Indistinguishability Amplification

Ueli Maurer[1], Krzysztof Pietrzak[2], and Renato Renner[3]

[1]Department of Computer Science, ETH Zurich,
Email: maurer@inf.ethz.ch
[2]CWI Amsterdam, Email: pietrzak@cwi.nl
[3]University of Cambridge, Email: r.renner@damtp.cam.ac.uk

Abstract Many aspects of cryptographic security proofs can be seen as the proof that a certain 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 information-theoretic distinguishing advantage (from an ideal system) for a combined system, assuming upper bounds of certain types for the component systems. For a general type of combination operation of systems, including the XOR of functions or the cascade of permutations, we prove two amplification theorems. The first is a product theorem, in the spirit of XOR-lemmas: The distinguishing advantage of the combination of two systems is at most twice the product of the individual distinguishing advantages. This bound 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 distinguishers.

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

1 Introduction

1.1 Indistinguishability Amplification for Random Variables

This paper is concerned with the indistinguishability of systems that interact with their environment. As a motivation for this paper, we consider an indistinguishability amplification result for random variables. A random variable can be understood as the special case of a system, which is non-interactive. Lemma 1 below states that the distance from uniform, of random variables, can be amplified by combining two or more (independent) moderately uniform random variables.

To state the lemma, we recall the following definitions.

Definition 1. The *statistical distance* of two random variables X and X' over $\mathcal{X}$ is defined as

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

The *distance* of a random variable X *from uniform* is $d(X) := \delta(X, U)$, where U is a uniform random variable on $\mathcal{X}$.

The advantage of the best distinguisher for X and X' is $\delta(X, X')$.

Lemma 1. *For any two independent random variables X and Y over a finite domain $\mathcal{X}$ and any quasi-group operation*[1] $\star$ *on $\mathcal{X}$,*

$$d(X \star Y) \ \leq \ 2\, d(X)\, d(Y).$$

This bound is tight, as the following example illustrates.

Example 1. Consider two independent biased bits, X with a 40/60-bias and Y with a 30/70-bias. Then $d(X) = 0.1$, $d(Y) = 0.2$, and $d(X \oplus Y) = 0.04$ $(= 2 \cdot 0.1 \cdot 0.2)$, since $X \oplus Y$ is 54/46-biased.

Corollary 2 of this paper can be seen as a natural generalization of Lemma 1. It states (for example) that if $\mathbf{F}$ and $\mathbf{G}$ are systems, for each of which the best distinguisher's advantage in distinguishing it from a uniform random function is bounded by ϵ and ϵ', respectively, then the system $\mathbf{F} \star \mathbf{G}$ obtained by using $\mathbf{F}$ and $\mathbf{G}$ in parallel and combining their outputs with $\star$, can be distinguished with advantage at most $2\epsilon\epsilon'$ from a uniform random function (for the same number of queries issued by the distinguisher). Actually, the proof of Corollary 2, restricted to random variables, appears to be a natural proof for Lemma 1.

As the abstraction underlying the quasi-group operation we introduce the concept of a *neutralizing combination* of two systems, which means that if any one (or both) of the systems is an ideal system (e.g. a uniform random function), then the combined system is also ideal. This is for example true for $X \star Y$: If either X or Y is uniform, then so is $X \star Y$.

1.2 Contributions of this Paper

The *amplification* of security properties is an important theme in cryptography. Examples of amplification results are XOR-lemmas, Vaudenay's product theorem for random permutations [Vau99], and the theorems proving adaptive security from non-adaptive security assumptions of [MP04] and [MOPS06].

This paper generalizes, strengthens, and unifies these results and provides a framework for proving such amplification results. We explore the general problem of proving various indistinguishability amplification results for systems. In contrast to earlier works, we do not restrict ourselves to *stateless* systems. The term "amplification" is used with two different meanings:

- **Reduction of the distinguishing advantage.** We prove a general theorem (Theorem 1), in the spirit of Lemma 1, which states that the distinguishing advantage of a neutralizing combination of two systems is at most twice the product of the individual distinguishing advantages.

[1] A quasi-group operation $\star$ on a set $\mathcal{X}$ is a function $\mathcal{X}^2 \to \mathcal{X} : (a, b) \mapsto c = a \star b$ such given a and c (b and c), b (a) is uniquely determined. An important example is the bit-wise XOR of bit-strings. Any group operation is also a quasi-group operation.

– **Attack strengthening.** We prove a general theorem (Theorem 2), which states that the adaptive distinguishing advantage of a neutralizing combination of two systems is bounded by the sum of the individual distinguishing advantages for a weaker distinguisher class (e.g. non-adaptive, or for permutations, one-sided instead of two-sided queries).

Our results are stated in the random systems framework of [Mau02] (see Section 2). They hold in the information-theoretic setting, with computationally unbounded distinguishers. In practice one is often interested in *computational* indistinguishability. Although the results from this paper do not directly translate to the computational setting[2], they have implications in the computational setting as well.

A main technical tool of this paper is a tight relation between the distinguishing advantage and the game-winning probability, discussed in the following section.

1.3 Discrete Systems, Indistinguishability, and Game-winning

Many cryptographic systems (e.g. a block cipher, the CBC-MAC construction, or more complex games) can be modeled as *discrete systems*. A discrete system interacts with its environment by taking a sequence of inputs and producing, for each new input, an output (for a single, a fixed, or an unbounded number of such interactions).

Two major paradigms for cryptographic security definitions are:

– **Indistinguishability:** An ideal-world system is indistinguishable from a real-world system. For example, a secure encryption scheme can be seen as realizing a secure channel (ideal world) from an authenticated channel (real world).
– **Game-winning:** Breaking a system means that the adversary must achieve a certain goal, i.e., win a certain game. For example, a MAC is secure if the adversary cannot generate a fresh message together with the correct MAC, even if he can query the system arbitrarily.

The first type of security definition requires to prove that the distinguishing advantage of a certain class of distinguishers for two systems is very small. The second type of security definition requires to prove that no adversary of a certain type can win the game, except with very small probability.

In this paper we establish a tight relation between the above two problems in the information-theoretic setting. More precisely, game-winning can be modeled as an internal monotone condition in a system. Indeed, an important paradigm in indistinguishability proofs is the definition of such an internal monotone condition in a 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

[2] Actually, some results from this paper are known to be false in the computational case under standard assumptions [Pie05].

the probability that $\mathbf{D}$ provokes this condition. A key technical tool of the paper (Lemma 5) is to show that this holds also in the other direction: For two systems $\mathbf{S}$ and $\mathbf{T}$ one can always define new systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$, which are equivalent to $\mathbf{S}$ and $\mathbf{T}$, 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 $\hat{\mathbf{S}}$ (or $\hat{\mathbf{T}}$), and
(ii) the systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$ are equivalent as long as the respective MBOs are 0.

1.4 Related Work and Applications

This section is perhaps best read after reading the technical part of the paper.

Lemma 5 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 was not tight by a logarithmic factor and whose proof was quite technical, based on martingales. This paper settles a main open problem from [MP04], as Lemma 5 is tight.

The product theorem for sequential composition of stateless permutations, implied by Corollary 3, 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 stateless permutations), and do not extend to our general setting. While Vaudenay's decorrelation theory [Vau03] is purely information-theoretic, its application is for the design of actual (computationally secure) block-ciphers. In the same sense, our results have applications in the computational setting, where one considers computationally bounded adversaries.

In the computational setting, a product theorem for the sequential composition of permutations was proved by Luby and Rackoff [LR86]. Myers [Mye03] proved a product theorem[3] for a construction which is basically the parallel composition but with some extra random values XOR-ed to the inputs.

Our stronger results (compared to [MP04]) on adaptive security by composition, namely Corollaries 4 and 5, immediately apply to all results that made use of the bounds of [MP04]. For example, the construction of Kaplan, Naor and Reingold [KNR05] of randomness-efficient constructions of almost k-wise independent permutations, achieve *a priori* only non-adaptive security, but the authors observe that one can apply the results from [MP04] in order to obtain adaptive security. This paper allows to improve the bound of [KNR05]. Another application of Corollary 5 is in the already mentioned decorrelation theory where it implies better security against adaptive attacks, even if the considered block-cipher only satisfies a non-adaptive notion of decorrelation.

The question whether composition implies adaptive security also in the computational setting (i.e. for pseudorandom systems) has been investigated

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

in [Mye04,Pie05]. Unlike for the product amplification results, these attack-strengthening results do not hold for pseudorandom systems in general, though some positive results have also been achieved in this setting [Pie06].

