Indistinguishability Amplification

Ueli Maurer
maurer@inf.ethz.ch
ETH Zürich

Krzysztof Pietrzak
pietrzak@di.ens.fr
ENS Paris

Renato Renner
r.renner@damtp.cam.ac.uk
Cambridge

Abstract

A random system is the abstraction of the input-output behavior of any kind of discrete system, in particular cryptographic systems. Many aspects of cryptographic security analyses and proofs can be seen as the proof that a certain random system (e.g. a block cipher) is indistinguishable from an ideal system (e.g. a random permutation), for different types of distinguishers.

This paper presents a new generic approach to proving upper bounds on the distinguishing advantage of a combined system, assuming upper bounds of various types on the component systems. For a general type of combination operation of systems (including the combination of functions or the cascade of permutations), we prove two amplification theorems.

The first is a direct-product theorem, similar in spirit to the XOR-Lemma: The distinguishing advantage (or security) of the combination of two (possibly stateful) systems is twice the product of the individual distinguishing advantages, which is optimal.

The second theorem states that the combination of systems is secure against some strong class of distinguishers, assuming only that the components are secure against some weaker class of attacks. As a corollary we obtain tight bounds on the adaptive security of the cascade and parallel composition of non-adaptively (or only random-query) secure component systems.

A key technical tool of the paper is to show a tight two-way correspondence, previously only known to hold in one direction, between the distinguishing advantage of two systems and the probability of provoking an appropriately defined event on one of the systems.

1 Introduction

1.1 Motivation

A *random system*,[1] the abstraction of the input-output behavior of a system, can be seen as the generalization of a random variable. In contrast to a random variable, which is non-interactive, a random system interacts with its observer and can keep a state.

Many cryptographic systems (e.g. a block cipher or the CBC-MAC construction) can be modeled as random systems, stateful or stateless, and security proofs often amount to proving the indistinguishability of two such systems. While in practice one is mostly interested in *computational* indistinguishability, the core of the proofs is often a proof of *information-theoretic* indistinguishability. For example, Vaudenay's decorrelation theory [Vau98, Vau99, Vau03] is purely information-theoretic, but its application is for the design of actual block ciphers. This paper is concerned with information-theoretic indistinguishability but has also implications for the computational case.

An important theme in cryptography is the *amplification* of security properties of a certain scheme. Examples of amplification results are the XOR-Lemma, Vaudenay's direct-product theorem for random permutations [Vau99] and the "adaptive from non-adaptive security" theorems of [MP04] and [MOPS06]. In this paper we investigate such *indistinguishability amplifications*. In contrast to earlier works, we do not restrict ourselves to *stateless* systems. The term "amplification" is used with two different meanings: either the distinguishing advantage is reduced by the construction, or the allowed type of distinguishers is strengthened. This paper generalizes, strengthens, and unifies the above mentioned results and provides a framework for proving such amplification results.

[1] Throughout the paper, the term "random" is used in the same sense as it is used in the term "random variable", without implying uniformity of a distribution.

1.2 Contributions of this Paper

An important paradigm in indistinguishability proofs is the definition of an internal monotone condition in a random system (sometimes also called a "bad event") such that for any distinguisher $\mathbf{D}$ the distinguishing advantage can be shown to be upper bounded by the probability that $\mathbf{D}$ provokes this bad event. A key technical tool of the paper (Lemma 2) is to show that this holds also in the other direction: for two systems $\mathbf{F}$ and $\mathbf{G}$ one can always define new systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$, which are equivalent to $\mathbf{F}$ and $\mathbf{G}$ respectively, but have an additional monotone binary output (MBO) such that (i) for any distinguisher $\mathbf{D}$ the distinguishing advantage for $\mathbf{F}$ and $\mathbf{G}$ is *equal* to the probability that $\mathbf{D}$ sets the MBO to 1 in $\hat{\mathbf{F}}$ (or $\hat{\mathbf{G}}$) and (ii) the systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ are equivalent as long as the respective MBOs are 0.[2]

This lemma is used to prove the two main theorems of this paper, one for each type of amplification mentioned above. Instead of stating the theorems here in the introduction, we state two corollaries understandable without much notation.

For a class of attacks ATK we denote with $\Delta_q^{\mathsf{ATK}}(\mathbf{F},\mathbf{G})$ the distinguishing advantage of the best ATK-distinguisher making q queries to the system $\mathbf{F}$ or $\mathbf{G}$. The attacks we consider are adaptive and non-adaptive chosen-plaintext attacks ($\mathsf{CPA}, \mathsf{nCPA}$), random queries ($\mathsf{KPA}$) and, if the system at hand is a permutation, also adaptive and non-adaptive chosen-ciphertext attacks (CCA and nCCA).

With $\mathbf{F} \triangleright \mathbf{G}(X) \overset{\text{def}}{=} \mathbf{G}(\mathbf{F}(X))$ we denote the sequential, and with $\mathbf{F} \star \mathbf{G}(X) \overset{\text{def}}{=} \mathbf{F}(X) \star \mathbf{G}(X)$ the parallel composition (where $\star$ stands for any group operation, like XOR). Let $\mathbf{R}$ (resp. $\mathbf{P}$) denote a uniform random function (resp. permutation). See Definition 2 for a definition of (stateful) random functions and permutations.

Corollary 1 (Direct Product)**.** *For* $\mathsf{ATK} \in \{\mathsf{nCPA}, \mathsf{CPA}, \mathsf{nCCA}, \mathsf{CCA}\}$ *and any stateless random permutations* $\mathbf{F}, \mathbf{G}$

$$\Delta_q^{\mathsf{ATK}}(\mathbf{F} \triangleright \mathbf{G}, \mathbf{P}) \leq 2 \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{F}, \mathbf{P}) \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{G}, \mathbf{P}) \tag{1}$$

For $\mathsf{ATK} \in \{\mathsf{nCPA}, \mathsf{CPA}\}$ *and any (possibly stateful) random functions* $\mathbf{F}, \mathbf{G}$

$$\Delta_q^{\mathsf{ATK}}(\mathbf{F} \star \mathbf{G}, \mathbf{R}) \leq 2 \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{F}, \mathbf{R}) \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{G}, \mathbf{R})$$

Corollary 2 (Adaptive Security by Composition)**.** *For any stateless random permutations* $\mathbf{F}$ *and* $\mathbf{G}$

$$\Delta_q^{\mathsf{CPA}}(\mathbf{F} \triangleright \mathbf{G}, \mathbf{P}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P}) + \Delta_q^{\mathsf{KPA}}(\mathbf{G}, \mathbf{P}) \tag{2}$$

$$\Delta_q^{\mathsf{CCA}}(\mathbf{F} \triangleright \mathbf{G}^{-1}, \mathbf{P}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P}) + \Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{P}) \tag{3}$$

For any (possibly stateful) random functions $\mathbf{F}, \mathbf{G}$

$$\Delta_q^{\mathsf{CPA}}(\mathbf{F} \star \mathbf{G}, \mathbf{R}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{R}) + \Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{R})$$

Note that the statements (1)-(3) for sequential composition are only given for *stateless* random permutations. This is necessary, as (1)-(3) are wrong[3] in the stateful case.[4]

1.3 Related Work and Applications

Lemma 2 from this paper improves on Lemma 9 of [MP04] where a relation between distinguishing advantage and monotone binary outputs (there called conditions) was introduced, but which "lost" a logarithmic factor and whose proof was quite technical based on martingales. We had to leave the

[2] We give another interpretation of Lemma 2. If two random variables X and X' have statistical distance δ, this can be interpreted as "X and X' are equal with probability $1-\delta$." More precisely, there exists a (common) random experiment with (at least) two random variables X and X' with correct marginal distributions such that $P(X = X') = 1 - \delta$. Lemma 2 can be interpreted as the generalization of this statement to random systems. For any distinguisher $\mathbf{D}$, two random systems $\mathbf{F}$ and $\mathbf{G}$ are equal with probability $1-\delta$, where δ is $\mathbf{D}$'s distinguishing advantage.

[3] Actually, in (1) and (2) only $\mathbf{G}$, but not $\mathbf{F}$ must be stateless, whereas (3) is wrong if either $\mathbf{F}$ or $\mathbf{G}$ can have state.

[4] The step where statlessness is required in the proof of (1) is Proposition 2 from this paper. For (2) and (3) statelessness is used in the proof of the Corollary 2 given in Appendix B.

question whether one can avoid this loss as an open problem in [MP04], Lemma 2 answers it in the affirmative.

ADAPTIVE SECURITY BY COMPOSITION TYPE PROBLEMS. Our stronger result immediately applies to all results that made use of the bounds from [MP04] (where we gave a weaker version of Corollary 2). For example for [KNR05], where Kaplan, Naor and Reingold consider the problem of randomness-efficient constructions of almost k-wise independent permutations. Their construction a priori only achieves non-adaptive security, but they observe that one can apply the results from [MP04] in order to get adaptive security. Another application of Corollary 2 is in the already mentioned decorrelation theory [Vau98, Vau99, Vau03], where it implies better security against adaptive attacks of some special form, even if the block-cipher considered only satisfies a non-adaptive notion of decorrelation.

The question whether composition implies adaptive security also in the computation setting (i.e. for pseudorandom systems) has been investigated in [Mye04, Pie05]. Unlike for the direct product results, this type of results do not to hold for pseudorandom systems in general, though some positive results have also been achieved in this setting [Pie06].

DIRECT PRODUCT PROBLEMS. The direct product property for sequential composition of stateless permutations, i.e. eq.(1), was proved earlier by Vaudenay within his *decorrelation* framework (see [Vau98] for the non-adaptive and [Vau99] for the adaptive case). Vaudenay's proofs — which use matrix norms — are tailored to the construction and attack at hand (i.e. sequential composition), and do not extend to our general setting.

In the computational setting, where one considers polynomially time-bounded adversaries, a direct product theorem for the sequential composition of permutations was proven by Luby and Rackoff [LR86]. Myers [Mye03] proves a direct product theorem[5] for a construction which is basically parallel composition but with some extra random values XOR-ed to the inputs.

SUBSEQUENT WORK. Theorem 2 can be used to prove the adaptive security of more complicated constructions than the sequential and parallel composition considered in Corollary 2. In [MOPS06] Theorem 2 is used to prove that the four round Feistel network with non-adaptively secure round functions is adaptively secure (that paper also shows that in the computational setting this is no longer true).

A result using Lemma 2 of a completely different vain than the problems considered in this paper is given in [PS06], where the security of some constructions for range extension of KPA-secure functions is proven in the information theoretic setting (again, in the computational setting those results no longer hold).

