking Π and a relation R as

Algorithm $A_1^{\mathcal{D}_{sk}(\cdot)}(pk)$	Algorithm $A_2^{\mathcal{D}_{sk}(\cdot)}(s_1, y)$
$(M, s_1) \leftarrow A_1'^{\mathcal{D}_{sk}(\cdot)}(pk)$	$(\boldsymbol{y}, s_2) \leftarrow A_2'^{\mathcal{D}_{sk}(\cdot)}(s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
return (M, s_1)	$(v, s_3) \leftarrow A_3'^{\mathcal{D}_{sk}(\cdot)}(s_2, \boldsymbol{x})$; $\boldsymbol{s} \leftarrow (v, s_3)$
	return $(\boldsymbol{y}, \boldsymbol{s})$

Relation $R(x, \boldsymbol{x}, M, \boldsymbol{s})$
 if $|\boldsymbol{s}| \ne 2$ then return 0
 else parse $\boldsymbol{s}$ as (v, s_3)
 if $F'(x, M, s_3) = v$ then return 1
 else return 0

for $(\mathrm{PCAX}, \mathrm{ATK}) = (\mathrm{PCA2}, \mathrm{CCA2})$, otherwise as

Algorithm $A_1^{\mathcal{O}_1}(pk)$	Algorithm $A_2(s_1, y)$
$(M, s_1) \leftarrow A_1'^{\mathcal{O}_1}(pk)$	$(\boldsymbol{y}, s_2) \leftarrow A'_2(s_1, y)$
return (M, s_1)	choose randomness r for A'_3
	$\boldsymbol{s} \leftarrow (r, s_2)$
	return $(\boldsymbol{y}, \boldsymbol{s})$

Relation $R(x, \boldsymbol{x}, M, \boldsymbol{s})$
 if $|\boldsymbol{s}| \ne 2$ then return 0
 else parse $\boldsymbol{s}$ as (r, s_2)
 $(v, s_3) \leftarrow A'_3(s_2, \boldsymbol{x}; r)$
 if $F'(x, M, s_3) = v$ then return 1
 else return 0

Since A' is bounded by time $p(k)$ and F' is computable in time $p(k)$, it follows that M is samplable in time $p(k)$ and A and R are also polynomial-time. Moreover, since A' is legitimate, A is also legitimate. We note that A can answer queries from A' by using her own oracle. It is now straightforward to see from the above construction of A and R that

$$\Pr[\mathsf{Expt}^{\mathrm{SSS}^{*}\text{-ATK-1}}_{\Pi,F',A'}(k) : w = 1] = \Pr[\mathsf{Expt}^{\mathrm{SNM}^{\circledast}\text{-ATK-1}}_{\Pi,R,A}(k) : w = 1].$$

It follows from Definition 4 that if Π is secure in the sense of SNM$^{\circledast}$-ATK, then there exist a polynomial p' and a simulator $S = (S_1, S_2)$ of the above adversary A, bounded by time $p'(k)$, such that $\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible. By using such S, let us next construct a simulator $S' = (S'_1, S'_2, S'_3)$ of A' as

Algorithm $S'_1(pk')$	Algorithm $S'_2((\boldsymbol{y}', v, s_3))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$; $(\boldsymbol{y}, \boldsymbol{s}) \leftarrow S_2(s_1)$	return $(\boldsymbol{y}', (v, s_3))$
if $\lvert\boldsymbol{s}\rvert \neq 2$ then return $(M, ((), \varepsilon, \varepsilon))$	
else parse $\boldsymbol{s}$ as (v, s_3)	Algorithm $S'_3((v, s_3), \boldsymbol{x})$
$\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\boldsymbol{x})$	return (v, s_3)
return $(M, (\boldsymbol{y}', v, s_3))$	

for $(\text{PCAX}, \text{ATK}) = (\text{PCA2}, \text{CCA2})$, otherwise as

Algorithm $S'_1(pk')$	Algorithm $S'_2((\boldsymbol{y}', v, s_3))$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$; $(M, s_1) \leftarrow S_1(pk)$; $(\boldsymbol{y}, \boldsymbol{s}) \leftarrow S_2(s_1)$	return $(\boldsymbol{y}', (v, s_3))$
if $\lvert\boldsymbol{s}\rvert \neq 2$ then return $(M, ((), \varepsilon, \varepsilon))$	
else parse $\boldsymbol{s}$ as (r, s_2)	Algorithm $S'_3((v, s_3), \boldsymbol{x})$
$\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$; $(v, s_3) \leftarrow A'_3(s_2, \boldsymbol{x}; r)$; $\boldsymbol{y}' \leftarrow \mathcal{E}_{pk'}(\boldsymbol{x})$	return (v, s_3)
return $(M, (\boldsymbol{y}', v, s_3))$	