Theorem 2 can be used to prove the adaptive security of more complicated constructions than the sequential and parallel composition considered in this paper. In [MOPS06], (a generalization of) 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 5 of a completely different vain than the problems considered in this paper is given in [PS07], where the security of some constructions for range extension of weak random functions is proven in the information theoretic setting (again, in the computational setting those results no longer hold).

2 Random Systems

This section follows and extends [Mau02], in slightly different notation.

2.1 Random Systems

Essentially every kind of discrete system (say $\mathbf{S}$), in particular a cryptographic system, can be described as follows. It takes inputs $X_1, X_2, \ldots$ (from some alphabet[4] $\mathcal{X}$) and generates, for each new input X_i, an output Y_i (from some alphabet $\mathcal{Y}$). The output Y_i depends (possibly probabilistically) on the current input X_i and on the internal state. Such a system is called an $(\mathcal{X}, \mathcal{Y})$-system.

In most contexts, only the *observable* input-output behavior, but not the internal state representation, is of interest. For example, if one considers the distinguishing advantage of a certain distinguisher $\mathbf{D}$ for two systems $\mathbf{S}$ and $\mathbf{T}$, then all that matters is the input-output behavior of the systems $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{T}$. Hence the input-output behavior is the abstraction of a system that needs to be captured. This is analogous, for example, to a memoryless channel $\mathbf{C}$ in communication theory whose abstraction is captured by a conditional probability distribution $\mathsf{p}^{\mathbf{C}}_{Y|X}$ of the output Y, given the input X, independently of the physical description of the channel. A system is more complex than a channel; what is the abstraction of a (discrete) system?

A system is described exactly by the conditional probability distributions of the ith output Y_i, given $X_1, \ldots, X_i$ and $Y_1, \ldots, Y_{i-1}$, for all i. We use the shorthand notation $X^i := [X_1, \ldots, X_i]$. This is captured by the following definition from [Mau02].

[4] It is not a restriction to consider fixed input and output alphabets. This allows to model also systems where inputs and outputs come from different alphabets for different i.

Definition 2. An $(\mathcal{X},\mathcal{Y})$-*random*[5] *system* $\mathbf{S}$ is a (generally infinite) sequence of conditional probability distributions[6] $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}$ for $i \geq 1$.[7]

This description of a system is exact and minimal in the sense that two systems with different input-output behavior correspond to two different random systems, and two different random systems have different input-output behavior.

Note that the name $\mathbf{S}$ is used interchangeably for a system $\mathbf{S}$ (which can be described arbitrarily, for example by its internal workings) and the corresponding random system. This should cause no confusion. It is therefore also meaningful to say that two systems are equivalent if they have the same behavior, even though their internal structure may be different.

Definition 3. Two systems $\mathbf{S}$ and $\mathbf{T}$ are *equivalent*, denoted $\mathbf{S} \equiv \mathbf{T}$, if they correspond to the same random system, i.e., if for all $i \geq 1$[8]

$$\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}} = \mathsf{p}^{\mathbf{T}}_{Y_i|X^iY^{i-1}}.$$

The results of this paper are stated for random systems, but we emphasize that they hold for arbitrary systems, as the only property of a system that is relevant here is the input-output behavior. When several random systems appear in the same random experiment, they are (tacitly) assumed to be independent. In a more general theory, random systems could be dependent.

A random system $\mathbf{S}$ can be characterized equivalently by the sequence $\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}$, for $i \geq 1$, of conditional probability distributions. This description is often convenient, but is not minimal.[9] The conversion between the two forms is given by

$$\mathsf{p}^{\mathbf{S}}_{Y^i|X^i} = \prod_{j=1}^{i} \mathsf{p}^{\mathbf{S}}_{Y_j|X^jY^{j-1}} \qquad \text{and} \qquad \mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}} = \frac{\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}}{\mathsf{p}^{\mathbf{S}}_{Y^{i-1}|X^{i-1}}}. \tag{1}$$

$\mathbf{S}$ and $\mathbf{T}$ are equivalent if and only if $\mathsf{p}^{\mathbf{S}}_{Y^i|X^i} = \mathsf{p}^{\mathbf{T}}_{Y^i|X^i}$ for $i \geq 1$.

2.2 Special Random Systems

Definition 4. A *random function* $\mathcal{X} \to \mathcal{Y}$ is a random system which answers consistently in the sense that $X_i = X_j \implies Y_i = Y_j$. A random function is

[5] 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.

[6] We use a lower-case p to stress the fact that these conditional distributions by themselves do not define a random experiment in which probabilities are defined.

[7] For arguments x^{i-1} and y^{i-1} such that $\mathsf{p}^{\mathbf{S}}_{Y^{i-1}|X^{i-1}}(y^{i-1}, x^{i-1}) = 0$, $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}$ need not be defined.

[8] This equality is an equality of (partial) functions, where two conditional probability distributions are considered to be equal if they are equal for all arguments for which both are defined.

[9] The distributions $\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}$ must satisfy a consistency condition for the different i.

stateless if it corresponds to a random variable taking on as values function tables $\mathcal{X} \to \mathcal{Y}$. A *random permutation* on $\mathcal{X}$ is a random function $\mathcal{X} \to \mathcal{X}$ mapping distinct inputs to distinct outputs: $X_i \neq X_j \Longrightarrow Y_i \neq Y_j$.

Note that in general a random function is not stateless. For example, a system defined by $Y_i = X_1$ for all i is not stateless.

We discuss a few examples of random systems.

Example 2. A *$\mathcal{Y}$-beacon*, usually denoted as $\mathbf{B}$, is a random system which outputs a new independent uniformly (over $\mathcal{Y}$) output Y_i for every new input X_i: $\mathsf{p}^{\mathbf{B}}_{Y_i|X^iY^{i-1}} = 1/|\mathcal{Y}|$ for all choices of the arguments.

Example 3. A *uniform random function*, usually denoted as $\mathbf{R}$, from some domain $\mathcal{X}$ to some finite range $\mathcal{Y}$. Typically $\mathcal{X} = \{0,1\}^m$ for some m or $\mathcal{X} = \{0,1\}^*$, and $\mathcal{Y} = \{0,1\}^n$ for some n. If $\mathcal{X}$ is finite, then this corresponds to a randomly selected function table. We have

$$\mathsf{p}^{\mathbf{R}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1}) = \begin{cases} 1 & \text{if } x_i = x_j \text{ for some } j < i \text{ and } y_i = y_j \\ 0 & \text{if } x_i = x_j \text{ for some } j < i \text{ and } y_i \neq y_j \\ 1/|\mathcal{Y}| & \text{else.} \end{cases}$$

$\mathsf{p}^{\mathbf{R}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1})$ is undefined if $x_j = x_k$ and $y_j \neq y_k$ for $j < k < i$.

We point out that when analyzing constructions involving uniform random functions (or other random systems), there is no need to resort to this apparently complex description. Any complete description is fine. Using the concept of random systems buys precision and simplicity, without requiring technical complexity of the arguments.

Example 4. A *uniform random permutation*, usually denoted as $\mathbf{P}$, for domain and range $\mathcal{X}$, is a function randomly selected from all bijective functions $\mathcal{X} \to \mathcal{X}$.

2.3 Distinguishing Random Systems

We are interested in distinguishing two systems $\mathbf{S}$ and $\mathbf{T}$ by means of a distinguisher $\mathbf{D}$. In the sequel, we will usually tacitly assume that the two systems are compatible, i.e., have the same input and output alphabets.

A distinguisher $\mathbf{D}$ for distinguishing two $(\mathcal{X}, \mathcal{Y})$-systems generates X_1 as an input, receives the output Y_1, then generates X_2, receives Y_2, etc. Finally, after receiving Y_k, it outputs a binary decision bit, say W. More formally:

Definition 5. A *distinguisher* $\mathbf{D}$ *for* $(\mathcal{X}, \mathcal{Y})$*-random systems* 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 $\mathsf{p}^{\mathbf{D}}_{X_i|Y^iX^{i-1}}$) for all i.[10] $\mathbf{D}$ outputs a bit W after a certain number k of queries, based on the transcript (X^k, Y^k).

[10] In particular the first output $\mathsf{p}^{\mathbf{D}}_{X_1}$ is defined before $\mathbf{D}$ is fed with any input.

When a distinguisher $\mathbf{D}$ is connected to a system $\mathbf{S}$, which we denote simply as $\mathbf{DS}$, this defines a random experiment. The probabilities of an event $\mathcal{E}$ in this experiment will be denoted as $\mathsf{P}^{\mathbf{DS}}(\mathcal{E})$. We note that the probability distribution $\mathsf{P}^{\mathbf{DS}}_{X^kY^k}$ can be expressed by

$$\begin{aligned}\mathsf{P}^{\mathbf{DS}}_{X^kY^k}(x^k,y^k) &= \prod_{i=1}^{k} \mathsf{p}^{\mathbf{D}}_{X_i|X^{i-1}Y^{i-1}}(x_i,x^{i-1},y^{i-1})\ \mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}(y_i,x^i,y^{i-1}) \\ &= \mathsf{p}^{\mathbf{D}}_{X^k|Y^{k-1}}(x^k,y^{k-1})\ \mathsf{p}^{\mathbf{S}}_{Y^k|X^k}(y^k,x^k)\ , \end{aligned} \tag{2}$$

where the last equality follows from (1).