2 Random Systems Basics

2.1 Notation and Basic Definitions

NOTATION FOR SETS AND RANDOM VARIABLES. Capital calligraphic letters (e.g. $\mathcal{X}$) denote sets and the corresponding capital letter (e.g. X) denotes a random variable taking values in the set. Concrete values for X are usually denoted by the corresponding small letter x. A list $[X_1, \ldots, X_i]$ of random variables is often denoted as X^i.

NOTATION FOR PROBABILITIES. $\mathsf{P}[\alpha]$ denotes the probability of the event α. The distribution of a random variable V is denoted P_V, we use $\mathsf{P}_V[v] \stackrel{\text{def}}{=} \mathsf{P}[V = v]$, similarly for conditional probabilities $\mathsf{P}_{V|W}[v, w] \stackrel{\text{def}}{=} \mathsf{P}[V = v|W = w]$.

Because we will consider different random experiments where the same (names for) random variables appear, we sometimes explicitly write the random experiment $\mathcal{E}$ considered as a superscript, e.g. $\mathsf{P}^{\mathcal{E}}_V[v]$. Equivalence of distributions means equivalence on all inputs, i.e.

$$\mathsf{P}^{\mathcal{E}_1}_{V|W} = \mathsf{P}^{\mathcal{E}_2}_{V|W} \iff \forall v \in \mathcal{V}, w \in \mathcal{W} : \ \mathsf{P}^{\mathcal{E}_1}_{V|W}[v, w] = \mathsf{P}^{\mathcal{E}_2}_{V|W}[v, w]$$

To avoid confusion, we use $\mathsf{p}^{\mathcal{E}}$ instead of $\mathsf{P}^{\mathcal{E}}$ if $\mathcal{E}$ is not fully defined random experiment but only a conditional distribution.

[5] Which in some sense is stronger than the amplification from [LR86], see [Mye03] for a discussion.

2.2 Random Systems

A *random system*, an abstraction introduced in [Mau02], is a system which takes inputs $X_1, X_2, \ldots$ and generates, for each new input X_i, an output Y_i which depends probabilistically on the inputs and outputs seen so far. We define random systems in terms of the distribution of the outputs Y_i conditioned on $X^i Y^{i-1}$ (i.e. the actual query X_i and all previous input/output pairs $X_1 Y_1, \ldots, X_{i-1} Y_{i-1}$). For a random system $\mathbf{F}$, this distribution is given by:

$$\mathsf{p}^{\mathbf{F}}_{Y_i|X_i Y^{i-1}}(y_i, x^i, y^i) \overset{\text{def}}{=} \mathsf{P}[\mathbf{F}(x_i) = y_i | \forall 1 \leq j < i : \mathbf{F}(x_j) = y_j]$$

Definition 1. *A* $(\mathcal{X}, \mathcal{Y})$-random system $\mathbf{F}$ *is a (possibly infinite) sequence of conditional probability distributions* $\mathsf{p}^{\mathbf{F}}_{Y_i|X^i Y^{i-1}}$ *for* $i \geq 1$*. Two random systems* $\mathbf{F}$ *and* $\mathbf{G}$ *are* equivalent*, denoted* $\mathbf{F} \equiv \mathbf{G}$*, if*

$$\mathsf{p}^{\mathbf{F}}_{Y_i|X^i Y^{i-1}} = \mathsf{p}^{\mathbf{G}}_{Y_i|X^i Y^{i-1}} \quad \textit{for all} \quad i \geq 1$$

Definition 2 (Random function/permutation)**.** *A* stateless random function $\mathcal{X} \to \mathcal{Y}$ *(*random permutation *on* $\mathcal{X}$*) is a random variable which takes as values functions* $\mathcal{X} \to \mathcal{Y}$ *(permutations* $\mathcal{X} \to \mathcal{Y}$ *where* $\mathcal{X} \equiv \mathcal{Y}$*). Throughout, the symbols* $\mathcal{R}$ *and* $\mathcal{P}$ *are used for the set of all random functions and the set of all random permutations respectively (*$\mathcal{X}, \mathcal{Y}$ *to be understood). Note that* $\mathcal{P} \subset \mathcal{R}$*.*

A uniform random function (URF) $\mathbf{R} : \mathcal{X} \to \mathcal{Y}$ *(A* uniform random permutation (URP) $\mathbf{P}$ *on* $\mathcal{X}$*) is a random function with uniform distribution over all functions from* $\mathcal{X}$ *to* $\mathcal{Y}$ *(permutations on* $\mathcal{X}$*). Throughout, the symbols* $\mathbf{R}$ *and* $\mathbf{P}$ *are used for the systems defined above.*

A stateful *random function* $\mathcal{X} \to \mathcal{Y}$ *is a* $(\mathcal{X}, \mathcal{Y})$*-random system which is consistent, i.e. it outputs the same value when queried on the same value twice. A* stateful *random permutation is defined similarly, but has the additional property of being a permutation. Throughout,* $\mathcal{R}_S$ *and* $\mathcal{P}_S$ *are used for the set of all stateful random functions and the set of all stateful random permutations respectively. Note that* $\mathcal{P} \subset \mathcal{P}_S, \mathcal{R} \subset \mathcal{R}_S$ *and* $\mathcal{P}_S \subset \mathcal{R}_S$*.*

Definition 3 $(\triangleright, \star)$**.** *The* sequential composition *of two (or more) random systems* $\mathbf{F}$ *and* $\mathbf{G}$*, denoted* $\mathbf{F} \triangleright \mathbf{G}$*, is defined naturally: The output of* $\mathbf{F}$ *is connected to the input of* $\mathbf{G}$*. For functions this corresponds to function composition.*

Similarly, the parallel composition *of* $\mathbf{F}$ *and* $\mathbf{G}$ *for some group operation* $\star$ *on the output alphabet, denoted* $\mathbf{F} \star \mathbf{G}$*, is the system obtained by feeding the input to both* $\mathbf{F}$ *and* $\mathbf{G}$ *and combining the outputs using* $\star$*.*

2.3 Indistinguishability of Random Systems

The statistical (or variational) distance of two random variables X and X' over $\mathcal{X}$ is defined as

$$\|P_X - P_{X'}\| := \frac{1}{2} \sum_{x \in \mathcal{X}} |P_X(x) - P_{X'}(x)| .$$

It is equal to the advantage of the best distinguisher for X and X'. The generalization to random systems is more involved as one must explicitly introduce a distinguisher $\mathbf{D}$ to be able to define how (in)distinguishable two systems are.

Definition 4 (Distinguisher)**.** *A* $(\mathcal{Y}, \mathcal{X})$-distinguisher *is a* $(\mathcal{Y}, \mathcal{X})$*-random system which is one query ahead, meaning that it is defined by* $\mathsf{p}^{\mathbf{D}}_{X_i|Y^{i-1} X^{i-1}}$ *(instead of* $\mathsf{p}^{\mathbf{D}}_{X_i|Y^i X^{i-1}}$*) for all i. In particular the first output* $\mathsf{p}^{\mathbf{D}}_{X_1}$ *is defined before* $\mathbf{D}$ *is fed with any input.*

The random experiment when a $(\mathcal{Y}, \mathcal{X})$*-distinguisher* $\mathbf{D}$ *interactively queries a* $(\mathcal{X}, \mathcal{Y})$*-random system* $\mathbf{F}$ *is denoted as*

$$\mathbf{D} \diamond \mathbf{F}$$

For the proof of the direct product theorem it is convenient to define $(\mathcal{Y}, \mathcal{X})$-double-distinguishers *which have access to two* $(\mathcal{X}, \mathcal{Y})$*-random systems, and which can arbitrary schedule their queries between the two systems. For a* $(\mathcal{Y}, \mathcal{X})$*-double-distinguisher* $\mathbf{D}$ *we denote with* $\mathbf{D} \diamond [\mathbf{F}, \mathbf{G}]$ *the random experiment where* $\mathbf{D}$

queries the two $(\mathcal{X},\mathcal{Y})$-random systems **F** *and* **G**. *Note that if we instantiate just one of the two systems, we get a standard distinguisher, i.e.* $\mathbf{D}\diamond[\,\cdot\,,\mathbf{F}]$ *and* $\mathbf{D}\diamond[\mathbf{F},\,\cdot\,]$ *is a $(\mathcal{Y},\mathcal{X})$-distinguisher for any compatible* **F**.

One can distinguish different classes of distinguishers by posing restrictions on how the distinguisher can access the system. In particular the following attacks will be of interest to us:

- CPA (Adaptively Chosen Plaintext Attack): This is the most general attack. Here the distinguisher can make any first query X_1 and receives the output Y_1, then it chooses a query X_2 depending on X_1 and Y_1 and receives Y_2, and so on. In general, he can choose the ith query X_i depending on X^{i-1} and Y^{i-1}.

- nCPA (Non-Adaptively Chosen Plaintext Attack): The distinguisher must choose all queries in advance.

- KPA (Known Plaintext Attack): The distinguisher obtains only distinct random inputs (i.e., the choice of input is beyond its control) to the system and the corresponding outputs.

If **F** is a permutation, then also its inverse $\mathbf{F}^{-1}$ is well defined. So in the case where the system queried is guaranteed to be a permutation, we can consider an even more powerful attack where the distinguisher can query the system at hand from both directions.

- CCA (Adaptively Chosen Ciphertext Attack): Is defined as CPA but the distinguisher can query from both sides.

- nCCA (Non-Adaptively Chosen Ciphertext Attack): Non adaptive version of CCA.

Definition 5 (ATK_2)**.** *Let* ATK *be one of the classes of distinguishers considered above, then a double distinguisher* **D** *(see Definition 4) is in the class* ATK_2 *if for any* **F** *the systems* $\mathbf{D}\diamond[\mathbf{F},\,\cdot\,]$ *and* $\mathbf{D}\diamond[\,\cdot\,,\mathbf{F}]$ *are* ATK *distinguishers.*

For given $k \geq 1$, the two random experiments $\mathbf{D}\diamond\mathbf{F}$ and $\mathbf{D}\diamond\mathbf{G}$ each define a transcript X^kY^k, a random variable with alphabet $\mathcal{X}^k \times \mathcal{Y}^k$.