Since S is bounded by time $p'(k)$ and $\mathcal{K}$, $\mathcal{E}_{pk'}$ and $\mathcal{D}_{sk}$ are polynomial-time, it follows that M is samplable in time $p'(k)$ and S' is also polynomial-time. Then, the above construction of S' and R gives that

$$\Pr[\mathsf{Expt}^{\text{SSS}^*\text{-PCAX-0}}_{\Pi,F',S'}(k) : w = 1] \geq \Pr[\mathsf{Expt}^{\text{SNM}^{\circledast}\text{-ATK-0}}_{\Pi,R,S}(k) : w = 1]$$

(where equality holds if and only if S' always fails when $\lvert\boldsymbol{s}\rvert \neq 2$), and so

$$\mathsf{Adv}^{\text{SSS}^*\text{-PCAX}}_{\Pi,F',A',S'}(k) \leq \mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k).$$

Consequently, if Π is secure in the sense of SNM$^{\circledast}$-ATK, then $\mathsf{Adv}^{\text{SNM}^{\circledast}\text{-ATK}}_{\Pi,R,A,S}(k)$ is negligible, and so is $\mathsf{Adv}^{\text{SSS}^*\text{-PCAX}}_{\Pi,F',A',S'}(k)$. This completes the proof of (II).

(III) CSS*-PCAX $\Longrightarrow$ CNM*-ATK: Let p be a polynomial of k. Let $A' = (A'_1, A'_2)$ be a legitimate CNM*-ATK adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A', let us construct an CSS*-PCAX adversary $A = (A_1, A_2, A_3)$ attacking Π as

Algorithm $A_1^{\mathcal{O}_1}(pk)$	Algorithm $A_2^{\mathcal{O}_2}(s_1, y)$	Algorithm $A_3^{\mathcal{O}_2}(R, \boldsymbol{x})$
$(M, s_1) \leftarrow A_1'^{\mathcal{O}_1}(pk)$	$(\boldsymbol{y}, R) \leftarrow A_2'^{\mathcal{O}_2}(s_1, y)$	return $(F_{R,\boldsymbol{x}}, 1)$
return (M, s_1)	return $(\boldsymbol{y}, R)$	

where the function $F_{R,\boldsymbol{x}}$ output by A_3 is given by

Function $F_{R,\boldsymbol{x}}(x)$
return $R(x, \boldsymbol{x})$

Since A' is bounded by time $p(k)$, it follows that M is samplable in time $p(k)$ and A is also polynomial-time. Moreover, since A' is legitimate, A is also legitimate.

We note that A can answer queries from A' by using her own oracle. It is now straightforward to see from the above construction of A that

$$\Pr[\mathsf{Expt}^{\mathrm{CNM^*\text{-}ATK\text{-}1}}_{\Pi,A'}(k) : w = 1] = \Pr[\mathsf{Expt}^{\mathrm{CSS^*\text{-}PCAX\text{-}1}}_{\Pi,A}(k) : w = 1].$$

Consequently, if Π is secure in the sense of CSS*-PCAX, then $\mathsf{Adv}^{\mathrm{CSS^*\text{-}PCAX}}_{\Pi,A}(k)$ is negligible, and so is $\mathsf{Adv}^{\mathrm{CNM^*\text{-}ATK}}_{\Pi,A'}(k)$. This completes the proof of (III).

(IV) CNM*-ATK $\Longrightarrow$ CSS*-PCAX: Let p be a polynomial of k. Let $A' = (A'_1, A'_2, A'_3)$ be a legitimate CSS*-PCAX adversary attacking an encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$, bounded by time $p(k)$. By using A', let us construct an SSS*-ATK adversary $A = (A_1, A_2)$ attacking Π as

Algorithm $A_1^{\mathcal{O}_1}(pk)$	Algorithm $A_2(s_1, y)$
$(M, s_1) \leftarrow A_1'^{\mathcal{O}_1}(pk)$	$(\boldsymbol{y}, s_2) \leftarrow A'_2(s_1, y)$
return (M, s_1)	choose randomness r for A'_3
	return $(\boldsymbol{y}, R_{r,s_2})$

where the relation R_{r,s_2} output by A_2 is given by

Relation $R_{r,s_2}(x, \boldsymbol{x})$
 $(f, v) \leftarrow A'_3(s_2, \boldsymbol{x}; r)$
 if $f(x) = v$ then return 1
 else return 0