The performance of a distinguisher, called the advantage, can be defined in two equivalent ways, both of which will be useful for us. We first state the standard definition.

Definition 6. The *advantage* of distinguisher $\mathbf{D}$ for random systems $\mathbf{S}$ and $\mathbf{T}$, for k queries, denoted $\Delta^{\mathbf{D}}_k(\mathbf{S},\mathbf{T})$, is defined as

$$\Delta^{\mathbf{D}}_k(\mathbf{S},\mathbf{T}) := \left|\mathsf{P}^{\mathbf{DS}}(W=1) - \mathsf{P}^{\mathbf{DT}}(W=1)\right|.$$

For a class $\mathcal{D}$ of distinguishers, the advantage of the best $\mathbf{D}$ in $\mathcal{D}$, asking at most k queries, is denoted as

$$\Delta^{\mathcal{D}}_k(\mathbf{S},\mathbf{T}) := \max_{\mathbf{D}\in\mathcal{D}} \Delta^{\mathbf{D}}_k(\mathbf{S},\mathbf{T}).$$

For the class of *all* distinguishers we simply write $\Delta_k(\mathbf{S},\mathbf{T})$.

To state an equivalent definition of the advantage we need the following definition.

Definition 7. For two compatible systems $\mathbf{S}$ and $\mathbf{T}$, $\langle \mathbf{S}/\mathbf{T}\rangle$ denotes the random system which is equal to system $\mathbf{S}$ or $\mathbf{T}$ with probability $\frac{1}{2}$ each. To make the independent unbiased binary random variable, say Z, selecting between $\mathbf{S}$ (for $Z=0$) and $\mathbf{T}$ (for $Z=1$) explicit, we write $\langle \mathbf{S}/\mathbf{T}\rangle_Z$.[11]

The advantage $\Delta^{\mathbf{D}}_k(\mathbf{S},\mathbf{T})$ can be defined equivalently in terms of the probability that $\mathbf{D}$, interacting with the mixed system $\langle \mathbf{S}/\mathbf{T}\rangle_Z$, guesses Z correctly:

Lemma 2. *For every distinguisher* $\mathbf{D}$,[12]

$$\Delta^{\mathbf{D}}_k(\mathbf{S},\mathbf{T}) = 2\left|\mathsf{P}^{\mathbf{D}\langle \mathbf{S}/\mathbf{T}\rangle_Z}(W=Z) - \tfrac{1}{2}\right|.$$

[11] It is helpful to think of Z as the position of a switch selecting between the systems $\mathbf{S}$ and $\mathbf{T}$.

[12] The normalization factor 2 assures that the advantage is between 0 and 1. The absolute value in $|\mathsf{P}^{\mathbf{D}\langle \mathbf{S}/\mathbf{T}\rangle_Z}(W=Z)-\frac{1}{2}|$ takes into account the fact that one can always invert the output of a distinguisher whose success probability is below $\frac{1}{2}$.

Proof. Let p_z for $z \in \{0,1\}$ denote the probability that $W = 1$ if $Z = z$. Then $\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) = |p_0 - p_1|$ and $\mathsf{P}^{\mathbf{D}\langle\mathbf{S}/\mathbf{T}\rangle_Z}(W = Z) = \frac{1}{2}(1 - p_0 + p_1)$, hence $2\left|\mathsf{P}^{\mathbf{D}\langle\mathbf{S}/\mathbf{T}\rangle_Z}(W = Z) - \frac{1}{2}\right| = |p_0 - p_1|$. □

The following distinguisher classes are usually of special interest:

Definition 8. By NA we denote the class of computationally unbounded non-adaptive distinguishers which select all queries $X_1, \ldots, X_k$ in advance (i.e., independent of the outputs Y_i).[13] By RI we denote the class of computationally unbounded distinguishers which (cannot select the queries but) are given uniformly random values $X_1, \ldots, X_k$ (and the corresponding outputs $Y_1, \ldots, Y_k$).

Clearly, $\mathsf{RI} \subseteq \mathsf{NA}$. The class NA is sometimes called nCPA (non-adaptive chosen-plaintext attack) in the literature and the class RI is sometimes called KPA (known-plaintext attack).

The following lemma captures the simple fact that if one has to distinguish the systems $\mathbf{S}$ and $\langle\mathbf{S}/\mathbf{T}\rangle_Z$, then the advantage is only half of the advantage when distinguishing $\mathbf{S}$ and $\mathbf{T}$. In a sense, $\langle\mathbf{S}/\mathbf{T}\rangle_Z$ is half-way between $\mathbf{S}$ and $\mathbf{T}$.

Lemma 3. *For every* $\mathbf{D}$, $\Delta_k^{\mathbf{D}}(\mathbf{S}, \langle\mathbf{S}/\mathbf{T}\rangle_Z) = \frac{1}{2}\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$.

Proof. This follows from the linearity of the probability of $\mathbf{D}$ outputting a 1: we have $\mathsf{P}^{\mathbf{D}\langle\mathbf{S}/\mathbf{T}\rangle_Z}(W = 1) = \frac{1}{2}(\mathsf{P}^{\mathbf{DS}}(W = 1) + \mathsf{P}^{\mathbf{DT}}(W = 1))$.

2.4 Game-winning and Monotone Binary Outputs

An important paradigm in certain security definitions is the notion of winning a game. Without loss of generality, a game with one player (e.g. the adversary) can be described by an $(\mathcal{X},\mathcal{Y})$-system which interacts with its environment by taking inputs $X_1, X_2, \ldots$ (considered as moves) and answering with outputs $Y_1, Y_2, \ldots$. In addition, after every input it also outputs a bit indicating whether the game has been won. This bit is monotone in the sense that it is initially set to 0 and that, once it has turned to 1 (the game is won), it can not turn back to 0. This motivates the following definition, which captures the notion of game-winning.

Definition 9. For a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-system $\mathbf{S}$ the binary component A_i of the output (Y_i, A_i) is called a *monotone binary output (MBO)* if $A_i = 1$ implies $A_j = 1$ for $j \geq i$. For such a system $\mathbf{S}$ with MBO we define two derived systems:

(i) $\mathbf{S}^-$ is the $(\mathcal{X},\mathcal{Y})$-system resulting from $\mathbf{S}$ by ignoring the MBO.
(ii) $\mathbf{S}^{\dashv}$ is the $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-system which masks the $\mathcal{Y}$-output to a dummy symbol ($\perp$) as soon as the MBO turns to 1. More precisely, the following function is applied to the outputs of $\mathbf{S}$:

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

[13] One can view such a distinguisher as making a single (compound) query $(x_1, \ldots, x_k)$.

Definition 10. Two systems $\mathbf{S}$ and $\mathbf{T}$ with MBOs are called *restricted equivalent* if $\mathbf{S}^{\dashv} \equiv \mathbf{T}^{\dashv}$, i.e., if they are equivalent as long as the MBO is 0.

A system (or player) $\mathbf{D}$ interacting with $\mathbf{S}$, trying to win the game defined by $\mathbf{S}$, is like a distinguisher, except that it need not have a binary output W. Whether or not $\mathbf{D}$ "sees" the MBO is irrelevant; one can think of $\mathbf{D}$ interacting with $\mathbf{S}^-$ instead of $\mathbf{S}$. One could call such a $\mathbf{D}$ a "player" or a "provoker", as it tries to provoke the MBO to become 1, but for consistency we will continue to call $\mathbf{D}$ a distinguisher.

Definition 11. For a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-random system $\mathbf{S}$ with an MBO (called A_i) and for a distinguisher $\mathbf{D}$, we denote with $\nu_k^{\mathbf{D}}(\mathbf{S})$ the probability that $\mathbf{D}$ wins the game within k queries:

$$\nu_k^{\mathbf{D}}(\mathbf{S}) := \mathsf{P}^{\mathbf{DS}}(A_k = 1).$$

For a class $\mathcal{D}$ of distinguishers, the winning probability of the best $\mathbf{D}$ in $\mathcal{D}$ within k queries is denoted as

$$\nu_k^{\mathcal{D}}(\mathbf{S}) := \max_{\mathbf{D} \in \mathcal{D}} \nu_k^{\mathbf{D}}(\mathbf{S}).$$

For the class of *all* distinguishers we simply write $\nu_k(\mathbf{S})$.