Definition 6 (Δ)**.** *For $k \geq 1$, the* advantage *of* **D** *after k queries in distinguishing* **F** *from* **G**, *denoted* $\Delta_k^{\mathbf{D}}(\mathbf{F},\mathbf{G})$, *is the statistical difference between the transcripts.*[6]

$$\Delta_k^{\mathbf{D}}(\mathbf{F},\mathbf{G}) \overset{def}{=} \|\mathsf{P}^{\mathbf{D}\diamond\mathbf{F}}_{X^kY^k} - \mathsf{P}^{\mathbf{D}\diamond\mathbf{G}}_{X^kY^k}\| \tag{4}$$

The advantage of the best ATK*-distinguisher making k queries for* **F** *and* **G** *is*

$$\Delta_k^{\mathsf{ATK}}(\mathbf{F},\mathbf{G}) \overset{def}{=} \max_{\mathbf{D}\in\mathsf{ATK}} \Delta_k^{\mathbf{D}}(\mathbf{F},\mathbf{G})$$

2.4 Monotone Binary Outputs for Random Systems

In the realization of a random system (e.g., the Luby-Rackoff construction of a random permutation from three random functions), it is often convenient to consider a certain condition that can hold within the system, for example that all inputs to an internal component are distinct (no collision). Such conditions are very useful for proving the indistinguishability of random systems [Mau02]. In this paper we also consider conditions for random systems, but the conditions considered here do usually not have a simple interpretation (like there was a collision). Rather, the conditions are defined such, that the system at hand is equivalent to some other system while the conditions holds.

We only only consider *monotone* conditions, which once they fail to be satisfied, they remain so. We model such monotone conditions by considering random systems with a monotone binary output

[6] This definition is equivalent to asking the distinguisher **D** to make a binary decision, and also equivalent to requiring that the distinguisher guesses correctly with which system it interacts, both being equally likely.

(MBO). Monotonicity means that the binary output is initially 0 and may eventually turn to 1. For a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-random system with an MBO $\mathbf{S}$ it is useful to consider $\mathbf{S}^{\rightarrow}$ which denotes the projection of $\mathbf{S}$ to the $\mathcal{Y}$-output (i.e., ignoring the binary output). It is also useful to define a system $\mathbf{S}^{\dashv}$ which is equivalent to $\mathbf{S}$, but where the $\mathcal{Y}$-output is blinded (i.e. set to some dummy value $\perp$) when the binary output turns to 1.

Definition 7. *A random-system with a monotone binary output (MBO)* $\mathbf{S}$ *is a* $(\mathcal{X}, \mathcal{Y} \times \mathcal{A})$*-random system where* $\mathcal{A} = \{0,1\}$ *and the MBO satisfies* $A_i = 1 \Rightarrow A_{i+1} = 1$ *(i.e. is monotone). From such a* $\mathbf{S}$ *we get the* $(\mathcal{X}, \mathcal{Y})$*-random system* $\mathbf{S}^{\rightarrow}$ *and the* $(\mathcal{X}, \{\mathcal{Y} \cup \perp\} \times \mathcal{A})$*-random system* $\mathbf{S}^{\dashv}$ *as follows.*

- $\mathbf{S}^{\rightarrow}$ *is* $\mathbf{S}$ *with the following function applied to the output:* $(y, b) \rightarrow y$*.*
- $\mathbf{S}^{\dashv}$ *is* $\mathbf{S}$ *with the following function applied to the output:*

$$(y,b) \mapsto (y', b) \quad \textit{where} \quad y' = \begin{cases} y & \textit{if } b = 0 \\ \perp & \textit{if } b = 1 \end{cases}$$

We will always use the letters $\mathbf{S}$ and $\mathbf{T}$ for random systems with an MBO, and $\mathbf{F}, \mathbf{G}$ for random systems without an MBO. Also the special systems $\mathbf{D}$ (always a distinguisher), $\mathbf{R}, \mathbf{P}$ (URF and URP) and $\mathbf{I}$ ("ideal system" or identity function) do not have MBOs.

We will often start with a random system without an MBO, and add an MBO to it. We then add a " ˆ " to denote the derived system. For example $\hat{\mathbf{F}}$ will always denote a random system with an MBO which satisfies $\hat{\mathbf{F}}^{\rightarrow} \equiv \mathbf{F}$.

Definition 8. *For a* $(\mathcal{X}, \mathcal{Y} \times \mathcal{A})$*-random sytem* $\mathbf{S}$ *with an MBO, we denote with* $\nu_k^{\mathbf{D}}(\mathbf{S})$ *the advantage of a* $(\mathcal{Y}, \mathcal{X})$*-distinguisher* $\mathbf{D}$ *in setting the MBO to* 1 *with* k *queries and by* $\nu_k^{\mathsf{ATK}}(\mathbf{S})$ *the advantage of the best* ATK *distinguisher*[7]

$$\nu_k^{\mathbf{D}}(\mathbf{S}) = \mathsf{P}^{\mathbf{D}\diamond\mathbf{S}}[A_k = 1] \qquad \textit{and} \qquad \nu_k^{\mathsf{ATK}}(\mathbf{S}) = \max_{\mathbf{D}\in\mathsf{ATK}} \nu_k^{\mathbf{D}}(\mathbf{S}) \tag{5}$$

For a double-distinguisher $\mathbf{D}$*, we denote with* $\nu_k^{\mathbf{D}}(\mathbf{S}, \mathbf{T})$ *the advantage of* $\mathbf{D}$ *in setting both MBOs to* 1 *making* k *queries to each system respectively*

$$\nu_k^{\mathbf{D}}(\mathbf{S}, \mathbf{T}) = \mathsf{P}^{\mathbf{D}\diamond[\mathbf{S},\mathbf{T}]}[A_k = 1 \wedge B_k = 1] \qquad \textit{and} \qquad \nu_k^{\mathsf{ATK}_2}(\mathbf{S}, \mathbf{T}) = \max_{\mathbf{D}\in\mathsf{ATK}_2} \nu_k^{\mathbf{D}}(\mathbf{S}, \mathbf{T})$$

where $A_1, A_2, \ldots$ *and* $B_1, B_2, \ldots$ *denote the MBO of* $\mathbf{S}$ *and* $\mathbf{T}$ *respectively.*

If two random systems $\mathbf{S}$ and $\mathbf{T}$ with MBOs are equivalent while the MBOs are 0, then clearly setting the MBO to 1 is equally difficult in both, i.e.

$$\mathbf{S}^{\dashv} \equiv \mathbf{T}^{\dashv} \Rightarrow \forall\, \mathbf{D}, k \;:\; \nu_k^{\mathbf{D}}(\mathbf{S}) = \nu_k^{\mathbf{D}}(\mathbf{T}). \tag{6}$$

By the following lemma from [Mau02] this probability is also an upper bound on the distinguishing advantage of $\mathbf{D}$ for the two systems.

Lemma 1. *If* $\mathbf{S}^{\dashv} \equiv \mathbf{T}^{\dashv}$*, then for any distinguisher* $\mathbf{D}$ *and any* $k \in \mathbb{N}$

$$\Delta_k^{\mathbf{D}}(\mathbf{S}^{\rightarrow}, \mathbf{T}^{\rightarrow}) \le \nu_k^{\mathbf{D}}(\mathbf{S}) = \nu_k^{\mathbf{D}}(\mathbf{T}). \tag{7}$$

Proof. Let $A_1, A_2, \ldots$ and $B_1, B_2, \ldots$ denote the MBO of $\mathbf{S}$ and $\mathbf{T}$ respectively

$$\begin{aligned}
\Delta_k^{\mathbf{D}}(\mathbf{S}^{\rightarrow}, \mathbf{T}^{\rightarrow}) &= \frac{1}{2} \sum_{\mathcal{X}^k, \mathcal{Y}^k} \left| \mathsf{P}^{\mathbf{D}\diamond\mathbf{S}}_{X^kY^k} - \mathsf{P}^{\mathbf{D}\diamond\mathbf{T}}_{X^kY^k} \right| \\
&\le \underbrace{\frac{1}{2} \sum_{\mathcal{X}^k, \mathcal{Y}^k} \left| \mathsf{P}^{\mathbf{D}\diamond\mathbf{S}}_{A_k=0X^kY^k} - \mathsf{P}^{\mathbf{D}\diamond\mathbf{T}}_{B_k=0X^kY^k} \right|}_{=0 \text{ as } \mathbf{S}^{\dashv}\equiv\mathbf{T}^{\dashv}} + \frac{1}{2} \sum_{\mathcal{X}^k, \mathcal{Y}^k} \left| \mathsf{P}^{\mathbf{D}\diamond\mathbf{S}}_{A_k=1X^kY^k} - \mathsf{P}^{\mathbf{D}\diamond\mathbf{T}}_{B_k=1X^kY^k} \right| \\
&\le \frac{1}{2} \left(\mathsf{P}^{\mathbf{D}\diamond\mathbf{S}}_{A_k=1} + \mathsf{P}^{\mathbf{D}\diamond\mathbf{T}}_{B_k=1} \right) = \frac{1}{2} \left(\nu_k^{\mathbf{D}}(\mathbf{S}) + \nu_k^{\mathbf{D}}(\mathbf{T}) \right) = \nu_k^{\mathbf{D}}(\mathbf{S}) = \nu_k^{\mathbf{D}}(\mathbf{T}).
\end{aligned}$$

□

[7] Note that the random experiment $\mathbf{D}\diamond\mathbf{S}$ in (5) is not well defined, as the domain of $\mathbf{D}$ is $\mathcal{Y}$ but $\mathbf{S}$ has range $\mathcal{Y} \times \mathcal{A}$. The meaning here is that $\mathbf{D}$ does not get the MBO $A_1, A_2, \ldots$ as input.

By the above lemma, if we add MBOs to random systems $\mathbf{F}$ and $\mathbf{G}$ in order to get random systems with MBOs which satisfy $\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{G}}^{\dashv}$, then for any $k \in \mathbb{N}$

$$\Delta_k^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) \leq \nu_k^{\mathbf{D}}(\hat{\mathbf{F}}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{G}}). \tag{8}$$

If we compose two random systems with MBOs into a new one using some construction like $\triangleright$ or $\star$, then we can naturally define a new MBO on the composed system as a (monotone and binary valued) function of the MBOs of the components; We will only use the logical AND ($\wedge$) and OR ($\vee$). The function used is written over the composition operator. For example, let $\mathbf{S}$ and $\mathbf{T}$ be random systems with MBOs, then $\mathbf{S} \overset{\wedge}{\triangleright} \mathbf{T}$ denotes the random system $\mathbf{S}^{\rightharpoondown} \triangleright \mathbf{T}^{\rightharpoondown}$ with an MBO which is 1 if the MBO of $\mathbf{S}$ AND the MBO of $\mathbf{T}$ is one.

If we combine a random system with an MBO $\mathbf{S}$ with a random system $\mathbf{F}$ (without an MBO), then the combined system is the combination of $\mathbf{S}^{\rightharpoondown}$ with $\mathbf{F}$, and this system has an MBO which is simply the MBO of $\mathbf{S}$. For example $\mathbf{S} \triangleright \mathbf{F}$ denotes the random system $\mathbf{S}^{\rightharpoondown} \triangleright \mathbf{F}$ with an MBO which is the MBO of the component $\mathbf{S}$.

