Free-Start Distinguishing: Combining Two Types of Indistinguishability Amplification

Peter Gaži[1,2] and Ueli Maurer[1]

[1] ETH Zürich, Switzerland
Department of Computer Science
{gazipete,maurer}@inf.ethz.ch
[2] Comenius University, Bratislava, Slovakia
Department of Computer Science

Abstract. The term indistinguishability amplification refers to a setting where a certain construction combines two (or more) cryptographic primitives of the same type to improve their indistinguishability from an ideal primitive. Various constructions achieving this property have been studied, both in the information-theoretic and computational setting. In the former, a result due to Maurer, Pietrzak and Renner describes the amplification achieved by a very general class of constructions called neutralizing. Two types of amplification are observed: a product theorem (bounding the advantage in distinguishing the construction by twice the product of individual advantages) and the amplification of the distinguisher class (the obtained construction is secure against a wider class of distinguishers).

In this paper, we combine these two aspects of information-theoretic indistinguishability amplification. We derive a new bound for the general case of a neutralizing construction that keeps the structure of a product theorem, while also capturing the amplification of the distinguisher class. This improves both bounds mentioned above.

The new technical notion we introduce, central to our analysis, is the notion of free-start distinguishing of systems. This describes the setting where the distinguisher is allowed to choose any common state for both systems and then it is supposed to distinguish these systems starting from that chosen state.

Keywords: Information-theoretic cryptography, indistinguishability amplification, neutralizing constructions, projected systems, free-start distinguishing.

1 Introduction

Indistinguishability Amplification. An important goal of cryptography is to provide real objects (e.g. functions, permutations) such that their behavior is indistinguishable from the corresponding ideal object (e.g. a truly random function or permutation) by a distinguisher interacting with these objects. One reasonable way to approach this task is to devise constructions that allow us

K. Kurosawa (Ed.): ICITS 2009, LNCS 5973, pp. 28–44, 2010.

to combine objects of the same type to obtain a new one, with provably better indistinguishability properties. This is called *indistinguishability amplification*.

A natural candidate for such an indistinguishability-amplifying construction for permutations is the composition, while for random functions it is the quasi-group combination of the outputs (e.g. XOR of the output bitstrings). Both these constructions are widely used in the design of practical cryptographic primitives, such as blockciphers. Therefore, the indistinguishability amplification achieved by these constructions deserves being studied in detail. Both these examples as well as other natural constructions are special cases of the general concept of a *neutralizing construction*, introduced in [6].

In the information-theoretic setting, the most general treatment of indistinguishability amplification is due to Maurer, Pietrzak and Renner [6]. In their work, two different types of indistinguishability amplification are presented. Both are proved for the general class of neutralizing constructions, but for simplicity we describe their contribution on the special case of the XOR of random functions $\mathbf{F} \oplus \mathbf{G}$. First, a product theorem is proved, stating that the advantage in distinguishing $\mathbf{F} \oplus \mathbf{G}$ from the uniform random function $\mathbf{R}$ is upper-bounded by twice the *product* of the individual distinguishing advantages for these functions. Second, an amplification of the distinguishing class is observed, proving that the advantage in distinguishing $\mathbf{F} \oplus \mathbf{G}$ from $\mathbf{R}$ adaptively is upper-bounded by the *sum* of advantages in distinguishing $\mathbf{F}$ and $\mathbf{G}$ from $\mathbf{R}$ non-adaptively.

Our Contribution. First, we extend the random system framework from [3], in which we perform our analysis. We introduce the concept of a system *projected to a specific state*. Loosely speaking, any properly defined discrete system $\mathbf{S}$ and a transcript t of interaction with this system together define a new system, which behaves as the original system $\mathbf{S}$ would behave after this interaction t. We refer to this new system as $\mathbf{S}$ projected to the state described by t. In particular, any one-player game can be modelled as a special type of a discrete system. Therefore, we are also able to model the intuitive situation where a player can continue playing a given game from a specific position (where the game is not won yet) or where it can pick an arbitrary such position in the game tree and try to win the game from there.

This leads to the central new notion in this paper, *free-start distinguishing*. Informally, the free-start distinguishing advantage of two systems is the best advantage a distinguisher can achieve, assuming that it is allowed to project both the distinguished systems to any one state consistent with both of them and then try to distinguish the resulting systems.