3 Relating Indistinguishability and Game-winning

3.1 From Game-winning to Indistinguishability

The following lemma was proved in [Mau02]. Versions of this lemma for special types of systems appeared subsequently.

Lemma 4. *Let $\mathbf{S}$ and $\mathbf{T}$ be two $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-random systems with MBOs. If $\mathbf{S}^{\dashv} \equiv \mathbf{T}^{\dashv}$, then*

$$\Delta_k^{\mathbf{D}}(\mathbf{S}^-, \mathbf{T}^-) \;\leq\; \nu_k^{\mathbf{D}}(\mathbf{S}) \;=\; \nu_k^{\mathbf{D}}(\mathbf{T})$$

for all distinguishers $\mathbf{D}$ for $(\mathcal{X}, \mathcal{Y})$-random systems.[14] *In particular, for any distinguisher class $\mathcal{D}$, $\Delta_k^{\mathcal{D}}(\mathbf{S}^-, \mathbf{T}^-) \leq \nu_k^{\mathcal{D}}(\mathbf{S})$, hence $\Delta_k(\mathbf{S}^-, \mathbf{T}^-) \leq \nu_k(\mathbf{S})$ and $\Delta_k^{\mathsf{NA}}(\mathbf{S}^-, \mathbf{T}^-) \leq \nu_k^{\mathsf{NA}}(\mathbf{S})$.*

Proof. According to Lemma 2, $\Delta_k^{\mathbf{D}}(\mathbf{S}, \mathbf{T})$ can be computed in terms of the probability that $\mathbf{D}$ guesses the switch Z in $\langle \mathbf{S}/\mathbf{T} \rangle_Z$ correctly. The condition $\mathbf{S}^{\dashv} \equiv \mathbf{T}^{\dashv}$ implies that if the MBO of $\langle \mathbf{S}/\mathbf{T} \rangle_Z$ is 0, then the output of $\langle \mathbf{S}/\mathbf{T} \rangle_Z$ is independent of Z, and therefore in this case $\mathbf{D}$ cannot do better than guess randomly. (If the MBO is 1, the success probability is bounded by 1.) Hence, if we denote by p the probability that $\mathbf{D}$ sets the MBO to 1, the probability that $\mathbf{D}$ guesses Z correctly is bounded by $\frac{1}{2}(1-p)+p = \frac{1}{2} + \frac{1}{2}p$, where $p = \nu_k^{\mathbf{D}}(\langle \mathbf{S}/\mathbf{T} \rangle_Z) = \nu_k^{\mathbf{D}}(\mathbf{S}) = \nu_k^{\mathbf{D}}(\mathbf{T})$. Applying Lemma 2 completes the proof. $\square$

[14] Recall that it is well-defined what it means for such a distinguisher to play the game for $\mathbf{S}$ which is defined *with* an MBO.

3.2 From Indistinguishability to Game-winning

The following lemma states, in a certain sense, a converse to Lemma 4, and is a key tool for the proofs of the main results. While Lemma 4 holds for every distinguisher, whether computationally bounded or not, and whether or not its binary output is determined optimally based on the transcript, the converse only holds in the information-theoretic setting and if we assume that the decision bit is computed optimally. More precisely, it is a statement about the statistical distance of transcripts.

Definition 12. Let

$$\delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) := \|\mathsf{P}^{\mathbf{DS}}_{X^kY^k} - \mathsf{P}^{\mathbf{DT}}_{X^kY^k}\|$$

be the statistical distance of the transcripts (X^kY^k) when $\mathbf{D}$ interacts with $\mathbf{S}$ and $\mathbf{T}$, respectively. For a class $\mathcal{D}$ of distinguishers we define[15]

$$\delta_k^{\mathcal{D}}(\mathbf{S},\mathbf{T}) := \max_{\mathbf{D}\in\mathcal{D}} \delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}).$$

Note that in general we have $\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) \leq \delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$, but for a computationally unbounded distinguisher $\mathbf{D}$ that chooses the output bit optimally, we have $\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) = \delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$. In particular,

$$\Delta_k(\mathbf{S},\mathbf{T}) = \delta_k(\mathbf{S},\mathbf{T}) \quad \text{and} \quad \Delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T}) = \delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T}).$$

Lemma 5. *For any two $(\mathcal{X},\mathcal{Y})$-systems $\mathbf{S}$ and $\mathbf{T}$ there exist $(\mathcal{X},\mathcal{Y}\times\{0,1\})$-random systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$ with MBOs such that*

(i) $\hat{\mathbf{S}}^- \equiv \mathbf{S}$,
(ii) $\hat{\mathbf{T}}^- \equiv \mathbf{T}$,
(iii) $\hat{\mathbf{S}}^{\dashv} \equiv \hat{\mathbf{T}}^{\dashv}$, *and*
(iv) $\delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{S}}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{T}})$ *for all* $\mathbf{D}$.[16]

To illustrate the idea of the proof of Lemma 5, we consider an analogous statement (in fact, a special case) where probability distributions P_X and Q_X (over some alphabet $\mathcal{X}$) take the place of the random systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$. In this case, the systems with MBO can be replaced by joint distributions $\hat{P}_{XA}$ and $\hat{Q}_{XA}$, where A is binary. Indeed, if we define these distributions by

$$\begin{aligned}
\hat{P}_{XA}(x,0) &= \hat{Q}_{XA}(x,0) = \min(P_X(x),Q_X(x))\\
\hat{P}_{XA}(x,1) &= P_X(x) - \min(P_X(x),Q_X(x))\\
\hat{Q}_{XA}(x,1) &= Q_X(x) - \min(P_X(x),Q_X(x))
\end{aligned}$$

(for any $x \in \mathcal{X}$) it is easy to verify that $\hat{P}_X = P_X$ and $\hat{Q}_X = Q_X$, which corresponds to (i) and (ii), respectively. Furthermore, and trivially, $\hat{P}_{XA}(\cdot,0) =$

[15] For the class of *all* distinguishers we simply write $\delta_k(\mathbf{S},\mathbf{T})$.

[16] This also implies, for example, $\Delta_k(\mathbf{S},\mathbf{T}) = \nu_k(\hat{\mathbf{S}})$ and $\Delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T}) = \nu_k^{\mathsf{NA}}(\hat{\mathbf{S}})$.

$\hat{Q}_{XA}(\cdot, 0)$, which is (iii). Finally, because the statistical distance can be written as

$$\delta(P_X, Q_X) = 1 - \sum_x \min(P_X(x), Q_X(x)) \ , \tag{3}$$

the equivalent of (iv) follows from the fact that the right-hand side of (3) equals $\hat{P}_A(1) = \hat{Q}_A(1)$.

Proof (of Lemma 5). The idea is to define the system $\hat{\mathbf{S}}$ with MBO A_i (and, likewise, $\hat{\mathbf{T}}$) such that, for all $i \geq 1$,

$$\begin{aligned} \mathsf{p}^{\hat{\mathbf{S}}}_{Y^i A_i | X^i}(y^i, 0, x^i) &:= m_{x^i, y^i} \\ \mathsf{p}^{\hat{\mathbf{S}}}_{Y^i A_i | X^i}(y^i, 1, x^i) &:= \mathsf{p}^{\mathbf{S}}_{Y^i | X^i}(y^i, x^i) - m_{x^i, y^i} , \end{aligned} \tag{4}$$

where

$$m_{x^i, y^i} := \min(\mathsf{p}^{\mathbf{S}}_{Y^i | X^i}(y^i, x^i), \mathsf{p}^{\mathbf{T}}_{Y^i | X^i}(y^i, x^i)) \ .$$

We will verify below that this can always be done consistently.