3 Technical Lemmata

The following lemma considers the other direction of Lemma 1 and states that to any random systems $\mathbf{F}$ and $\mathbf{G}$, we can add conditions in order to get $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$, such that one can achieve equality in (8).

Lemma 2 (Indistinguishability vs. Conditions)**.** *Let $\mathbf{F}$ and $\mathbf{G}$ be $(\mathcal{X}, \mathcal{Y})$-random systems. Then there exit systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ with MBOs such that $\hat{\mathbf{F}}^{\rightharpoondown} \equiv \mathbf{F}$ and $\hat{\mathbf{G}}^{\rightharpoondown} \equiv \mathbf{G}$,*

$$\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{G}}^{\dashv} ,$$

and, for any distinguisher $\mathbf{D}$ and any $k \in \mathbb{N}$,

$$\Delta_k^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{F}}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{G}}) .$$

Proof. We first give an explicit construction of the random systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ (see (11) below). Then, in a second part of the proof, show that these systems have all the desired properties.

To simplify the formulas, we introduce the following abbreviations which we use whenever the choice of inputs $x_1, x_2, \ldots$ is clear from the context. For any $k \in \mathbb{N}_0$, $x^k := (x_1, \ldots, x_k) \in \mathcal{X}^k$, $y^k = (y_1, \ldots, y_k) \in \mathcal{Y}^k$, and $\mathbf{H} \in \{\mathbf{F}, \mathbf{G}\}$, let

$$\begin{aligned} \mathsf{p}_{y^k}^{\mathbf{H}} &:= \begin{cases} \mathsf{p}_{Y^k|X^k=x^k}^{\mathbf{H}}(y^k) & \text{if } k \geq 1 \\ 1 & \text{otherwise} \end{cases} \\ \alpha_{y^k} &:= \min(\mathsf{p}_{y^k}^{\mathbf{F}}, \mathsf{p}_{y^k}^{\mathbf{G}}) \ . \end{aligned}$$

Note that, for any $k \geq 1$, fixed inputs $x^k = (x_1, \ldots, x_k)$, and fixed outputs $y^{k-1} = (y_1, \ldots, y_{k-1})$, we have $\sum_{y_k} \mathsf{p}_{y^k}^{\mathbf{H}} = \mathsf{p}_{y^{k-1}}^{\mathbf{H}}$, where the sum ranges over the last element y_k of the k-tuple y^k. This implies

$$\sum_{y_k} \alpha_{y^k} = \sum_{y_k} \min(\mathsf{p}_{y^k}^{\mathbf{F}}, \mathsf{p}_{y^k}^{\mathbf{G}}) \leq \min(\sum_{y_k} \mathsf{p}_{y^k}^{\mathbf{F}}, \sum_{y_k} \mathsf{p}_{y^k}^{\mathbf{G}}) = \alpha_{y^{k-1}} \ , \tag{9}$$

and, hence, $\sum_{y_k} \left(\mathsf{p}_{y^k}^{\mathbf{H}} - \alpha_{y^k}\right) \geq \mathsf{p}_{y^{k-1}}^{\mathbf{H}} - \alpha_{y^{k-1}}$. Consequently, the probability $\mathsf{p}_{y^{k-1}}^{\mathbf{H}} - \alpha_{y^{k-1}}$ can be split into nonnegative values $\gamma_{y^k}^{\mathbf{H}} \leq \mathsf{p}_{y^k}^{\mathbf{H}} - \alpha_{y^k}$, i.e.,[8]

$$\sum_{y_k} \gamma_{y^k}^{\mathbf{H}} = \mathsf{p}_{y^{k-1}}^{\mathbf{H}} - \alpha_{y^{k-1}} \ . \tag{10}$$

[8] Note that the choice of the probabilities γ_{y^k} is generally not unique.

The random systems with MBO, $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$, are now constructed such that, for any given input tuple $X^k = x^k$, the event that the output Y^k equals y^k *and* the MBO B_k equals 0 has probability α_{y^k}. That is, formally, for $\mathbf{H} \in \{\mathbf{F}, \mathbf{G}\}$, the constructed random system $\hat{\mathbf{H}}$ should satisfy equation (12) below. This is achieved with the definition

$$\mathsf{p}^{\hat{\mathbf{H}}}_{Y_k B_k | X^k = x^k, Y^{k-1} = y^{k-1}, B^{k-1} = b^{k-1}}(y_k, b_k) := \begin{cases} \frac{\alpha_{y^k}}{\alpha_{y^{k-1}}} & \text{if } b^{k-1} = (0, \ldots, 0), b_k = 0 \\ \frac{\mathsf{p}^{\mathbf{H}}_{y^k} - \alpha_{y^k} - \gamma^{\mathbf{H}}_{y^k}}{\alpha_{y^{k-1}}} & \text{if } b^{k-1} = (0, \ldots, 0), b_k = 1 \\ \frac{\gamma^{\mathbf{H}}_{y^k}}{\mathsf{p}^{\mathbf{H}}_{y^{k-1}} - \alpha_{y^{k-1}}} & \text{if } b^{k-1} \neq (0, \ldots, 0), b_k = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

for any $k \in \mathbb{N}$, $x^k \in \mathcal{X}^k$, $y^k \in \mathcal{Y}^k$, and $b^k \in \{0, 1\}^k$ (with the convention $\frac{0}{0} = 0$). It follows directly from the definition of the values $\gamma^{\mathbf{H}}_{y^k}$ that all quantities on the r.h.s. are nonnegative. Moreover, because of (10) and $\sum_{y_k} \alpha_{y^k} + \sum_{y_k} \big(\mathsf{p}^{\mathbf{H}}_{y^k} - \alpha_{y^k} - \gamma^{\mathbf{H}}_{y^k}\big) = \alpha_{y^{k-1}}$, the conditional probabilities sum up to one, i.e., the conditional distribution is well defined. It is also easy to check that the binary output is indeed monotone.

We now turn to the second part of the proof, where we verify that the random systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{G}}$ defined by (11) satisfy the properties stated in the lemma. For this, it is convenient to consider the conditional probability distributions $\mathsf{p}^{\hat{\mathbf{F}}}_{Y^k B^k | X^k = x^k}$ and $\mathsf{p}^{\hat{\mathbf{G}}}_{Y^k B^k | X^k = x^k}$ defined by these systems. By induction over $k \in \mathbb{N}$, it is easy to see that, for $\mathbf{H} \in \{\mathbf{F}, \mathbf{G}\}$,

$$\mathsf{p}^{\hat{\mathbf{H}}}_{Y^k B_k | X^k = x^k}(y^k, 0) = \alpha_{y^k} \tag{12}$$

$$\mathsf{p}^{\hat{\mathbf{H}}}_{Y^k B_k | X^k = x^k}(y^k, 1) = \mathsf{p}^{\mathbf{H}}_{y^k} - \alpha_{y^k} \ . \tag{13}$$

To check that $\hat{\mathbf{F}}^{\rightarrow} = \mathbf{F}$ and $\hat{\mathbf{G}}^{\rightarrow} = \mathbf{G}$ holds, it suffices to sum up the quantities (12)–(13) which gives

$$\mathsf{p}^{\hat{\mathbf{H}}}_{Y^k | X^k = x^k}(y^k) = \mathsf{p}^{\mathbf{H}}_{Y^k | X^k = x^k}(y^k) \ ,$$

for any $k \in \mathbb{N}$, $x^k \in \mathcal{X}^k$ and $y^k \in \mathcal{Y}^k$.

Furthermore, for the property $\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{G}}^{\dashv}$, it suffices to verify that

$$\mathsf{p}^{\hat{\mathbf{F}}}_{Y^k B_k | X^k = x^k}(y^k, 0) = \mathsf{p}^{\hat{\mathbf{G}}}_{Y^k B_k | X^k = x^k}(y^k, 0)$$

holds for any $k \in \mathbb{N}$. This is a direct consequence of (12).

Finally, we need to prove that, for any distinguisher $\mathbf{D}$, the distinguishing advantage $\Delta^{\mathbf{D}}_k(\mathbf{F}, \mathbf{G})$ after k steps equals the probability $\nu^{\mathbf{D}}_k(\hat{\mathbf{F}}) = \nu^{\mathbf{D}}_k(\hat{\mathbf{G}})$ that the MBO of $\hat{\mathbf{F}}$ or $\hat{\mathbf{G}}$ equals 1. Recall that $\Delta^{\mathbf{D}}_k(\mathbf{F}, \mathbf{G})$ is given by the statistical distance $\|\mathsf{P}^{\mathbf{D} \diamond \mathbf{F}}_{X^k Y^k} - \mathsf{P}^{\mathbf{D} \diamond \mathbf{G}}_{X^k Y^k}\|$, where, for $\mathbf{H} \in \{\mathbf{F}, \mathbf{G}\}$, the distribution $\mathsf{P}^{\mathbf{D} \diamond \mathbf{H}}_{X^n Y^n}$ is defined by

$$\begin{aligned} \mathsf{P}^{\mathbf{D} \diamond \mathbf{H}}_{X^k Y^k}(x^k, y^k) &= \prod_{i=1}^{k} \mathsf{P}^{\mathbf{D}}_{X_i | X^{i-1} = x^{i-1}, Y^{i-1} = y^{i-1}}(x_i) \cdot \mathsf{P}^{\mathbf{H}}_{Y_i | X^i = x^i, Y^{i-1} = y^{i-1}}(y_i) \\ &= \mathsf{p}^{\mathbf{D}}_{X^k | Y^{k-1} = y^{k-1}}(x^k) \cdot \mathsf{p}^{\mathbf{H}}_{Y^k | X^k = x^k}(y^k) \ , \end{aligned}$$

for $x^k \in \mathcal{X}^k$ and $y^k \in \mathcal{Y}^k$, where $\mathsf{p}^{\mathbf{D}}_{X^k | Y^{k-1} = y^{k-1}}(x^k) := \prod_{i=1}^{k} \mathsf{P}^{\mathbf{D}}_{X_i | X^{i-1} = x^{i-1}, Y^{i-1} = y^{i-1}}(x_i)$. Because the statistical distance between two distributions P_Z and $P_{Z'}$ can be written as $\|P_Z - P_{Z'}\| = 1 - \sum_z \min(P_Z(z), P_{Z'}(z))$, we find

$$\begin{aligned} \Delta^{\mathbf{D}}_k(\mathbf{F}, \mathbf{G}) = \|\mathsf{P}^{\mathbf{D} \diamond \mathbf{F}}_{X^k Y^k} - \mathsf{P}^{\mathbf{D} \diamond \mathbf{G}}_{X^k Y^k}\| &= 1 - \sum_{x^k, y^k} \min\big(\mathsf{P}^{\mathbf{D} \diamond \mathbf{F}}_{X^k Y^k}(x^k, y^k), \mathsf{P}^{\mathbf{D} \diamond \mathbf{G}}_{X^k Y^k}(x^k, y^k)\big) \\ &= 1 - \sum_{x^k, y^k} \mathsf{p}^{\mathbf{D}}_{X^k | Y^{k-1} = y^{k-1}}(x^k) \cdot \min\big(\mathsf{p}^{\mathbf{F}}_{Y^k | X^k = x^k}(y^k), \mathsf{p}^{\mathbf{G}}_{Y^k | X^k = x^k}(y^k)\big) \ . \end{aligned} \tag{14}$$