This concept, besides giving an interesting new viewpoint on the distinguishing of random systems, allows us to perform a more careful analysis of the indistinguishability amplification achieved by neutralizing constructions in the information-theoretic setting. We use the notion of free-start distinguishing to combine the two types of amplification described in [6]. We derive a new bound which keeps the structure of a product theorem, while involving also the non-adaptive distinguishing advantages, thus describing the amplification of the distinguisher class.

Motivation and Intuition. As observed in [6], there is a tight correspondence between distinguishing systems and winning an appropriately defined game. Distinguishing $\mathbf{F} \oplus \mathbf{G}$ from $\mathbf{R}$ can be reduced (by a factor of 2) to winning two games constructed from $\mathbf{F}$ and $\mathbf{G}$, while obtaining only the XOR of their outputs. As long as none of the games is won, the output of the construction is useless to the player, hence one of the games has to be won non-adaptively first. After achieving this, the player still has to win the other game, this time with access to some (possibly useful) outputs. Since winning each of these games is as hard as distinguishing the corresponding system from $\mathbf{R}$, one could conjecture a bound like

$$\Delta_k(\mathbf{F} \oplus \mathbf{G}, \mathbf{R}) \leq 2 \left(\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{R}) \cdot \Delta_k(\mathbf{G}, \mathbf{R}) + \Delta_k^{\mathsf{NA}}(\mathbf{G}, \mathbf{R}) \cdot \Delta_k(\mathbf{F}, \mathbf{R}) \right),$$

where $\Delta_k(\mathbf{S}, \mathbf{T})$ and $\Delta_k^{\mathsf{NA}}(\mathbf{S}, \mathbf{T})$ denote the adaptive and non-adaptive advantage in distinguishing $\mathbf{S}$ from $\mathbf{T}$ with k queries, respectively.

However, this is not correct, since winning the first game may involve getting the second game into a state where winning it becomes much easier than if played from scratch. We model this by allowing the player to choose the starting position in the second game freely, with the only restriction being that the game is not won yet in the chosen position. Translated back into the language of systems distinguishing, this gives us a valid bound

$$\Delta_k(\mathbf{F} \oplus \mathbf{G}, \mathbf{R}) \leq 2 \left(\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{R}) \cdot \Lambda_k(\mathbf{G}, \mathbf{R}) + \Delta_k^{\mathsf{NA}}(\mathbf{G}, \mathbf{R}) \cdot \Lambda_k(\mathbf{F}, \mathbf{R}) \right), \qquad (1)$$

where $\Lambda_k(\mathbf{S}, \mathbf{T})$ denotes the free-start distinguishing advantage for systems $\mathbf{S}$ and $\mathbf{T}$, as described above. In this paper we prove a general theorem for neutralizing constructions, of which the bound (1) is a simple corollary.

Related Work. There has been a lot of previous research on indistinguishability-amplifying constructions, both in the information-theoretic and the computational setting.

In the former, a product theorem for the composition of stateless permutations was proved by Vaudenay using the decorrelation framework [11]. The amplification of the distinguisher class was proved in [5] for a class of constructions and in [4] also for the four-round Feistel network. As mentioned above, the paper [6] addressed both these types of indistinguishability amplification for any neutralizing construction.

On the other hand, computational product theorems for various constructions were proved by Luby and Rackoff [2], Myers [8,9] and Dodis et al. [1]. For the general case of a neutralizing construction a product theorem was proved by Maurer and Tessaro [7]. The second type of amplification considered here, amplification of the distinguisher class, does not in general translate to the computational setting, as observed by Pietrzak [10].

2 Preliminaries

2.1 Basic Notation

Throughout the paper, we denote sets by calligraphic letters (e.g. $\mathcal{S}$). A k-tuple is denoted by $u^k = (u_1, \ldots, u_k)$, and the set of all k-tuples of elements of $\mathcal{U}$ is denoted by $\mathcal{U}^k$. The tuples can be concatenated, which we write as $u^k v^l = (u_1, \ldots, u_k, v_1, \ldots, v_l)$. By $\mathrm{ms}(i)$ we denote the set of monotone binary sequences of length i where zeroes are preceding ones, i.e., $\mathrm{ms}(i) = \{0^i, 0^{i-1}1, \ldots, 1^i\}$.