Note that properties (i), (ii), and (iii) follow immediately from these equations (similarly to the above argument for random variables). To verify (iv), we recall that the probabilities of $\mathsf{P}^{\mathbf{DS}}_{X^k Y^k}$ (and, likewise, $\mathsf{P}^{\mathbf{DT}}_{X^k Y^k}$) can be expressed by equation (2). Using formula (3) for the statistical distance we find

$$\begin{aligned} \delta^{\mathbf{D}}_k(\mathbf{S}, \mathbf{T}) &= \| \mathsf{P}^{\mathbf{DS}}_{X^k Y^k} - \mathsf{P}^{\mathbf{DT}}_{X^k Y^k} \| \\ &= 1 - \sum_{x^k, y^k} \min\bigl(\mathsf{P}^{\mathbf{DS}}_{X^k Y^k}(x^k, y^k), \mathsf{P}^{\mathbf{DT}}_{X^k Y^k}(x^k, y^k)\bigr) \\ &= 1 - \sum_{x^k, y^k} \mathsf{p}^{\mathbf{D}}_{X^k | Y^{k-1}}(x^k, y^{k-1}) \min\bigl(\mathsf{p}^{\mathbf{S}}_{Y^k | X^k}(y^k, x^k), \mathsf{p}^{\mathbf{T}}_{Y^k | X^k}(y^k, x^k)\bigr) \ . \end{aligned}$$

Property (iv) then follows because the probability that the MBO A_k of $\hat{\mathbf{S}}$ (and, likewise, $\hat{\mathbf{T}}$) equals 1 after k steps is given by

$$\begin{aligned} \nu^{\mathbf{D}}_k(\hat{\mathbf{S}}) &= 1 - \sum_{x^k, y^k} \mathsf{P}^{\mathbf{D}\hat{\mathbf{S}}}_{X^k Y^k A_k (x^k, y^k, 0)} \\ &= 1 - \sum_{x^k, y^k} \mathsf{p}^{\mathbf{D}}_{X^k | Y^{k-1}}(x^k, y^{k-1}) \mathsf{p}^{\hat{\mathbf{S}}}_{Y^k A_k | X^k}(y^k, 0, x^k) \ , \end{aligned}$$

which equals the above expression for $\delta^{\mathbf{D}}_k(\mathbf{S}, \mathbf{T})$.

It remains to verify that there exists a system $\hat{\mathbf{S}}$ satisfying (4) (the argument for $\hat{\mathbf{T}}$ follows by symmetry).

Note that (4) only determines the interrelation between the system's output Y_i and the value A_i of the MBO at the same step, but it does not specify the dependency on previous values A^{i-1}. In fact, there are various degrees of freedom in the definition of $\hat{\mathbf{S}}$, for instance in the choice of the probabilities

$r_{x^i,y^i} := \mathsf{p}^{\hat{\mathbf{S}}}_{Y^i A^{i-1}|X^i}(y^i, 0^{i-1}, x^i)$. Most generally, the probabilities defining $\hat{\mathbf{S}}$, conditioned on the event that the previous MBO equals 0, can be written as[17]

$$\mathsf{p}^{\hat{\mathbf{S}}}_{Y_i A_i|X^i Y^{i-1} A^{i-1}}(y_i, a_i, x^i, y^{i-1}, 0^{i-1}) := \begin{cases} \frac{m_{x^i,y^i}}{m_{x^{i-1},y^{i-1}}} & \text{if } a_i = 0 \\ \frac{r_{x^i,y^i} - m_{x^i,y^i}}{m_{x^{i-1},y^{i-1}}} & \text{if } a_i = 1, \end{cases}$$

for any $i \geq 1$, where $r_{x_i,y_i} \in [m_{x^i,y^i}, \mathsf{p}^{\mathbf{S}}_{Y^i|X^i}(y^i, x^i)]$ are parameters. To make sure that the conditional probabilities sum up to 1, we require

$$\sum_{y_i} r_{x^i,y^i} = m_{x^{i-1},y^{i-1}} \ , \tag{5}$$

for any fixed x^i and y^{i-1}. Note that such a choice of r_{x^i,y^i} always exists because the right side of (5) lies in the interval

$$m_{x^{i-1},y^{i-1}} \in \left[\sum_{y_i} m_{x^i,y^i}, \sum_{y_i} \mathsf{p}^{\mathbf{S}}_{Y^i|X^i}(y^i, x^i)\right] \ .$$

To complete the definition of $\hat{\mathbf{S}}$, we set, for any $i > 1$ and $a^{i-1} \neq 0$,

$$\mathsf{p}^{\hat{\mathbf{S}}}_{Y_i A_i|X^i Y^{i-1} A^{i-1}}(y_i, 1, x^i, y^{i-1}, a^{i-1}) := \frac{\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}(y^i, x^i) - r_{x^i,y^i}}{\mathsf{p}^{\mathbf{S}}_{Y^{i-1}|X^{i-1}}(y^{i-1}, x^{i-1}) - m_{x^{i-1},y^{i-1}}} \ .$$

Again, the conditional probabilities are well-defined because all values are non-negative and, by (5), sum up to 1. Furthermore, it is easy to see that the outputs A_i of $\hat{\mathbf{S}}$ are indeed monotone. Finally, by induction over i, it is straightforward to verify that $\hat{\mathbf{S}}$ satisfies (4), which concludes the proof. $\square$

We give another interpretation of Lemma 5. If two probability distributions P_X and Q_X have statistical distance δ then there exists a (common) random experiment with two random variables X' and X'', distributed according to P_X and Q_X, respectively, such that $X' = X''$ with probability $1 - \delta$. Lemma 5 can be interpreted as the generalization of this statement to random systems. For any distinguisher $\mathbf{D}$, two random systems $\mathbf{S}$ and $\mathbf{T}$ are equal with probability $1 - \delta$, where δ is $\mathbf{D}$'s distinguishing advantage.

4 Amplification of the Distinguishing Advantage

4.1 Neutralizing Constructions

Throughout the rest of the paper we let $\mathbf{C}(\cdot,\cdot)$ be a construction invoking two systems. For example $\mathbf{C}(\mathbf{F}, \mathbf{G})$ denotes the system obtained when $\mathbf{C}(\cdot,\cdot)$ invokes the two systems $\mathbf{F}$ and $\mathbf{G}$.

[17] We use the convention $\mathsf{p}^{\mathbf{S}}_{Y^0|X^0} \equiv 1$ and, in particular, $m_{x^0,y^0} = 1$.

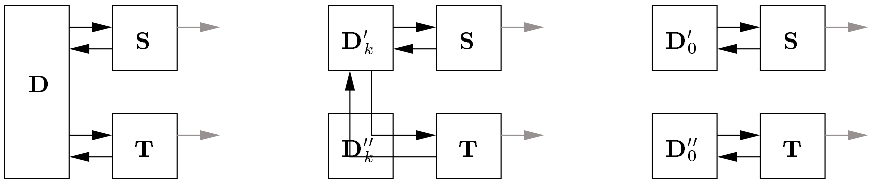

Figure1. Illustration for the proof of Lemma 6. $\mathbf{D}$ can be seen as a pair $(\mathbf{D}'_k, \mathbf{D}''_k)$ of distinguishers which can exchange up to $k = 2k''$ messages (simply set $\mathbf{D}'_k \equiv \mathbf{D}$ and $\mathbf{D}''_k$ to be the trivial system which only passes messages). The gray arrows indicate the MBOs.

Definition 13. A construction $\mathbf{C}(\cdot,\cdot)$ is called *neutralizing* for the pairs $(\mathbf{F},\mathbf{I})$ and $(\mathbf{G},\mathbf{J})$ of (independent) systems if

$$\mathbf{C}(\mathbf{F},\mathbf{J}) \equiv \mathbf{C}(\mathbf{I},\mathbf{G}) \equiv \mathbf{C}(\mathbf{I},\mathbf{J}) \equiv \mathbf{Q} \tag{6}$$

(for some $\mathbf{Q}$). Moreover, we denote by k' and k'' the maximal number of queries made to the first and the second subsystem, respectively, when the number of queries to $\mathbf{C}(\cdot,\cdot)$ is k.

4.2 Winning Independent Games

The following lemma states that the best combined strategy for winning two independent games is not better than applying the individually best strategies separately. We note that this is (of course) also true for real games, like playing black jack, but we phrase the result at an abstract (and hence very general) level.

We need some new notation: For two systems $\mathbf{S}$ and $\mathbf{T}$ with MBOs let $[\mathbf{S}\,\|\,\mathbf{T}]^\wedge$ be the system consisting of $\mathbf{S}$ and $\mathbf{T}$ being accessible independently, with an MBO which is 1 if and only if the MBOs of $\mathbf{S}$ and $\mathbf{T}$ are *both* 1. Let $\nu^{\mathcal{D}}_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge)$ denote the advantage of the best distinguisher in $\mathcal{D}$, making k' and k'' (arbitrarily scheduled) queries to $\mathbf{S}$ and $\mathbf{T}$, respectively, in setting the MBO to 1 (we simply write $\nu_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge)$ if $\mathcal{D}$ is the class of all distinguishers).