Furthermore, the probability that the MBO B_k of the system $\hat{\mathbf{H}}$, for $\mathbf{H} \in \{\mathbf{F}, \mathbf{G}\}$, equals 1 after k steps is given by

$$\nu_k^{\mathbf{D}}(\hat{\mathbf{H}}) = \mathsf{P}_{B_k}^{\mathbf{D}\diamond\hat{\mathbf{H}}}(1) = 1 - \sum_{x^k, y^k} \mathsf{P}_{X^kY^kB_k}^{\mathbf{D}\diamond\hat{\mathbf{H}}}(x^k, y^k, 0)$$

$$= 1 - \sum_{x^k, y^k} \mathsf{p}_{X^k|Y^{k-1}=y^{k-1}}^{\mathbf{D}}(x^k) \cdot \mathsf{p}_{Y^kB_k|X^k=x^k}^{\hat{\mathbf{H}}}(y^k, 0) \ . \quad (15)$$

Recall that we need to show that the quantity in (14) equals the quantity in (15). It thus suffices to verify that, for any $x^k \in \mathcal{X}^k$ and $y^k \in \mathcal{Y}^k$,

$$\min\left(\mathsf{p}_{Y^k|X^k=x^k}^{\mathbf{F}}(y^k), \mathsf{p}_{Y^k|X^k=x^k}^{\mathbf{G}}(y^k)\right) = \mathsf{p}_{Y^kB_k|X^k=x^k}^{\hat{\mathbf{H}}}(y^k, 0) \ .$$

This is however a direct consequence of (12) and the definition of α_{y^i}. □

Definition 9 ($\langle .,. \rangle$). *For two random systems* $\mathbf{F}$ *and* $\mathbf{G}$ *we denote with* $\langle \mathbf{F}, \mathbf{G} \rangle$ *the system which is defined by first sampling a bit* b *uniformly at random and then setting the system to be* $\mathbf{F}$ *if* $b = 0$ *and* $\mathbf{G}$ *if* $b = 1$.

The following trivial lemma will be useful:

Lemma 3. *For any systems* $\mathbf{F}$ *and* $\mathbf{G}$ *and any distinguisher* $\mathbf{D}$

$$\Delta_k^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) = 2 \cdot \Delta_k^{\mathbf{D}}(\mathbf{F}, \langle \mathbf{F}, \mathbf{G} \rangle). \quad (16)$$

Definition 10 (Composition Operator $\bowtie$). *A composition operator* $\bowtie$ *for a class of random systems* $\mathcal{I}$ *is a random system which expects oracle access to two systems from* $\mathcal{I}$. *With* $\mathbf{F} \bowtie \mathbf{G}$ *we denote* $\bowtie$ *where the first and second oracle are instantiated with* $\mathbf{F}$ *and* $\mathbf{G}$, *respectively.*

We will only consider $\bowtie$ *where (i) for any* $\mathbf{F}, \mathbf{G} \in \mathcal{I}$, *also* $\mathbf{F} \bowtie \mathbf{G} \in \mathcal{I}$, *and that (ii) on every invocation of* $\mathbf{F} \bowtie \mathbf{G}$ *the system* $\mathbf{F}$ *and* $\mathbf{G}$ *is invoked exactly once.*[9]

Definition 11 (Closure of $\mathcal{D}$ with $\bowtie$). *Let* $\mathcal{D}$ *be a class of distinguishers and* $\bowtie$ *a composition operator for* $\mathcal{I}$. *Let*

$$\mathcal{D}^{\bowtie,\mathcal{I}} := \{\mathcal{D}\diamond(\mathbf{F} \bowtie \ \cdot \ ), \mathcal{D}\diamond(\ \cdot \ \bowtie \mathbf{F}); \mathbf{F} \in \mathcal{I}\}$$

We say that $\bowtie$ *over* $\mathcal{I}$ *is closed for* $\mathcal{D}$ *if* $\mathcal{D}^{\bowtie,\mathcal{I}} \subseteq \mathcal{D}$.

For example if $\mathbf{D}$ is an $\mathsf{ATK} \in \{\mathsf{KPA}, \mathsf{nCPA}, \mathsf{CPA}\}$ distinguisher, then clearly so is $\mathbf{D}\diamond(\mathbf{F} \star \ \cdot \ )$ and $\mathbf{D}\diamond(\cdot \ \star\mathbf{F})$ for any $\mathbf{F} \in \mathcal{R}_S$. Also sequential composition is closed for the attacks considered in this paper (though for some attacks only in the stateless case).

Proposition 1.

(i) $\triangleright$ *over* $\mathcal{P}_S$ *is closed for* nCPA, CPA, *and* CCA, *and over* $\mathcal{P}$ *also for* KPA *and* nCCA.

(ii) $\star$ *over* $\mathcal{R}_S$ *is closed for* KPA, nCPA, *and* CPA.

To motivate Lemma 4 below, consider a gambler who plays black-jack at two different tables, and that he can schedule his moves in the two games arbitrarily, depending on the state of the other game. Assume that his only goal is to win *both* games. The following lemma states that playing independent optimal strategies at each table is optimal, i.e., that changing tables between moves does not help.[10] The proof of the Lemma is given in Appendix C.

[9] This definition could be generalized in many ways in order to capture more sophisticated constructions. For example in [MOPS06] we use an operator (the Feistel network) which combines more than two components and where (i) does not hold.

[10] This statement appears intuitive and perhaps simple to prove, but we point out that there exist very similar statements which are false and serve as paradoxes in probability theory. In particular, a simpler proof via two separate distinguishers which cannot communicate, but each with access to the other system's randomness, does not work. Induction appears unavoidable.

Lemma 4 (Independent vs. combined attacks)**.** *For any random systems* $\mathbf{S}$ *and* $\mathbf{T}$ *with MBOs, and any attack* $\mathsf{ATK} \in \{\mathsf{nCPA}, \mathsf{CPA}\}$*, or if* $\mathbf{S}^{\rightarrow}, \mathbf{T}^{\rightarrow}$ *are permutations* $\mathsf{ATK} \in \{\mathsf{nCCA}, \mathsf{CCA}\}$*,*[11]

$$\nu_k^{\mathsf{ATK}_2}(\mathbf{S}, \mathbf{T}) \;=\; \nu_k^{\mathsf{ATK}}(\mathbf{S}) \cdot \nu_k^{\mathsf{ATK}}(\mathbf{T})$$

For two systems $\mathbf{S}$ and $\mathbf{T}$ with MBOs consider an attacker whose task it is to set both MBOs to 1. The following lemma states the intuitive fact that — if $\bowtie$ is closed for the attacks considered — this task is not easier when having access to the combined system (using $\bowtie$) instead of having access to each system separately.

Lemma 5. *For any composition operator* $\bowtie$ *and any class* ATK *of distinguishers, if* $\bowtie$ *over* $\mathcal{I}$ *is closed for* ATK*, then we have for any* $k \in \mathbb{N}$ *and* $\mathbf{S}, \mathbf{T}$ *(where* $\mathbf{S}^{\rightarrow}, \mathbf{T}^{\rightarrow} \in \mathcal{I}$*)*

$$\nu_k^{\mathsf{ATK}}(\mathbf{S} \overset{\wedge}{\bowtie} \mathbf{T}) \leq \nu_k^{\mathsf{ATK}_2}(\mathbf{S}, \mathbf{T})$$

Proof. Let $\mathbf{D}$ be an ATK distinguisher where $\nu_k^{\mathsf{ATK}}(\mathbf{S} \overset{\wedge}{\bowtie} \mathbf{T}) = \nu_k^{\mathbf{D}}(\mathbf{S} \overset{\wedge}{\bowtie} \mathbf{T})$ and consider the double-distinguisher $\mathbf{D}'$ where $\mathbf{D}' \diamond [\mathbf{S}, \mathbf{T}]$ simulates the random experiment $\mathbf{D} \diamond \mathbf{S}^{\rightarrow} \bowtie \mathbf{T}^{\rightarrow}$, then $\nu_k^{\mathbf{D}}(\mathbf{S} \overset{\wedge}{\bowtie} \mathbf{T}) = \nu_k^{\mathbf{D}'}(\mathbf{S}, \mathbf{T})$. As $\bowtie$ is closed for ATK, $\mathbf{D}' \diamond [\mathbf{S}, \,\cdot\,]$ and $\mathbf{D}' \diamond [\,\cdot\,, \mathbf{T}]$ are ATK distinguishers, and thus $\mathbf{D}'$ is a ATK_2 double-distinguisher. We conclude $\nu_k^{\mathsf{ATK}}(\mathbf{S} \overset{\wedge}{\bowtie} \mathbf{T}) = \nu_k^{\mathbf{D}'}(\mathbf{S}, \mathbf{T}) \leq \nu_k^{\mathsf{ATK}_2}(\mathbf{S}, \mathbf{T})$. □

Definition 12 (Transparent $\bowtie$)**.** *A composition operator* $\bowtie$ *is transparent for* $\mathcal{I}$*, if there is a system* $\mathbf{I} \in \mathcal{I}$ *(which we call the ideal system for* $\mathcal{I}$*) such that for all* $\mathbf{F} \in \mathcal{I}$

$$\mathbf{F} \bowtie \mathbf{I} \equiv \mathbf{I} \bowtie \mathbf{F} \equiv \mathbf{I}$$

As the cascade of a URP with any other (independent) stateless permutation is again a URP, and the parallel composition of a URF with any other function is a URF we get the following simple proposition.

Proposition 2.

- $\triangleright$ *is transparent for* $\mathcal{P}$ *(but not for* $\mathcal{P}_S$*) with the ideal system being* $\mathbf{P}$*.*
- $\star$ *is transparent for* $\mathcal{R}_S$ *with the ideal systems being* $\mathbf{R}$*.*

4 The Direct Product Theorem

Corollary 1 follows from Propositions 1 and 2 and the following theorem which we will prove in this section.