We usually denote random variables and concrete values they can take on by capital and small letters, respectively. Naturally, for any binary random variable B, we denote the event that it takes on the value 1 also by B. The complement of an event A is denoted by $\overline{A}$. For events A and B and random variables U and V with ranges $\mathcal{U}$ and $\mathcal{V}$, respectively, we denote by $\mathsf{P}_{UA|VB}$ the corresponding conditional probability distribution, seen as a function $\mathcal{U} \times \mathcal{V} \to \langle 0, 1\rangle$. Here the value $\mathsf{P}_{UA|VB}(u, v)$ is well-defined for all $u \in \mathcal{U}$ and $v \in \mathcal{V}$ such that $\mathsf{P}_{VB}(v) > 0$ and undefined otherwise. Two probability distributions P_U and $\mathsf{P}_{U'}$ on the same set $\mathcal{U}$ are equal, denoted $\mathsf{P}_U = \mathsf{P}_{U'}$, if $\mathsf{P}_U(u) = \mathsf{P}_{U'}(u)$ for all $u \in \mathcal{U}$. Conditional probability distributions are equal if the equality holds for all arguments for which both of them are defined. To emphasize the random experiment $\mathcal{E}$ in consideration, we usually write it in the superscript, e.g. $\mathsf{P}^{\mathcal{E}}_{U|V}(u, v)$. By a lowercase p we denote (conditional) probability distributions that by themselves do not define a random experiment.

2.2 Random Systems

In this subsection, we present the basic notions of the random systems framework introduced in [3], following the notational changes in [6]. The input-output behavior of any discrete system can be described by a *random system* in the spirit of the following definition.

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

The behavior of the random system is specified by the sequence of conditional probabilities $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1})$ (for $i \geq 1$) of obtaining the output $y_i \in \mathcal{Y}$ on query $x_i \in \mathcal{X}$ given the previous $i-1$ queries $x^{i-1} = (x_1, \ldots, x_{i-1}) \in \mathcal{X}^{i-1}$ and their corresponding outputs $y^{i-1} = (y_1, \ldots, y_{i-1}) \in \mathcal{Y}^{i-1}$.

We shall use boldface letters (e.g. $\mathbf{S}$) to denote both a discrete system and a random system corresponding to it. This should cause no confusion. We emphasize that although the results of this paper are stated for random systems, they hold for arbitrary systems, since the only property of a system that is relevant here is its input-output behavior. It is reasonable to consider two discrete systems equivalent if their input-output behaviors are the same, even if their internal structure differs.

Definition 2. *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* $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}} = \mathsf{p}^{\mathbf{T}}_{Y_i|X^iY^{i-1}}$ *for all* $i \geq 1$.

A random system can also be defined by a sequence of conditional probability distributions $\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}$ for $i \geq 1$. This description is often convenient, but is not minimal: the distributions $\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}$ must satisfy a consistency condition for different i. The conversion between these two forms can be described by

$$\mathsf{p}^{\mathbf{S}}_{Y^i|X^i} = \prod_{j=1}^{i} \mathsf{p}^{\mathbf{S}}_{Y_j|X^jY^{j-1}} \qquad \text{and} \qquad \mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}} = \frac{\mathsf{p}^{\mathbf{S}}_{Y^i|X^i}}{\mathsf{p}^{\mathbf{S}}_{Y^{i-1}|X^{i-1}}}. \tag{2}$$

A *random function* is a special type of random system that answers consistently, i.e., it satisfies the condition $X_i = X_j \Rightarrow Y_i = Y_j$. For example, $\mathbf{R}$ denotes a *uniform random function*, which answers every new query with an element uniformly chosen from its (finite) range. A *random permutation* on $\mathcal{X}$ is a random function $\mathcal{X} \to \mathcal{X}$ mapping distinct inputs to distinct outputs: $X_i \neq X_j \Rightarrow Y_i \neq Y_j$. For example, $\mathbf{P}$ denotes a *uniform random permutation*, which for a domain and range $\mathcal{X}$ realizes a function chosen uniformly at random from all bijective functions $\mathcal{X} \to \mathcal{X}$. Following [7], we say that a random function is *convex-combination stateless* (*cc-stateless*) if it corresponds to a random variable taking on as values function tables $\mathcal{X} \to \mathcal{Y}$. For example, both $\mathbf{R}$ and $\mathbf{P}$ are cc-stateless.