Lemma 6. *For any random systems* $\mathbf{S}$ *and* $\mathbf{T}$ *with MBOs, and any* k' *and* k''*,*

$$\nu_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge) \;=\; \nu_{k'}(\mathbf{S})\,\nu_{k''}(\mathbf{T}), \tag{7}$$

and

$$\nu^{\mathsf{NA}}_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge) \;=\; \nu^{\mathsf{NA}}_{k'}(\mathbf{S})\,\nu^{\mathsf{NA}}_{k''}(\mathbf{T}). \tag{8}$$

Proof. The non-adaptive case (8) follows from the adaptive case (7) by viewing the non-adaptive queries as a single adaptive query. To prove (7), let $\mathbf{D}$ be an optimal distinguisher for the task considered, i.e.

$$\nu_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge) \;=\; \nu^{\mathbf{D}}_{k',k''}([\mathbf{S}\,\|\,\mathbf{T}]^\wedge).$$

Let $A_1, \ldots, A_{k'}$ and $B_1, \ldots, B_{k''}$ denote the MBOs of $\mathbf{S}$ and $\mathbf{T}$, respectively. We can interpret $\mathbf{D}$ as a pair $(\mathbf{D}'_k, \mathbf{D}''_k)$ of distinguishers which can exchange up to $k = 2k''$ messages with each other, as shown in Figure 1. 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 $(\mathbf{D}'_k, \mathbf{D}''_k)$.

Now assume that there is a pair of 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 claim that then there also exist 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

$$\nu_{k',k''}([\mathbf{S} \| \mathbf{T}]^\wedge) \leq \nu_{k'}^{\mathbf{D}'_0}(\mathbf{S}) \cdot \nu_{k''}^{\mathbf{D}''_0}(\mathbf{T}) \leq \nu_{k'}(\mathbf{S}) \cdot \nu_{k''}(\mathbf{T}).$$

We actually have equality above as the other direction ($\geq$) is trivial. To prove the claim, assume that 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

$$\sum_{m,v} \mathsf{P}^{\mathcal{E}}[A_{k'} = 1 \wedge M = m \wedge V = v] \cdot \mathsf{P}^{\mathcal{E}}[B_{k''} = 1 | M = m \wedge V = v]. \tag{9}$$

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 and 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 sent (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 (9). $\square$

4.3 The Product Theorem

We can now state the first main result of the paper. Recall Definition 13.

Theorem 1. *If $\mathbf{C}(\cdot,\cdot)$ is neutralizing for the pairs $(\mathbf{F},\mathbf{I})$ and $(\mathbf{G},\mathbf{J})$ of systems, then, for all k,*

$$\Delta_k(\mathbf{C}(\mathbf{F},\mathbf{G}), \mathbf{C}(\mathbf{I},\mathbf{J})) \;\leq\; 2\, \Delta_{k'}(\mathbf{F},\mathbf{I})\, \Delta_{k''}(\mathbf{G},\mathbf{J}).$$

Proof. We consider the systems $\mathbf{H}_{Z,Z'} := \mathbf{C}(\langle \mathbf{I}/\mathbf{F}\rangle_Z, \langle \mathbf{J}/\mathbf{G}\rangle_{Z'})$, indexed by Z and Z', where Z and Z' are independent unbiased bits. Due to (6) we have $\mathbf{H}_{11} \equiv \mathbf{C}(\mathbf{F},\mathbf{G})$ and $\mathbf{H}_{00} \equiv \mathbf{H}_{01} \equiv \mathbf{H}_{10} \equiv \mathbf{Q} \equiv \mathbf{C}(\mathbf{I},\mathbf{J})$. One can hence easily verify that

$$\mathbf{H}_{Z,Z'} \equiv \langle\langle \mathbf{Q}/\mathbf{C}(\mathbf{F},\mathbf{G})\rangle_{Z'}/\mathbf{Q}\rangle_{Z\oplus Z'},$$

by checking the equivalence for all four values of the pair (Z, Z').

Lemma 3 implies that $\Delta_k(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{Q}) = 2 \cdot \Delta_k(\langle \mathbf{Q}/\mathbf{C}(\mathbf{F},\mathbf{G})\rangle_{Z'},\mathbf{Q})$, where, according to Lemma 2, $\Delta_k(\langle \mathbf{Q}/\mathbf{C}(\mathbf{F},\mathbf{G})\rangle_{Z'},\mathbf{Q})$ is equal to the optimal advantage in guessing $Z \oplus Z'$ with k queries to $\mathbf{H}_{Z,Z'}$, since Z' and $Z \oplus Z'$ are independent unbiased bits. For the analysis of this advantage we consider the form $\mathbf{H}_{Z,Z'} = \mathbf{C}(\langle \mathbf{I}/\mathbf{F}\rangle_Z, \langle \mathbf{J}/\mathbf{G}\rangle_{Z'})$.

Let $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$ be defined as guaranteed by Lemma 5, where $\hat{\mathbf{F}}^- \equiv \mathbf{F}$, $\hat{\mathbf{I}}^- \equiv \mathbf{I}$, $\hat{\mathbf{F}}^\dashv \equiv \hat{\mathbf{I}}^\dashv$, and $\delta_{k'}(\mathbf{F},\mathbf{I}) = \Delta_{k'}(\mathbf{F},\mathbf{I}) = \nu_{k'}(\hat{\mathbf{F}})$. Similarly, let $\hat{\mathbf{G}}$ and $\hat{\mathbf{J}}$ be defined such that $\hat{\mathbf{G}}^- \equiv \mathbf{G}$, $\hat{\mathbf{J}}^- \equiv \mathbf{J}$, $\hat{\mathbf{G}}^\dashv \equiv \hat{\mathbf{J}}^\dashv$, and $\delta_{k''}(\mathbf{G},\mathbf{J}) = \Delta_{k''}(\mathbf{G},\mathbf{J}) = \nu_{k''}(\hat{\mathbf{G}})$. We define the system

$$\hat{\mathbf{H}}_{Z,Z'} := \mathbf{C}(\langle \hat{\mathbf{I}}/\hat{\mathbf{F}}\rangle_Z, \langle \hat{\mathbf{J}}/\hat{\mathbf{G}}\rangle_{Z'})$$

with two MBOs. If we define $\hat{\mathbf{H}}^-_{Z,Z'}$ as $\hat{\mathbf{H}}_{Z,Z'}$ with *both* MBOs ignored, then $\hat{\mathbf{H}}^-_{Z,Z'} \equiv \mathbf{H}_{Z,Z'}$.

Since the MBOs can always be ignored, guessing $Z \oplus Z'$ can only become easier in $\hat{\mathbf{H}}_{Z,Z'}$ (compared to $\mathbf{H}_{Z,Z'}$.) If we assume further that whenever an MBO turns to 1, the corresponding bit (Z or Z') is also output (i.e., given to the distinguisher for free), this can only improve the advantage further.

If either MBO is 0, the advantage in guessing that bit (Z or Z') is 0, and hence also the advantage in guessing $Z \oplus Z'$ is 0. Thus the optimal strategy for guessing $Z \oplus Z'$ is to provoke *both* MBOs (i.e., win both games), and the probability that this succeeds is the advantage in guessing $Z \oplus Z'$.

We can now consider making the distinguisher's task even easier. Instead of having to provoke the two MBOs in the system $\hat{\mathbf{H}}_{Z,Z'}$, we give the distinguisher direct access to the systems $\langle \hat{\mathbf{I}}/\hat{\mathbf{F}}\rangle_Z$ and $\langle \hat{\mathbf{J}}/\hat{\mathbf{G}}\rangle_{Z'}$, allowing k' and k'' queries, respectively. Lemma 6 implies that in this setting, using individual optimal strategies is optimal. The probabilities of provoking the MBOs by individually optimal strategies are $\nu_{k'}(\hat{\mathbf{F}}) = \Delta_{k'}(\mathbf{F},\mathbf{I})$ and $\nu_{k''}(\hat{\mathbf{G}}) = \Delta_{k''}(\mathbf{G},\mathbf{J})$, respectively, hence the advantage in guessing $Z \oplus Z'$ is $\Delta_{k'}(\mathbf{F},\mathbf{I})\Delta_{k''}(\mathbf{G},\mathbf{J})$. Taking into account the factor 2 from above (due to Lemma 3) this completes the proof. □

We say that a construction $\mathbf{C}(\cdot,\cdot)$ is *feed-forward* if, within the evaluation of a single query to $\mathbf{C}(\mathbf{F},\mathbf{G})$, no input to $\mathbf{F}$ (or $\mathbf{G}$) depends on a previous output of $\mathbf{F}$ (or $\mathbf{G}$) of the same evaluation of $\mathbf{C}(\mathbf{F},\mathbf{G})$. We will only consider constructions $\mathbf{C}(\cdot,\cdot)$ that make a single call to the invoked systems per invocation of $\mathbf{C}(\cdot,\cdot)$, and such constructions are always feed-forward. The proof of the following result is omitted.