Theorem 1 (Direct Product)**.** *For any* transparent *composition operator* $\bowtie$ *(over* $\mathcal{I}$ *with ideal system* $\mathbf{I}$*) and any* $\mathsf{ATK} \in \{\mathsf{nCPA}, \mathsf{CPA}\}$ *or if* $\mathcal{I} \subseteq \mathcal{P}_S$ *also* $\mathsf{ATK} \in \{\mathsf{nCCA}, \mathsf{CCA}\}$*, if* $\bowtie$ *over* $\mathcal{I}$ *is* closed *for* ATK*, we have for any* $\mathbf{F}, \mathbf{G} \in \mathcal{I}$

$$\Delta_q^{\mathsf{ATK}}(\mathbf{F} \bowtie \mathbf{G}, \mathbf{I}) \;\leq\; 2 \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{F}, \mathbf{I}) \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{G}, \mathbf{I})$$

Lemma 6. *For* $\bowtie$ *as in the theorem and random systems with MBOs* $\mathbf{S}_0, \ldots, \mathbf{S}_3$ *where* $\mathbf{S}_0{}^{\rightarrow}, \ldots, \mathbf{S}_3{}^{\rightarrow} \in \mathcal{I}$ *and*

$$\mathbf{S}_0{}^{\dashv} \equiv \mathbf{S}_1{}^{\dashv} \qquad\qquad \mathbf{S}_2{}^{\dashv} \equiv \mathbf{S}_3{}^{\dashv} \tag{17}$$

Then

$$\langle \mathbf{S}_0 \overset{\wedge}{\bowtie} \mathbf{S}_2, \mathbf{S}_1 \overset{\wedge}{\bowtie} \mathbf{S}_3 \rangle^{\dashv} \equiv \langle \mathbf{S}_0 \overset{\wedge}{\bowtie} \mathbf{S}_3, \mathbf{S}_1 \overset{\wedge}{\bowtie} \mathbf{S}_2 \rangle^{\dashv} \tag{18}$$

[11] Let us stress that the lemma is wrong for $\mathsf{ATK} = \mathsf{KPA}$, as a KPA_2 attacker (who must make random queries) can use correlated queries (say the same query for both systems), and thus also correlate the probability of the two MBOs becoming 1.

Proof. We can write (18) as $\mathbf{T}_0^{\dashv} \equiv \mathbf{T}_1^{\dashv}$ where $\mathbf{T}_c$ is $\mathbf{S}_i \overset{\triangle}{\bowtie} \mathbf{S}_j$ for random $i \in \{0,1\}, j \in \{2,3\}$ conditioned on $i + j \bmod 2 = c$.

Now assume we query $\mathbf{T}_c$ for a random c, and at some point the MBO of component $\mathbf{S}_i$ becomes 1 (but the MBO of $\mathbf{S}_j$ is still 0). From then on $\mathbf{T}_c$ behaves either as $S_0 \overset{\triangle}{\bowtie} S_j$ or $S_1 \overset{\triangle}{\bowtie} S_j$, but this gives no no information on $c = i + j \mod 2$, as if for example $i = 0$, then $c = 0$ if $j = 2$ but $c = 1$ if $j = 3$, but as the MBO of S_j is 0 and by assumption $\mathbf{S_2}^{\dashv} \equiv \mathbf{S_3}^{\dashv}$, no information on j has been leaked yet. Similarly, when only the MBO of S_j becomes 1 no information about c can be learned. With this observation we see that $\mathbf{T}_0$ and $\mathbf{T}_1$ behave exactly the same as long as not both MBOs are 1, i.e $\mathbf{T_0}^{\dashv} \equiv \mathbf{T_1}^{\dashv}$. □

Proof of Theorem 1. Recall that $\bowtie$ being transparent for $\mathcal{I}$ means that for any $\mathbf{F} \in \mathcal{I}$

$$\mathbf{F} \bowtie \mathbf{I} \equiv \mathbf{I} \bowtie \mathbf{F} \equiv \mathbf{I} \bowtie \mathbf{I} \equiv \mathbf{I}. \tag{19}$$

Consider any $\mathbf{F}, \mathbf{G} \in \mathcal{I}$, by Lemma 2 we can add MBOs to those systems in order to get $\hat{\mathbf{F}}, \hat{\mathbf{G}}, \hat{\mathbf{I}}$ and $\hat{\mathbf{I}}'$ (here $\hat{\mathbf{F}}^{\rightarrow} \equiv \mathbf{F}, \hat{\mathbf{G}}^{\rightarrow} \equiv \mathbf{G}, \hat{\mathbf{I}}^{\rightarrow} \equiv \hat{\mathbf{I}}'^{\rightarrow} \equiv \mathbf{I}$) which satisfy

$$\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{I}}^{\dashv} \quad \text{and} \quad \hat{\mathbf{G}}^{\dashv} \equiv \hat{\mathbf{I}}'^{\dashv} \tag{20}$$

and for all $\mathbf{D}$

$$\nu_q^{\mathbf{D}}(\hat{\mathbf{I}}) = \nu_q^{\mathbf{D}}(\hat{\mathbf{F}}) = \Delta_q^{\mathbf{D}}(\mathbf{F}, \mathbf{I}) \tag{21}$$

$$\nu_q^{\mathbf{D}}(\hat{\mathbf{I}}') = \nu_q^{\mathbf{D}}(\hat{\mathbf{G}}) = \Delta_q^{\mathbf{D}}(\mathbf{G}, \mathbf{I}) \tag{22}$$

Now the theorem follows as (In the system $\mathbf{I} \bowtie \mathbf{I}$ below the two components are realized by *independent* instantiations of $\mathbf{I}$)

$$\begin{aligned}
& \Delta_q^{\mathsf{ATK}}(\mathbf{F} \bowtie \mathbf{G}, \mathbf{I}) \\
\overset{Lem.3}{=} \quad & 2 \cdot \Delta_q^{\mathsf{ATK}}(\langle \mathbf{F} \bowtie \mathbf{G}, \mathbf{I} \rangle, \mathbf{I}) \\
\overset{(19)}{=} \quad & 2 \cdot \Delta_q^{\mathsf{ATK}}(\langle \mathbf{F} \bowtie \mathbf{G}, \mathbf{I} \bowtie \mathbf{I} \rangle, \langle \mathbf{F} \bowtie \mathbf{I}, \mathbf{I} \bowtie \mathbf{G} \rangle) \\
\overset{Lem.1\ \&\ 6}{\leq} \quad & 2 \cdot \nu_q^{\mathsf{ATK}}(\langle \hat{\mathbf{F}} \overset{\triangle}{\bowtie} \hat{\mathbf{G}}, \hat{\mathbf{I}} \overset{\triangle}{\bowtie} \hat{\mathbf{I}}' \rangle) \\
\overset{Def.9\ \&\ union\ bound}{\leq} \quad & \nu_q^{\mathsf{ATK}}(\hat{\mathbf{F}} \overset{\triangle}{\bowtie} \hat{\mathbf{G}}) + \nu_q^{\mathsf{ATK}}(\hat{\mathbf{I}} \overset{\triangle}{\bowtie} \hat{\mathbf{I}}') \\
\overset{Lem.5}{\leq} \quad & \nu_q^{\mathsf{ATK}_2}(\hat{\mathbf{F}}, \hat{\mathbf{G}}) + \nu_q^{\mathsf{ATK}_2}(\hat{\mathbf{I}}, \hat{\mathbf{I}}') \\
\overset{Lem.4}{=} \quad & \nu_q^{\mathsf{ATK}}(\hat{\mathbf{F}}) \cdot \nu_q^{\mathsf{ATK}}(\hat{\mathbf{G}}) + \nu_q^{\mathsf{ATK}}(\hat{\mathbf{I}}) \cdot \nu_q^{\mathsf{ATK}}(\hat{\mathbf{I}}') \\
\overset{(21,22)}{=} \quad & 2 \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{F}, \mathbf{I}) \cdot \Delta_q^{\mathsf{ATK}}(\mathbf{G}, \mathbf{I})
\end{aligned}$$

□

5 When Two Weak Make One Strong

In this section we will prove a theorem which provides some conditions which, if satisfied, imply that the composition of components which are only secure against some weak class of attacks, is secure against stronger attacks. Corollary 2 follows from this theorem (for α and α' from the statement of the theorem being 0) by basic arguments as shown in Appendix B.

Theorem 2. *Consider three classes of attacks, a strong one* ATK *and two weak ones denoted* wATK *and* wATK'*. Let* $\bowtie$ *be any composition operator (over* $\mathcal{I}$ *with ideal system* $\mathbf{I}$*) where* $\bowtie$ *over* $\mathcal{I}$ *is closed for* wATK *and* wATK'*. If there exist*[12] $\alpha, \alpha' \in \mathbb{R}$ *such that for all all random systems with an MBO* $\mathbf{S}$ *(where* $\mathbf{S}^{\rightarrow} \in \mathcal{I}$*) and all* $k \in \mathbb{N}$

$$\nu_k^{\mathsf{ATK}}(\mathbf{S} \bowtie \mathbf{I}) \leq \nu_k^{\mathsf{wATK}}(\mathbf{S} \bowtie \mathbf{I}) + \alpha \quad \textit{and} \quad \nu_k^{\mathsf{ATK}}(\mathbf{I} \bowtie \mathbf{S}) \leq \nu_k^{\mathsf{wATK}'}(\mathbf{I} \bowtie \mathbf{S}) + \alpha' \tag{23}$$

[12] In this paper we will only use the case where $\alpha = \alpha' = 0$. We prove the more general case of the theorem with non-zero values for α and α' as it is not harder to prove and has applications in other works, in particular [MOPS06].

then for all $\mathbf{F}, \mathbf{G} \in \mathcal{I}$ *and all* $k \in \mathbb{N}$

$$\Delta_k^{\mathsf{ATK}}(\mathbf{F} \bowtie \mathbf{G}, \mathbf{I} \bowtie \mathbf{I}) \leq \Delta_k^{\mathsf{wATK}}(\mathbf{F}, \mathbf{I}) + \Delta_k^{\mathsf{wATK}'}(\mathbf{G}, \mathbf{I}) + \alpha + \alpha'$$