We can define a *distinguisher* $\mathbf{D}$ for an $(\mathcal{X}, \mathcal{Y})$-system as a $(\mathcal{Y}, \mathcal{X})$-system which is one query ahead, i.e., it is defined by the conditional probability distributions $\mathsf{p}^{\mathbf{D}}_{X_i|X^{i-1}Y^{i-1}}$ for all $i \geq 1$. In particular, the first query of $\mathbf{D}$ is determined by $\mathsf{p}^{\mathbf{D}}_{X_1}$. After a certain number of queries (say k), the distinguisher outputs a bit W_k depending on the transcript X^kY^k. For a random system $\mathbf{S}$ and a distinguisher $\mathbf{D}$, let $\mathbf{DS}$ be the random experiment where $\mathbf{D}$ interacts with $\mathbf{S}$. The distribution of X^kY^k in this experiment can be expressed by

$$\begin{aligned} \mathsf{P}^{\mathbf{DS}}_{X^kY^k}(x^k, y^k) &= \prod_{i=1}^{k} \mathsf{p}^{\mathbf{D}}_{X_i|X^{i-1}Y^{i-1}}(x_i, x^{i-1}, y^{i-1}) \mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1}) \\ &= \mathsf{p}^{\mathbf{D}}_{X^k|Y^{k-1}}(x^k, y^{k-1}) \cdot \mathsf{p}^{\mathbf{S}}_{Y^k|X^k}(y^k, x^k), \end{aligned} \tag{3}$$

where the last equality follows from (2).

We consider two special classes of distinguishers. By NA we denote the class of all (computationally unbounded) non-adaptive distinguishers which select all queries $X_1, \ldots, X_k$ in advance, i.e., independent of the outputs $Y_1, \ldots, Y_k$. By RI we denote the class of all (computationally unbounded) distinguishers which cannot select queries but are given uniformly random values $X_1, \ldots, X_k$ and the corresponding outputs $Y_1, \ldots, Y_k$. These distinguisher classes correspond to the attacks nCPA (non-adaptive chosen-plaintext attack) and KPA (known-plaintext attack) from the literature, respectively.

For two $(\mathcal{X}, \mathcal{Y})$-systems $\mathbf{S}$ and $\mathbf{T}$, the *distinguishing advantage* of $\mathbf{D}$ in distinguishing systems $\mathbf{S}$ and $\mathbf{T}$ by k queries is defined as

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

We shall denote by $\Delta_k^{\mathcal{D}}(\mathbf{S},\mathbf{T})$ and $\Delta_k(\mathbf{S},\mathbf{T})$ the maximal advantage over the class $\mathcal{D}$ of distinguishers and over all distinguishers issuing at most k queries, respectively. On the other hand, we define

$$\delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) := ||\mathsf{P}^{\mathbf{DS}}_{X^kY^k} - \mathsf{P}^{\mathbf{DT}}_{X^kY^k}|| = \frac{1}{2}\sum_{x^ky^k} |\mathsf{P}^{\mathbf{DS}}_{X^kY^k}(x^k,y^k) - \mathsf{P}^{\mathbf{DT}}_{X^kY^k}(x^k,y^k)|$$

to be the *statistical distance of transcripts* when $\mathbf{D}$ interacts with $\mathbf{S}$ and $\mathbf{T}$, respectively. Again, $\delta_k^{\mathcal{D}}(\mathbf{S},\mathbf{T})$ and $\delta_k(\mathbf{S},\mathbf{T})$ denote the maximal value over the class $\mathcal{D}$ of distinguishers and over all distinguishers, respectively. The statistical distance of transcripts is closely related to the distinguishing advantage: in general we have $\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) \le \delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$, 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, we have $\Delta_k(\mathbf{S},\mathbf{T}) = \delta_k(\mathbf{S},\mathbf{T})$, $\Delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T}) = \delta_k^{\mathsf{NA}}(\mathbf{S},\mathbf{T})$ and $\Delta_k^{\mathsf{RI}}(\mathbf{S},\mathbf{T}) = \delta_k^{\mathsf{RI}}(\mathbf{S},\mathbf{T})$. Finally, using (3) to expand the definition of $\delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$, we obtain