Corollary 1. *Consider the setting of Theorem 1. If $\mathbf{C}(\cdot,\cdot)$ is a feed-forward construction, then the inequality also holds for non-adaptive strategies:*

$$\Delta_k^{\mathsf{NA}}(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})) \;\leq\; 2\,\Delta_{k'}^{\mathsf{NA}}(\mathbf{F},\mathbf{I})\,\Delta_{k''}^{\mathsf{NA}}(\mathbf{G},\mathbf{J}).$$

4.4 Implications of the Product Theorem

Recall that $\mathbf{R}$ ($\mathbf{P}$) denotes a uniform random function (permutation).

Definition 14. For two $(\mathcal{X},\mathcal{Y})$-systems $\mathbf{F}$ and $\mathbf{G}$ and a quasi-group operation $\star$ on $\mathcal{Y}$, we define $\mathbf{F}\star\mathbf{G}$ as the system obtained by feeding each input to both systems and combining the outputs using $\star$.

Corollary 2. *For any random functions $\mathbf{F}$ and $\mathbf{G}$, any quasi-group operation $\star$, and for all k,*

$$\Delta_k(\mathbf{F}\star\mathbf{G},\mathbf{R}) \;\leq\; 2\,\Delta_k(\mathbf{F},\mathbf{R})\,\Delta_k(\mathbf{G},\mathbf{R})$$

and

$$\Delta_k^{\mathsf{NA}}(\mathbf{F}\star\mathbf{G},\mathbf{R}) \;\leq\; 2\,\Delta_k^{\mathsf{NA}}(\mathbf{F},\mathbf{R})\,\Delta_k^{\mathsf{NA}}(\mathbf{G},\mathbf{R}).$$

The same statements hold for general random systems $\mathbf{F}$ and $\mathbf{G}$ when $\mathbf{R}$ is replaced by (a beacon) $\mathbf{B}$.

Proof. Let $\mathbf{I}:=\mathbf{R}$ and $\mathbf{J}:=\mathbf{R}$ in Theorem 1 and $\mathbf{C}(\mathbf{F},\mathbf{G}):=\mathbf{F}\star\mathbf{G}$. Condition (6) is satisfied since $\mathbf{F}\star\mathbf{R}\equiv\mathbf{R}$, $\mathbf{R}\star\mathbf{G}\equiv\mathbf{R}$, and $\mathbf{R}\star\mathbf{R}\equiv\mathbf{R}$. This proves the first inequality. The second inequality follows from Corollary 1 since $\mathbf{F}\star\mathbf{G}$ is clearly a feed-forward construction. The proof of the last statement is analogous. □

Definition 15. For two $(\mathcal{X},\mathcal{X})$-random permutations $\mathbf{F}$ and $\mathbf{G}$ we define $\mathbf{F}\triangleright\mathbf{G}$ as the system obtained by cascading $\mathbf{F}$ and $\mathbf{G}$, i.e., the input to $\mathbf{F}\triangleright\mathbf{G}$ is fed to $\mathbf{F}$, its output is fed to $\mathbf{G}$, and $\mathbf{G}$'s output is the output of $\mathbf{F}\triangleright\mathbf{G}$. Moreover, for a random permutation $\mathbf{F}$, we denote by $\langle\mathbf{F}\rangle$ the random permutation which can be queried from "both sides", i.e., one can also provide an output and receive the corresponding input.[18]

Corollary 3. *For any compatible random permutations $\mathbf{F}$ and $\mathbf{G}$, where $\mathbf{G}$ is stateless, for all k,*

$$\Delta_k(\mathbf{F}\triangleright\mathbf{G},\mathbf{P}) \;\leq\; 2\,\Delta_k(\mathbf{F},\mathbf{P})\,\Delta_k(\mathbf{G},\mathbf{P})$$

and

$$\Delta_k^{\mathsf{NA}}(\mathbf{F}\triangleright\mathbf{G},\mathbf{P}) \;\leq\; 2\,\Delta_k^{\mathsf{NA}}(\mathbf{F},\mathbf{P})\,\Delta_k^{\mathsf{NA}}(\mathbf{G},\mathbf{P}).$$

If also $\mathbf{F}$ is stateless, then the corresponding two inequalities also hold when bi-directional permutations are considered.[19]

[18] This definition is motivated by considering chosen-plaintext and chosen-ciphertext attacks against a block-cipher. One-sided and two-sided attacks are sometimes also called CCA and nCCA, for the adaptive and the non-adaptive version.

[19] E.g., $\Delta_k(\langle\mathbf{F}\rangle\triangleright\langle\mathbf{G}\rangle,\langle\mathbf{P}\rangle) \;\leq\; 2\,\Delta_k(\langle\mathbf{F}\rangle,\langle\mathbf{P}\rangle)\,\Delta_k(\langle\mathbf{G}\rangle,\langle\mathbf{P}\rangle)$.

Proof. Let $\mathbf{I} := \mathbf{P}$ and $\mathbf{J} := \mathbf{P}$ in Theorem 1 and $\mathbf{C}(\mathbf{F}, \mathbf{G}) := \mathbf{F} \triangleright \mathbf{G}$. Condition (6) is satisfied since $\mathbf{F} \triangleright \mathbf{P} \equiv \mathbf{P}$, $\mathbf{P} \triangleright \mathbf{G} \equiv \mathbf{P}$, and $\mathbf{P} \triangleright \mathbf{P} \equiv \mathbf{P}$. Note that $\mathbf{P} \triangleright \mathbf{G} \equiv \mathbf{P}$ is only guaranteed to hold if $\mathbf{G}$ is stateless.[20] No restriction applies to $\mathbf{F}$. This proves the first inequality. The second inequality follows from Corollary 1 since the cascade construction is feed-forward. The proof of the last statement is similar but omitted. □

5 Amplification of the Distinguisher Class

The second main result of this paper states that if subsystems of a neutralizing construction are only indistinguishable from ideal systems by a *weak* distinguisher class, then the construction is indistinguishable for a *stronger* distinguisher class. Recall Definition 13.

Theorem 2. *If $\mathbf{C}(\cdot, \cdot)$ is neutralizing for the pairs $(\mathbf{F}, \mathbf{I})$ and $(\mathbf{G}, \mathbf{J})$ of systems, then, for all k and all distinguishers $\mathbf{D}$,*[21]

$$\delta_k^{\mathbf{D}}(\mathbf{C}(\mathbf{F}, \mathbf{G}), \mathbf{C}(\mathbf{I}, \mathbf{J})) \;\leq\; \delta_{k'}^{\mathbf{DC}(\cdot, \mathbf{J})}(\mathbf{F}, \mathbf{I}) + \delta_{k''}^{\mathbf{DC}(\mathbf{I}, \cdot)}(\mathbf{G}, \mathbf{J}).$$

Proof. As in the proof of Theorem 1, let $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$ be defined as guaranteed by Lemma 5, where $\hat{\mathbf{F}}^{-} \equiv \mathbf{F}$, $\hat{\mathbf{I}}^{-} \equiv \mathbf{I}$, $\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{I}}^{\dashv}$, and $\delta_k^{\mathbf{D}}(\mathbf{F}, \mathbf{I}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{F}}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{I}})$ for all $\mathbf{D}$. (Note that this $\mathbf{D}$ is different from that in the theorem.) Similarly, let $\hat{\mathbf{G}}$ and $\hat{\mathbf{J}}$ be defined such that $\hat{\mathbf{G}}^{-} \equiv \mathbf{G}$, $\hat{\mathbf{J}}^{-} \equiv \mathbf{J}$, $\hat{\mathbf{G}}^{\dashv} \equiv \hat{\mathbf{J}}^{\dashv}$, and $\delta_k^{\mathbf{D}}(\mathbf{G}, \mathbf{J}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{G}}) = \nu_k^{\mathbf{D}}(\hat{\mathbf{J}})$ for all $\mathbf{D}$.