Proof. Consider any $\mathbf{F}, \mathbf{G} \in \mathcal{I}$, by Lemma 2 we can add MBOs to those systems in order to get $\hat{\mathbf{F}}, \hat{\mathbf{G}}, \hat{\mathbf{I}}$ and $\hat{\mathbf{I}}'$ (here $\hat{\mathbf{F}}^{\rightarrow} \equiv \mathbf{F}, \hat{\mathbf{G}}^{\rightarrow} \equiv \mathbf{G}, \hat{\mathbf{I}}^{\rightarrow} \equiv \hat{\mathbf{I}'}^{\rightarrow} \equiv \mathbf{I}$) which satisfy

$$\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{I}}^{\dashv} \quad \text{and} \quad \hat{\mathbf{G}}^{\dashv} \equiv \hat{\mathbf{I}'}^{\dashv} \tag{24}$$

and for all $\mathbf{D}$

$$\begin{aligned} \nu_q^{\mathbf{D}}(\hat{\mathbf{I}}) = \nu_q^{\mathbf{D}}(\hat{\mathbf{F}}) &= \Delta_q^{\mathbf{D}}(\mathbf{F}, \mathbf{I}) && (25) \\ \nu_q^{\mathbf{D}}(\hat{\mathbf{I}}') = \nu_q^{\mathbf{D}}(\hat{\mathbf{G}}) &= \Delta_q^{\mathbf{D}}(\mathbf{G}, \mathbf{I}) && (26) \end{aligned}$$

From eq. (24) we get

$$(\hat{\mathbf{F}} \overset{\vee}{\bowtie} \hat{\mathbf{G}})^{\dashv} \equiv (\hat{\mathbf{I}} \overset{\vee}{\bowtie} \hat{\mathbf{I}}')^{\dashv}. \tag{27}$$

Now the theorem follows as

$$\begin{aligned} & \Delta_q^{\mathsf{ATK}}(\mathbf{F} \bowtie \mathbf{G}, \mathbf{I} \bowtie \mathbf{I}) && (28) \\ \overset{Lem.1\ \&\ (27)}{=} \quad & \nu_q^{\mathsf{ATK}}(\hat{\mathbf{I}} \overset{\vee}{\bowtie} \hat{\mathbf{I}}') && (29) \\ \overset{union\ bound}{\leq} \quad & \nu_q^{\mathsf{ATK}}(\hat{\mathbf{I}} \bowtie \mathbf{I}) + \nu_q^{\mathsf{ATK}}(\mathbf{I} \bowtie \hat{\mathbf{I}}') && (30) \\ \overset{(23)}{=} \quad & \nu_q^{\mathsf{wATK}}(\hat{\mathbf{I}} \bowtie \mathbf{I}) + \nu_q^{\mathsf{wATK}'}(\mathbf{I} \bowtie \hat{\mathbf{I}}') + \alpha + \alpha' && (31) \\ \overset{Def.11}{\leq} \quad & \nu_q^{\mathsf{wATK}}(\hat{\mathbf{I}}) + \nu_q^{\mathsf{wATK}'}(\hat{\mathbf{I}}') + \alpha + \alpha' && (32) \\ \overset{(25,26)}{=} \quad & \Delta_q^{\mathsf{wATK}}(\mathbf{F}, \mathbf{I}) + \Delta_q^{\mathsf{wATK}'}(\mathbf{G}, \mathbf{I}) + \alpha + \alpha' && (33) \end{aligned}$$

□

References

[KNR05] Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. In *Random-Approx 2005*, 2005.

[LR86] Michael Luby and Charles Rackoff. Pseudo-random permutation generators and cryptographic composition. In *Proc, 18th ACM Symposium on the Theory of Computing (STOC)*, pages 356–363, 1986.

[Mau02] Ueli Maurer. Indistinguishability of random systems. In *Advances in Cryptology — EUROCRYPT '02*, pages 110–132, 2002.

[MOPS06] Ueli Maurer, Yvonne Anne Oswald, Krzysztof Pietrzak, and Johan Sjödin. Luby-Rackoff ciphers with weak round functions. In *Advances in Cryptology — EUROCRYPT '06*, volume 4004, pages 391–408, 2006.

[MP04] Ueli Maurer and Krzysztof Pietrzak. Composition of random systems: When two weak make one strong. In *Theory of Cryptograpy — TCC '04*, volume 2951, pages 410–427, 2004.

[Mye03] Steven Myers. Efficient amplification of the security of weak pseudo-random function generators. *Journal of Cryptology*, 16(1):1–24, 2003.

[Mye04] Steven Myers. Black-box composition does not imply adaptive security. In *Advances in Cryptology — EUROCRYPT 04*, volume 3027, pages 189–206, 2004.

[Pie05] Krzysztof Pietrzak. Composition does not imply adaptive security. In *Advances in Cryptology — CRYPTO '05*, volume 3621, pages 55–65, 2005.

[Pie06] Krzysztof Pietrzak. Composition implies adaptive security in minicrypt. In *Advances in Cryptology — EUROCRYPT '06*, volume 4004, pages 328–338, 2006.

[PS06] Krzysztof Pietrzak and Johan Sjödin. Domain extension for weak PRFs; the good, the bad, and the ugly, 2006. Manuscript.

[Vau98] Serge Vaudenay. Provable security for block ciphers by decorrelation. In *STACS*, pages 249–275, 1998.

[Vau99] Serge Vaudenay. Adaptive-attack norm for decorrelation and super-pseudorandomness. In *Selected Areas in Cryptography*, pages 49–61, 1999.

[Vau03] Serge Vaudenay. Decorrelation: A theory for block cipher security. *J. Cryptology*, 16(4):249–286, 2003.

A On the Tightness of Corollaries 1 and 2

In this section we have a look at the tightness of Corollaries 1 and 2 (and thus also the underlying Theorems 1 and 2).

Corollary 2 (Adaptive Security by Composition) is Tight. We consider only the statement for sequential composition. This is, we must construct a random permutation $\mathbf{F}$ such that for some q we have $\Delta_q^{\mathsf{CPA}}(\mathbf{F} \triangleright \mathbf{F}, \mathbf{P}) \approx 2 \cdot \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P})$. The idea is to first define a random permutation $\mathbf{G}$ which can be distinguished from a URP $\mathbf{P}$ with adaptive, but not with non-adaptive queries. Then we define $\mathbf{F}$ as being the identity function $\mathbf{I}$ with some probability ϵ, and $\mathbf{G}$ with probability $1-\epsilon$. Now $\mathbf{F} \triangleright \mathbf{F}$ is $\mathbf{G}$ if one of the components is $\mathbf{I}$ (as $\mathbf{G} \triangleright \mathbf{I} \equiv \mathbf{I} \triangleright \mathbf{G} \equiv \mathbf{G}$), this happens with probability $2(1-\epsilon)\epsilon$ which is basically 2ϵ for small ϵ. If $\mathbf{G}$ is easily distinguished with q adaptive queries, then $\Delta_q^{\mathsf{CPA}}(\mathbf{F} \triangleright \mathbf{F}, \mathbf{P}) \approx \mathsf{P}[\mathbf{F} \triangleright \mathbf{F} \equiv \mathbf{G}] \approx 2\epsilon$ and if $\mathbf{G}$ is hard to distinguish from $\mathbf{P}$ with q non-adaptive queries then $\Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P}) \approx \mathsf{P}[\mathbf{F} \equiv \mathbf{I}] = \epsilon$, thus $\Delta_q^{\mathsf{CPA}}(\mathbf{F} \triangleright \mathbf{F}, \mathbf{P}) \approx 2 \cdot \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P})$ as required.

We still have to give a realization for $\mathbf{G}$, i.e. a random permutation where $\Delta_q^{\mathsf{CPA}}(\mathbf{G}, \mathbf{P}) \approx 1$ and $\Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{P}) \approx 0$ for some q. We now give such a $\mathbf{G}$ where this is the case for any $q \geq 2$ and $q \ll N/2$ where N denotes the domain size of the permutations considered. Let $\mathbf{G}$ be a uniformly random permutation conditioned on some distinct element 0 satisfying $\mathbf{G}(\mathbf{G}(0)) = 0$ (i.e. 0 lies on a cycle of length 2). Using standard techniques (e.g. [Mau02]) it is easy to show that $\Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{P}) \leq 2q/N$ and $\Delta_2^{\mathsf{CPA}}(\mathbf{G}, \mathbf{P}) \geq 1 - 2/N$.

There is a similar example which proves the tightness of the statement for parallel composition of Corollary [**?**]. Corollary 1 (Direct Product) is Tight. Consider the following example from [Mye03]; Let $\mathbf{F} : \{0,1\}^n \to \{0,1\}$ be a uniformly random function, but with the restriction that on input 0^n the output is always 0. Note that then, for $\star$ being XOR, also $\mathbf{F} \star \mathbf{F} \equiv \mathbf{F}$. So, as $\mathsf{P}[\mathbf{R}(0^n) = 0] = 0.5$ and 0^n is the only query where $\mathbf{F}$ and $\mathbf{R}$ differ, we get that for $\mathsf{ATK} \in \{\mathsf{nCPA}, \mathsf{CPA}\}$ and any $q \geq 1$

$$\Delta_q^{\mathsf{ATK}}(\mathbf{F}, \mathbf{R}) = 0.5 \qquad \text{and thus as } \mathbf{F} \star \mathbf{F} \equiv \mathbf{F} \text{ also} \qquad \Delta_q^{\mathsf{ATK}}(\mathbf{F} \star \mathbf{F}, \mathbf{R}) = 0.5$$

Because $2 \cdot (0.5)^2 = 0.5$ this shows that the statement for parallel composition of Corollary 1 is tight. This counterexample does not directly translate to a counterexample for sequential composition, but consider the two following examples:

Let $\mathbf{F}$ be a permutation over $\{0,1\}^n$ which keeps the first bit fixed, but otherwise is uniformly random. For this $\mathbf{F}$ we have $\mathbf{F} \triangleright \mathbf{F} \equiv \mathbf{F}$ and thus $\Delta_1^{\mathsf{ATK}}(\mathbf{F}, \mathbf{P}) = \Delta_1^{\mathsf{ATK}}(\mathbf{F} \triangleright \mathbf{F}, \mathbf{P}) = 0.5$.

Let $\mathbf{G}$ be a uniformly random permutation which is *even* (i.e. can be decomposed into an even number of transpositions), then also $\mathbf{G} \triangleright \mathbf{G}$ is even. As exactly half the permutations (over any domain of size at least 2) are even we get $\Delta_N^{\mathsf{ATK}}(\mathbf{G}, \mathbf{P}) = \Delta_N^{\mathsf{ATK}}(\mathbf{G} \triangleright \mathbf{G}, \mathbf{P}) = 0.5$ where N denotes the size of the domain considered.

This two examples show that Corollary 1 is tight for sequential composition if the number of queries is 1 or N. We do not know if the bound (for this particular construction) is tight "in-between", i.e. if $1 \ll q \ll N$. It is possible that for this range the constant 2 can be replaced with $1+\epsilon$ where $\epsilon = \epsilon(q,N)$ is some function of q, N and $\epsilon(q,N) \ll 1$ for large N and $1 \ll q \ll N$.