$$\begin{aligned}\delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) &= \frac{1}{2}\sum_{x^ky^k} \mathsf{p}^{\mathbf{D}}_{X^k|Y^{k-1}}(x^k,y^{k-1}) \cdot \left|\mathsf{p}^{\mathbf{S}}_{Y^k|X^k}(y^k,x^k) - \mathsf{p}^{\mathbf{T}}_{Y^k|X^k}(y^k,x^k)\right| \\ &= \sum_{x^ky^k} \mathsf{p}^{\mathbf{D}}_{X^k|Y^{k-1}}(x^k,y^{k-1}) \cdot \left(\mathsf{p}^{\mathbf{S}}_{Y^k|X^k}(y^k,x^k) - \mathsf{p}^{\mathbf{T}}_{Y^k|X^k}(y^k,x^k)\right), \quad (4)\end{aligned}$$

where the last summation goes only over all x^ky^k such that $\mathsf{p}^{\mathbf{S}}_{Y^k|X^k}(y^k,x^k) > \mathsf{p}^{\mathbf{T}}_{Y^k|X^k}(y^k,x^k)$ holds.

For two $(\mathcal{X},\mathcal{Y})$-systems $\mathbf{S}$ and $\mathbf{T}$ and a uniform random bit B, $\langle \mathbf{S}/\mathbf{T}\rangle_B$ denotes the random system which is equal to $\mathbf{S}$ if $B=0$ and equal to $\mathbf{T}$ otherwise. If mentioning the random variable B explicitly is not necessary, we only write $\langle \mathbf{S}/\mathbf{T}\rangle$. The following simple lemma comes from [6].

Lemma 1. *For every distinguisher* $\mathbf{D}$*, we have:*

(i) $\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T}) = 2\left|\mathsf{P}^{\mathbf{D}\langle \mathbf{S}/\mathbf{T}\rangle_B}(W_k = B) - \frac{1}{2}\right|$,
(ii) $\Delta_k^{\mathbf{D}}(\mathbf{S},\langle \mathbf{S}/\mathbf{T}\rangle_B) = \frac{1}{2}\Delta_k^{\mathbf{D}}(\mathbf{S},\mathbf{T})$.

We denote by $\mathbf{C}(\cdot,\cdot)$ a *construction* that invokes two other systems as its subsystems. If we instantiate these subsystems by $\mathbf{S}_1$ and $\mathbf{S}_2$, we denote the resulting system by $\mathbf{C}(\mathbf{S}_1,\mathbf{S}_2)$. Upon each query to $\mathbf{C}(\cdot,\cdot)$, the construction may adaptively issue 0 or more queries to its subsystems. A construction is *neutralizing* for pairs of systems $(\mathbf{F},\mathbf{I})$ and $(\mathbf{G},\mathbf{J})$ if $\mathbf{C}(\mathbf{F},\mathbf{J}) \equiv \mathbf{C}(\mathbf{I},\mathbf{G}) \equiv \mathbf{C}(\mathbf{I},\mathbf{J})$. Moreover, let k' and k'' denote the maximal number of queries made to the first and second subsystem, respectively, during the first k queries issued to the construction (if defined). There are two important examples of neutralizing constructions that we shall consider in this paper:

Quasi-group combination. For $(\mathcal{X},\mathcal{Y})$-random systems $\mathbf{F}$ and $\mathbf{G}$ and for a quasi-group[1] operation $\star$ on $\mathcal{Y}$, the construction $\mathbf{F}\star\mathbf{G}$ feeds any query it

[1] A binary operation $\star$ on $\mathcal{X}$ is a quasi-group operation if for every $a,c\in\mathcal{X}$ (every $b,c\in\mathcal{X}$) there is a unique $b\in\mathcal{X}$ ($a\in\mathcal{X}$) such that $a\star b = c$.

receives to both subsystems and then combines their outputs using $\star$ to determine its own output. This is a neutralizing construction for random functions $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{I} \equiv \mathbf{J} \equiv \mathbf{R}$.