We can consider the following two systems with MBO: $\hat{\mathbf{H}}_{00} := \mathbf{C}(\hat{\mathbf{I}}, \hat{\mathbf{J}})$ and $\hat{\mathbf{H}}_{11} := \mathbf{C}(\hat{\mathbf{F}}, \hat{\mathbf{G}})$, where for each system the MBO is defined as the OR of the two internal MBOs. We have $\hat{\mathbf{H}}_{00}^{\dashv} \equiv \hat{\mathbf{H}}_{11}^{\dashv}$ because $\hat{\mathbf{F}}^{\dashv} \equiv \hat{\mathbf{I}}^{\dashv}$ and $\hat{\mathbf{G}}^{\dashv} \equiv \hat{\mathbf{J}}^{\dashv}$. Therefore, since $\hat{\mathbf{H}}_{00}^{-} \equiv \mathbf{C}(\mathbf{I}, \mathbf{J})$ and $\hat{\mathbf{H}}_{11}^{-} \equiv \mathbf{C}(\mathbf{F}, \mathbf{G})$, Lemma 4 implies that

$$\delta_k^{\mathbf{D}}(\mathbf{C}(\mathbf{F}, \mathbf{G}), \mathbf{C}(\mathbf{I}, \mathbf{J})) \;\leq\; \nu_k^{\mathbf{D}}(\hat{\mathbf{H}}_{00}).$$

It remains to determine a bound on $\nu_k^{\mathbf{D}}(\hat{\mathbf{H}}_{00})$. The MBO in $\hat{\mathbf{H}}_{00}$ (i.e., in $\mathbf{C}(\hat{\mathbf{I}}, \hat{\mathbf{J}})$) is provoked if either of the two internal MBOs is provoked. We can apply the union bound and consider the provocation of each MBO separately. More precisely, we consider the following systems with MBO: $\mathbf{C}(\hat{\mathbf{I}}, \mathbf{J})$ and $\mathbf{C}(\mathbf{I}, \hat{\mathbf{J}})$. Then $\nu_k^{\mathbf{D}}(\hat{\mathbf{H}}_{00})$ is bounded by the sum of the probabilities that $\mathbf{D}$ provokes the MBO in each of these systems, i.e.,

$$\nu_k^{\mathbf{D}}(\hat{\mathbf{H}}_{00}) \;\leq\; \nu_k^{\mathbf{D}}(\mathbf{C}(\hat{\mathbf{I}}, \mathbf{J})) + \nu_k^{\mathbf{D}}(\mathbf{C}(\mathbf{I}, \hat{\mathbf{J}})).$$

The proof is completed, using Lemma 5, by noting that $\nu_k^{\mathbf{D}}(\mathbf{C}(\hat{\mathbf{I}}, \mathbf{J})) = \nu_{k'}^{\mathbf{DC}(\cdot, \mathbf{J})}(\hat{\mathbf{I}}) = \delta_{k'}^{\mathbf{DC}(\cdot, \mathbf{J})}(\mathbf{F}, \mathbf{I})$ and $\nu_k^{\mathbf{D}}(\mathbf{C}(\mathbf{I}, \hat{\mathbf{J}})) = \nu_{k''}^{\mathbf{DC}(\mathbf{I}, \cdot)}(\hat{\mathbf{J}}) = \delta_{k''}^{\mathbf{DC}(\mathbf{I}, \cdot)}(\mathbf{G}, \mathbf{J})$. □

[20] As an example, consider a stateful random permutation $\mathbf{G}$ which internally builds a permutation function table by always taking the least unused element.

[21] Here, for example, $\mathbf{DC}(\cdot, \mathbf{J})$ denotes the distinguisher consisting of $\mathbf{D}$ connected to $\mathbf{C}(\cdot, \cdot)$ where the second subsystem is simulated as $\mathbf{J}$ and the system to be distinguished is placed as the first subsystem.

Note that since Theorem 2 applies to every distinguisher, it also applies to any distinguisher class $\mathcal{D}$, for instance the class of all distinguishers. Recalling that $\Delta_k(\mathbf{S},\mathbf{T}) = \delta_k(\mathbf{S},\mathbf{T})$ and $\Delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T}) = \delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T})$, we obtain:

Corollary 4. *For any compatible random functions* $\mathbf{F}$ *and* $\mathbf{G}$ *and any quasi-group operation* $\star$, *and all* k,

$$\Delta_k(\mathbf{F}\star\mathbf{G},\mathbf{R}) \;\le\; \Delta_k^{\mathsf{NA}}(\mathbf{F},\mathbf{R}) + \Delta_k^{\mathsf{NA}}(\mathbf{G},\mathbf{R}).$$

Proof. We recall that the $\star$-combination is neutralizing: $\mathbf{F}\star\mathbf{R} \equiv \mathbf{R}\star\mathbf{G} \equiv \mathbf{R}\star\mathbf{R} \equiv \mathbf{R}$. It remains to show that the distinguisher classes correspond to the class of non-adaptive distinguishers.

For any $\mathbf{D}$, the distinguisher $\mathbf{DC}(\cdot,\mathbf{J})$ (i.e., the distinguisher $\mathbf{D}(\cdot\star\mathbf{R})$) for provoking the MBO in $\hat{\mathbf{F}}$ obtains only random outputs, independently of $\hat{\mathbf{F}}$. A distinguisher could simulate these random outputs itself, ignoring the output of $\mathbf{F}\star\mathbf{R}$, and hence corresponds to a non-adaptive distinguisher. The same argument also applies to the distinguisher $\mathbf{DC}(\mathbf{I},\cdot)$ for provoking the MBO in $\hat{\mathbf{G}}$. □

Corollary 5. *For any compatible random permutations* $\mathbf{F}$ *and* $\mathbf{G}$, *where* $\mathbf{G}$ *is stateless, for all* k,

$$\Delta_k(\mathbf{F}\triangleright\mathbf{G},\mathbf{P}) \;\le\; \Delta_k^{\mathsf{NA}}(\mathbf{F},\mathbf{P}) + \Delta_k^{\mathsf{RI}}(\mathbf{G},\mathbf{P}).$$

If also $\mathbf{F}$ *is stateless, then*[22]

$$\Delta_k(\langle\mathbf{F}\triangleright\mathbf{G}^{-1}\rangle,\langle\mathbf{P}\rangle) \;\le\; \Delta_k^{\mathsf{NA}}(\mathbf{F},\mathbf{P}) + \Delta_k^{\mathsf{NA}}(\mathbf{G},\mathbf{P}).$$

The last statement means that $\langle\mathbf{F}\triangleright\mathbf{G}^{-1}\rangle$ is adaptively indistinguishable (from both sides) if $\mathbf{F}$ and $\mathbf{G}$ are only non-adaptively indistinguishable (from one side).

Proof. We recall that the $\triangleright$-combination is neutralizing: $\mathbf{F}\triangleright\mathbf{P} \equiv \mathbf{P}\triangleright\mathbf{G} \equiv \mathbf{P}\triangleright\mathbf{P} \equiv \mathbf{P}$. It remains to show that the distinguisher classes correspond to the class NA of non-adaptive distinguishers and the class RI of random-input distinguishers, respectively.

For any $\mathbf{D}$, the distinguisher $\mathbf{DC}(\cdot,\mathbf{J})$, i.e., the distinguisher $\mathbf{D}(\cdot\triangleright\mathbf{P})$, obtains only random outputs, independently of $\mathbf{F}$. A distinguisher could simulate these random outputs itself, ignoring the output of $\mathbf{F}\triangleright\mathbf{P}$, and hence corresponds to a non-adaptive distinguisher.

Similarly, the distinguisher $\mathbf{DC}(\mathbf{I},\cdot)$, i.e., the distinguisher $\mathbf{D}(\mathbf{P}\triangleright\cdot)$, can only produce random inputs to $\mathbf{G}$, with the possibility of repeating a previous input. Because $\mathbf{G}$ is stateless, repeating an input does not help in provoking the MBO in $\mathbf{G}$.

The proof of the second statement is omitted. □

[22] Here $\mathbf{G}^{-1}$ is the inverse of $\mathbf{G}$, which is well defined as $\mathbf{G}$ is a stateless random permutation.

Acknowledgments

It is a pleasure to thank Yevgeniy Dodis, Ghislain Fourny, Thomas Holenstein, Dominik Raub, Johan Sjödin, and Stefano Tessaro for discussions about random systems.

References

[KNR05] Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. In *APPROX-RANDOM*, volume 3624 of *Lecture Notes in Computer Science*, pages 354–365. Springer, 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*, volume 2332 of *Lecture Notes in Computer Science*, pages 110–132. Springer, 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 of *Lecture Notes in Computer Science*, pages 391–408. Springer, 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 of *Lecture Notes in Computer Science*, pages 410–427. Springer, 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 of *Lecture Notes in Computer Science*, pages 189–206. Springer, 2004.

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

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

[PS07] Krzysztof Pietrzak and Johan Sjödin. Domain extension for weak PRFs; the good, the bad, and the ugly. In *Advances in Cryptology — EUROCRYPT '07*, volume 2442 of *Lecture Notes in Computer Science*, pages 517–533. Springer, 2007.

[Vau98] Serge Vaudenay. Provable security for block ciphers by decorrelation. In *STACS*, volume 1373 of *Lecture Notes in Computer Science*, pages 249–275. Springer, 1998.

[Vau99] Serge Vaudenay. Adaptive-attack norm for decorrelation and super-pseudorandomness. In *Selected Areas in Cryptography*, volume 1758 of *Lecture Notes in Computer Science*, pages 49–61. Springer, 1999.

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