Since A' is bounded by time $p(k)$, it follows that M is samplable in time $p(k)$ and A is also polynomial-time. Moreover, since A' is legitimate, A is also legitimate. We note that A can answer queries from A' by using her own oracle. It is now straightforward to see from the above construction of A and R that

$$\mathsf{Adv}^{\mathrm{CSS^*\text{-}PCAX}}_{\Pi,A'}(k) \le \mathsf{Adv}^{\mathrm{CNM^*\text{-}ATK}}_{\Pi,A}(k).$$

Consequently, if Π is secure in the sense of CNM*-ATK, then $\mathsf{Adv}^{\mathrm{CNM^*\text{-}ATK}}_{\Pi,A}(k)$ is negligible, and so is $\mathsf{Adv}^{\mathrm{CSS^*\text{-}PCAX}}_{\Pi,A'}(k)$. This completes the proof of (IV), and the proposition follows. □

C Indistinguishability-based characterization of comparison-based non-malleability

In this appendix, we describe an indistinguishability-based characterization of comparison-based non-malleability, denoted as IND†-PCAX in this paper, which was introduced in [15] for private-key encryption schemes. The difference between IND*-PCAX and IND†-PCAX is that an IND* adversary always fails if $\perp \in \boldsymbol{x}$, while the success of an IND† adversary is determined at random if $\perp \in \boldsymbol{x}$; namely, if $\perp \in \boldsymbol{x}$, then an IND† adversary succeeds with probability $\frac{1}{2}$ and fails with the same probability. A formal definition of IND†-PCAX is described below.

Definition 7 (IND†-PCAX [15]). *Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption scheme and $A = (A_1, A_2)$ be an adversary attacking Π. For $k \in \mathbb{N}$ and* PCAX $\in$ $\{\mathrm{PCA0}, \mathrm{PCA1}, \mathrm{PCA2}\}$*, consider the following experiment:*

Experiment $\mathsf{Expt}_{\Pi,A}^{\mathrm{IND}^\dagger\text{-PCAX}}(k)$
$(pk, sk) \leftarrow \mathcal{K}(1^k)$
$(x_0, x_1, s_1) \leftarrow A_1^{\mathcal{O}_1}(pk)$; $b \leftarrow \{0,1\}_U$; $y \leftarrow \mathcal{E}_{pk}(x_b)$
$(\boldsymbol{y}, s_2) \leftarrow A_2^{\mathcal{O}_2}(x_0, x_1, s_1, y)$; $\boldsymbol{x} \leftarrow \mathcal{D}_{sk}(\boldsymbol{y})$
$d \leftarrow A_3^{\mathcal{O}_2}(\boldsymbol{x}, s_2)$
if $\perp \in \boldsymbol{x}$ then $w \leftarrow \{0,1\}_U$
else if $d = b$ then $w \leftarrow 1$
else $w \leftarrow 0$

Here, A is supposed to be legitimate as in Definition 1, and $\mathcal{O}_1$ and $\mathcal{O}_2$ are defined as in Definition 3. Then, an encryption scheme Π is called secure in the sense of IND†-PCAX *if for all polynomial p and all probabilistic adversary A runnable in time $p(k)$, $\mathsf{Adv}_{\Pi,A}^{IND^\dagger\text{-}PCAX}(k)$ is negligible, where $\mathsf{Adv}_{\Pi,A}^{IND^\dagger\text{-}PCAX}$ denotes the advantage of A defined by*

$$\mathsf{Adv}_{\Pi,A}^{IND^\dagger\text{-}PCAX}(k) = 2\Pr[\mathsf{Expt}_{\Pi,A}^{IND^\dagger\text{-}PCAX}(k) : w = 1] - 1$$

The proof of the equivalence between CNM* and IND† for the private-key setting given in [15] straightforwardly applies to the public-key setting, yielding the following proposition.

Proposition 3. *IND†-PCAX $\iff$ CNM*-ATK for* (PCAX, ATK)$\in \{(\mathrm{PCA0}, \mathrm{CPA}),$ $(\mathrm{PCA1}, \mathrm{CCA1}),\ (\mathrm{PCA2}, \mathrm{CCA2})\}$.