Composition. For a $(\mathcal{X}, \mathcal{Y})$-random system $\mathbf{F}$ and a $(\mathcal{Y}, \mathcal{Z})$-random system $\mathbf{G}$, $\mathbf{F} \triangleright \mathbf{G}$ denotes the serial composition of systems: every input to $\mathbf{F} \triangleright \mathbf{G}$ is fed to $\mathbf{F}$, its output is fed to $\mathbf{G}$ and the output of $\mathbf{G}$ is the output of $\mathbf{F} \triangleright \mathbf{G}$. This is a neutralizing construction for a permutation $\mathbf{F}$, a cc-stateless permutation $\mathbf{G}$ and $\mathbf{I} \equiv \mathbf{J} \equiv \mathbf{P}$.

2.3 Monotone Boolean Outputs and Games

Among random systems, we shall be in particular interested in systems having a monotone bit as a part of their output, in the sense of the following definition from [6].

Definition 3. *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 all $j > i$. For convenience, we define $A_0 = 0$. For a system $\mathbf{S}$ with MBO we define two derived systems:*

(i) $\mathbf{S}^-$ *is the $(\mathcal{X}, \mathcal{Y})$-system obtained 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 $\mathbf{S}$:*

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

The reason for studying this particular type of systems is that any one-player game can be seen as a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-system $\mathbf{S}$ with a monotone binary output. Here the player makes moves $X_1, X_2, \ldots$ and receives game outputs $Y_1, Y_2, \ldots$. Additionally, the game after each move also outputs a monotone bit indicating whether the game has already been won. The goal of the player[2] is to provoke the change of this bit, which is initially 0. Note that it is irrelevant whether the player can see this bit, so we can think of it interacting only with the system $\mathbf{S}^-$.

For a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-system $\mathbf{S}$ with an MBO called A_i and for a player $\mathbf{D}$, we denote by $\nu_k^{\mathbf{D}}(\mathbf{S})$ the probability that $\mathbf{D}$ wins the game $\mathbf{S}$ within k queries, i.e., $\nu_k^{\mathbf{D}}(\mathbf{S}) = \mathsf{P}_{A_k}^{\mathbf{DS}}(1)$. As usually, $\nu_k^{\mathcal{D}}(\mathbf{S})$ and $\nu_k(\mathbf{S})$ denote the maximal winning probability over the class $\mathcal{D}$ of players and over all players, respectively.

The relationship between distinguishing two systems and winning an appropriately defined game was studied in [3] and later in [6], where the following lemma was proved.

Lemma 2. *For any two $(\mathcal{X}, \mathcal{Y})$-systems $\mathbf{S}$ and $\mathbf{T}$ there exist $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$ such that*

[2] Note that a player is formally the same type of object as a distinguisher, hence we shall use both terms, depending on the context.

(i) $\hat{\mathbf{S}}^{-} \equiv \mathbf{S}$
(ii) $\hat{\mathbf{T}}^{-} \equiv \mathbf{T}$
(iii) $\hat{\mathbf{S}}^{\dashv} \equiv \hat{\mathbf{T}}^{\dashv}$
(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* $\mathbf{D}$.

Intuitively, Lemma 2 states that any two systems $\mathbf{S}$ and $\mathbf{T}$ can be extended by adding an MBO to each of them that "signals" whether the system has deviated from the common behavior of both $\mathbf{S}$ and $\mathbf{T}$. The systems are equivalent as long as the MBOs are 0 and the probability that a distinguisher $\mathbf{D}$ turns one of these MBOs to 1 is equal to the statistical distance of transcripts of the experiments $\mathbf{DS}$ and $\mathbf{DT}$.

Moreover, it was proved in [6] that if any $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$ satisfy for every $i \geq 1$ the conditions (for $\hat{\mathbf{T}}$, the conditions are analogous)

$$\begin{aligned} \mathsf{p}^{\hat{\mathbf{S}}}_{Y^i A_i | X^i}(y^i, 0, x^i) &= m^{\mathbf{S},\mathbf{T}}_{x^i, y^i} \\ \mathsf{p}^{\hat{\mathbf{S}}}_{Y^i A_i | X^i}(y^i, 1, x^i) &= \mathsf{p}^{\mathbf{S}}_{Y^i | X^i}(y^i, x^i) - m^{\mathbf{S},\mathbf{T}}_{x^i, y^i} \end{aligned} \tag{5}$$

where

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

then they also satisfy the properties stated in Lemma 2. In fact, Lemma 2 was proved in [6] by demonstrating that the systems $\hat{\mathbf{S}}$ and $\hat{\mathbf{T}}$ satisfying (5) can always be constructed.