B Proof of Corollary 2

Proof of Corollary 2. We first prove the third statement of the corollary, namely that for any (possibly stateful) random functions $\mathbf{F}, \mathbf{G} \in \mathcal{R}_S$

$$\Delta_q^{\mathsf{CPA}}(\mathbf{F} \star \mathbf{G}, \mathbf{R}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{R}) + \Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{R}). \tag{34}$$

Eq. (34) follows by Theorem 2 using Proposition 1 (i.e. that $\star$ over $\mathcal{R}_S$ is closed for nCPA) and

$$\begin{aligned}\nu^{\mathsf{CPA}}(\hat{\mathbf{F}} \star \mathbf{R}) &= \nu^{\mathsf{nCPA}}(\hat{\mathbf{F}} \star \mathbf{R})\\ \nu^{\mathsf{CPA}}(\mathbf{R} \star \hat{\mathbf{G}}) &= \nu^{\mathsf{nCPA}}(\mathbf{R} \star \hat{\mathbf{G}})\end{aligned}$$

where the first statement (the second is symmetric) holds as the output of $\hat{\mathbf{F}}^{\rightarrow} \star \mathbf{R}$ is completely independent of the output of the subsystem $\hat{\mathbf{F}}^{\rightarrow}$, it is "blinded" by the uniformly random output of $\mathbf{R}$. Thus adaptivity cannot help in setting the MBO of $\hat{\mathbf{F}}$ to 1.

We now prove the two first statements of the corollary which state that for any stateless random permutations $\mathbf{F}, \mathbf{G} \in \mathcal{P}$

$$\Delta_q^{\mathsf{CPA}}(\mathbf{F} \triangleright \mathbf{G}, \mathbf{P}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P}) + \Delta_q^{\mathsf{KPA}}(\mathbf{G}, \mathbf{P}) \tag{35}$$

$$\Delta_q^{\mathsf{CCA}}(\mathbf{F} \triangleright \mathbf{G}^{-1}, \mathbf{P}) \leq \Delta_q^{\mathsf{nCPA}}(\mathbf{F}, \mathbf{P}) + \Delta_q^{\mathsf{nCPA}}(\mathbf{G}, \mathbf{P}) \tag{36}$$

As (35,36) are about stateless systems, we can wlog. assume that no distinguisher will ever repeat a query.[13] Now (35) follows by Theorem 2 using Proposition 1 and

$$\begin{aligned}\nu^{\mathsf{CPA}}(\hat{\mathbf{F}} \triangleright \mathbf{P}) &= \nu^{\mathsf{nCPA}}(\hat{\mathbf{F}} \triangleright \mathbf{P})\\ \nu^{\mathsf{CPA}}(\mathbf{P} \triangleright \hat{\mathbf{G}}) &= \nu^{\mathsf{KPA}}(\mathbf{P} \triangleright \hat{\mathbf{G}})\end{aligned}$$

Here the first statement holds as the output of $\hat{\mathbf{F}}^{\rightarrow} \triangleright \mathbf{P}$ is again independent of the output of the subsystem $\hat{\mathbf{F}}^{\rightarrow}$ as it is permuted by a URP $\mathbf{P}$. The second equation holds as in $\mathbf{P} \triangleright \hat{\mathbf{G}}^{\rightarrow}$, the input to the subsystem $\hat{\mathbf{G}}$ are all distinct (as we do not allow the distinguisher to repeat a query) and uniformly random no matter with what adaptive strategy one uses to query $\mathbf{P} \triangleright \hat{\mathbf{G}}$, thus the inputs to $\hat{\mathbf{G}}$ are distributed as in a KPA attack. Finally (36) follows by Lemma 2 using (note that $\mathbf{P} \equiv \mathbf{P}^{-1}$)

$$\begin{aligned}\nu^{\mathsf{CCA}}(\hat{\mathbf{F}} \triangleright \mathbf{P}) &= \nu^{\mathsf{nCPA}}(\hat{\mathbf{F}} \triangleright \mathbf{P})\\ \nu^{\mathsf{CCA}}(\mathbf{P} \triangleright \hat{\mathbf{G}}^{-1}) &= \nu^{\mathsf{nCPA}}(\mathbf{P} \triangleright \hat{\mathbf{G}}^{-1})\end{aligned}$$

To see the first statement (the second is symmetric as we consider CCA attacks), we observe that

$$\nu^{\mathsf{CCA}}(\hat{\mathbf{F}} \triangleright \mathbf{P}) = \nu^{\mathsf{CPA}}(\hat{\mathbf{F}} \triangleright \mathbf{P})$$

as making an inverse query to $\hat{\mathbf{F}}^{\rightarrow} \triangleright \mathbf{P}$ results in a value on the input which is random and independent of $\hat{\mathbf{F}}^{\rightarrow}$. So this queries give no additional advantage in setting the MBO of $\hat{\mathbf{F}}$ to 1 as they can be "simulated" by random forward queries. We have already argued that CPA is no better than nCPA here to show (35). □

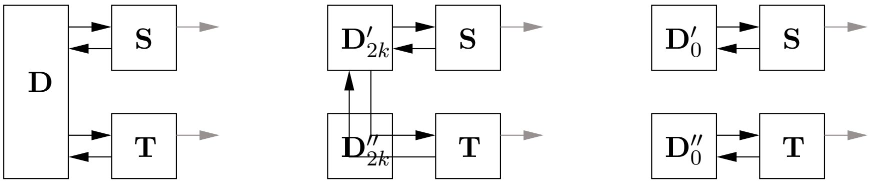

Figure 1: Illustration for proof of Lemma 4. An ATK_2 double-distinguisher $\mathbf{D}$ can be seen as a pair $\mathbf{D}'_{2k}$,$\mathbf{D}''_{2k}$ of ATK distinguishers which can additionally exchange up to $2k$ messages (simply set $\mathbf{D}'_{2k} \equiv \mathbf{D}$ and $\mathbf{D}''_{2k}$ to be the trivial system which only passes messages). The gray arrows indicate the MBOs.

C Proof of Lemma 4

Proof. Let $\mathbf{D}$ be an optimal ATK_2 double-distinguisher for the task considered, i.e.

$$\nu_k^{\mathsf{ATK}_2}(\mathbf{S},\mathbf{T}) \;=\; \nu_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$$

Let $A_1, A_2, \ldots$ and $B_1, B_2, \ldots$ denote the MBO of $\mathbf{S}$ and $\mathbf{T}$ respectively. We can see the ATK_2 double-distinguisher $\mathbf{D}$ as a pair of ATK distinguishers $\mathbf{D}'_{2k}$ and $\mathbf{D}''_{2k}$ which can exchange up to $2k$ messages with each other as shown in Figure 1 (where $\mathbf{D}'_{2k} \equiv \mathbf{D}$ and $\mathbf{D}''_{2k}$ is a dummy system which simply passes messages). As this is just a conceptual change, the advantage of setting both MBOs to 1 is exactly the same for $\mathbf{D}$ as for the pair $\mathbf{D}'_{2k}, \mathbf{D}''_{2k}$.

Now assume that there is a pair of ATK distinguishers $\mathbf{D}'_\ell$ and $\mathbf{D}''_\ell$ which can exchange up to ℓ messages and have advantage ϵ to provoke $A_k = 1 \wedge B_k = 1$ when querying $\mathbf{S}$ and $\mathbf{T}$ respectively (we just saw that such distinguishers exist for $\ell = 2k$ with optimal $\epsilon = \nu_k^{\mathsf{ATK}_2}(\mathbf{S},\mathbf{T})$). We claim that then there also exist ATK distinguishers $\mathbf{D}'_{\ell-1}$ and $\mathbf{D}''_{\ell-1}$ which exchange one message less but still have advantage at least ϵ to provoke $A_k = 1 \wedge B_k = 1$. Before we prove this claim, note that it implies the lemma as by induction there now exist $\mathbf{D}'_0$ and $\mathbf{D}''_0$ (which do not communicate at all) where

$$\nu_k^{\mathsf{ATK}_2}(\mathbf{S},\mathbf{T}) \le \nu_k^{\mathbf{D}'_0}(\mathbf{S}) \cdot \nu_k^{\mathbf{D}''_0}(\mathbf{T}) \le \nu_k^{\mathsf{ATK}}(\mathbf{S}) \cdot \nu_k^{\mathsf{ATK}}(\mathbf{T}).$$

We actually have equality above as the $\ge$ direction is trivial. To prove the claim, assume the (last) ℓ-th message is sent from $\mathbf{D}'_\ell$ to $\mathbf{D}''_\ell$. Let the random variable M denote this last message, and let V be the "view" of $\mathbf{D}''_\ell$ just before receiving the message. Let $\mathcal{E}$ denote this random experiment where $\mathbf{D}'_\ell$ and $\mathbf{D}''_\ell$ are querying $\mathbf{S}$ and $\mathbf{T}$ respectively. The probability that we have $A_k = 1 \wedge B_k = 1$ is

$$\sum_{m,v} \mathsf{P}^{\mathcal{E}}[A_k = 1 \wedge M = m \wedge V = v] \cdot \mathsf{P}^{\mathcal{E}}[B_k = 1 | M = m \wedge V = v]. \tag{37}$$

We used $\mathsf{P}^{\mathcal{E}}[B_k = 1 | A_k = 1 \wedge M = m \wedge V = v] = \mathsf{P}^{\mathcal{E}}[B_k = 1 | M = m \wedge V = v]$ which holds as $\mathbf{S}$ is independent of $\mathbf{T}$ and the whole interaction between these systems is captured by M, V. Now consider a new system $\mathbf{D}''_{\ell-1}$ which simulates $\mathbf{D}''_\ell$ but does not expect the (last) ℓ-th message M and instead replaces it with a message m' which maximizes the probability of $B_k = 1$ (given the view V). Also, let $\mathbf{D}'_{\ell-1}$ be the system $\mathbf{D}'_\ell$, but where the last message is not send (note that this change does not affect the probability of $A_k = 1$ or the distribution of V). The probability that the pair $(\mathbf{D}'_{\ell-1}$,$\mathbf{D}''_{\ell-1})$ can provoke $A_k = 1 \wedge B_k = 1$ is thus

$$\sum_{m,v} \mathsf{P}^{\mathcal{E}}[A_k = 1 \wedge M = m \wedge V = v] \cdot \max_{m'} \mathsf{P}^{\mathcal{E}}[B_k = 1 | M = m' \wedge V = v]$$

which is at least equal to (37). □

[13] This is not the case for stateful systems, as even though also in this case we will not learn anything new by repeating a query, such a query can change still change the state.