3 Projected Systems

Any system $\mathbf{S}$ and a transcript of the initial part of a possible interaction with it together define a new system that simulates the behavior of $\mathbf{S}$ from the state at the end of this interaction onwards. This is formalized in the following definition.

Definition 4. *For an* $(\mathcal{X}, \mathcal{Y})$*-random system* $\mathbf{S}$ *and* $(\overline{x}^j, \overline{y}^j) \in \mathcal{X}^j \times \mathcal{Y}^j$*, let* $\mathbf{S}[\overline{x}^j, \overline{y}^j]$ *denote the system* $\mathbf{S}$ projected to the state $\overline{x}^j \overline{y}^j$*, i.e. the random system that behaves like* $\mathbf{S}$ *would behave after answering the first* j *queries* $\overline{x}^j$ *by* $\overline{y}^j$*. Formally,* $\mathbf{S}[\overline{x}^j, \overline{y}^j]$ *is defined by the distributions*

$$\mathsf{p}^{\mathbf{S}[\overline{x}^j, \overline{y}^j]}_{Y_i | X^i Y^{i-1}}(y_i, x^i, y^{i-1}) := \mathsf{p}^{\mathbf{S}}_{Y_{j+i} | X^{j+i} Y^{j+i-1}}(y_i, \overline{x}^j x^i, \overline{y}^j y^{i-1})$$

if $\mathsf{p}^{\mathbf{S}}_{Y^j | X^j}(\overline{y}^j, \overline{x}^j) > 0$ *and undefined otherwise.*

This is most intuitive if we consider a game (i.e., a special type of system with an MBO), where the transcript represents a position in this game. For a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-system $\mathbf{S}$ representing a game, the MBO bits are also a part of the output, therefore we have to specify them when describing its answers to the first j queries. To denote a position where the game is not won yet, we set these bits to 0, obtaining the system $\mathbf{S}[\overline{x}^j, \overline{y}^j 0^j]$.

Definition 5. *Let* $\mathbf{S}$ *be a* $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$*-system with the MBO* A_i *and let* $\mathbf{D}$ *be a compatible player. Let* $j \leq k$ *be non-negative integers. For any* $x^j \in \mathcal{X}^j$ *and* $y^j \in \mathcal{Y}^j$ *such that* $\mathsf{p}^{\mathbf{S}}_{Y^j A_j | X^j}(y^j, 0, x^j) > 0$*, we call* $\nu^{\mathbf{D}}_{k-j}(\mathbf{S}[x^j, y^j 0^j])$ *the probability of* $\mathbf{D}$ winning the game $\mathbf{S}$ from the position $x^j y^j$ *within the remaining* $k - j$ *queries. Moreover, we also define the probability of winning* $\mathbf{S}$ *within* k *queries* with a free start *to be*

$$\lambda_k(\mathbf{S}) := \max_{j, x^j, y^j} \nu_{k-j}(\mathbf{S}[x^j, y^j 0^j]),$$

where the maximization[3] *goes over all* $j \leq k, x^j, y^j$ *such that the projected system* $\mathbf{S}[x^j, y^j 0^j]$ *is defined.*

Intuitively, if a player starts playing the game $\mathbf{S}$ from the position $x^j y^j$ (assuming the game is not won yet), $\nu_{k-j}(\mathbf{S}[x^j, y^j 0^j])$ describes the probability that it wins the game within the remaining $k - j$ queries if he plays optimally from now on. On the other hand, if the player is allowed to choose *any* position in the game tree within the first k queries (where the game is not won yet) and play from that position, it can win with probability $\lambda_k(\mathbf{S})$. Obviously $\lambda_k(\mathbf{S}) \geq \nu_k(\mathbf{S})$.

Let us now consider a construction $\mathbf{C}(\mathbf{S}_1, \mathbf{S}_2)$. In this section, we assume that $\mathbf{S}_1$ and $\mathbf{S}_2$ are two $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-systems (games) with MBOs A_i and B_i, respectively. Moreover, we assume that $\mathbf{C}(\mathbf{S}_1, \mathbf{S}_2)$ is a $(\mathcal{X}, \mathcal{Y} \times \{0,1\})$-construction and it combines the last binary outputs of its subsystems using the AND operation to determine its own binary output C_i. Note that although the construction may determine the number and ordering of the queries to its subsystems adaptively, we can assume that the order of the queries to the subsystems is well-defined for every run of the experiment. This justifies the following definition.

Definition 6. *In the experiment* $\mathbf{DC}(\mathbf{S}_1, \mathbf{S}_2)$*, let* F^i_j *denote the event that the game* $\mathbf{S}_i$ *was won during the first* j *queries to* $\mathbf{C}(\mathbf{S}_1, \mathbf{S}_2)$ *and it was the first of the games* $\mathbf{S}_1$*,* $\mathbf{S}_2$ *that was won.*

Note that if both games are to be won, one of them always has to be won first. Afterwards, the adversary needs to also win the second game in order to provoke the MBO of the whole construction. This is captured by the following lemma.

Lemma 3. *Let* $\mathbf{S}$ *denote the system* $\mathbf{C}(\mathbf{S}_1, \mathbf{S}_2)$ *with MBO as described above. Then we have*

$$\nu^{\mathbf{D}}_k(\mathbf{S}) \leq \mathsf{P}^{\mathbf{DS}}(F^1_k) \cdot \lambda_{k''}(\mathbf{S}_2) + \mathsf{P}^{\mathbf{DS}}(F^2_k) \cdot \lambda_{k'}(\mathbf{S}_1).$$

Proof. Since the MBO of $\mathbf{S}$ is the AND of the MBOs of the subsystems, we have

$$\begin{aligned} \nu^{\mathbf{D}}_k(\mathbf{S}) &\leq \mathsf{P}^{\mathbf{DS}}(F^1_k \wedge B_{k''}) + \mathsf{P}^{\mathbf{DS}}(F^2_k \wedge A_{k'}) \\ &= \mathsf{P}^{\mathbf{DS}}(F^1_k) \cdot \mathsf{P}^{\mathbf{DS}}(B_{k''} | F^1_k) + \mathsf{P}^{\mathbf{DS}}(F^2_k) \cdot \mathsf{P}^{\mathbf{DS}}(A_{k'} | F^2_k). \end{aligned}$$

[3] Note that depending on the game $\mathbf{S}$, any $j \in \{0, \ldots, k-1\}$ may maximize the term $\nu^{\mathbf{D}}_{k-j}(\mathbf{S}[x^j, y^j 0^j])$.

It remains to upper-bound the terms $\mathsf{P}^{\mathbf{DS}}(B_{k''}|F_k^1)$ and $\mathsf{P}^{\mathbf{DS}}(A_{k'}|F_k^2)$. Let X_i and Y_i be the random variables corresponding to the i-th input and $\mathcal{Y}$-output of $\mathbf{S}$, respectively; and let M_i and N_i (U_i and V_i) be the random variables corresponding to the i-th input and $\mathcal{Y}$-output of $\mathbf{S}_1$ ($\mathbf{S}_2$), respectively. Let T denote the random variable corresponding to the initial part of the transcript of the experiment from its beginning until the MBO A is provoked or until the end of the experiment, whichever comes first. This transcript contains all the queries X_i to the construction, all the corresponding answers (Y_i, C_i), as well as all the query-answer pairs $(M_i, (N_i, A_i))$ and $(U_i, (V_i, B_i))$ of the subsystems, in the order as they appeared during the execution. Conditioning over all possible values of T, we have

$$\mathsf{P}^{\mathbf{DS}}(B_{k''}|F_k^1) = \sum_t \mathsf{P}^{\mathbf{DS}}_{T|F_k^1}(t) \cdot \mathsf{P}^{\mathbf{DS}}_{B_{k''}|TF_k^1}(t). \tag{6}$$

Let now t be fixed such that $\mathsf{P}^{\mathbf{DS}}_{T|F_k^1}(t) > 0$, we need to prove $\mathsf{P}^{\mathbf{DS}}_{B_{k''}|TF_k^1}(t) \leq \lambda_{k''}(\mathbf{S}_2)$. Let us consider a player $\mathbf{D}'$ defined as follows: it simulates the behavior of the player $\mathbf{DC}(\mathbf{S}_1, \cdot)$. However, as long as the MBO A is not provoked, all its choices are fixed to follow the transcript t. After these "cheated" choices, as soon as the MBO A is provoked (and t ends), it simulates $\mathbf{D}$, $\mathbf{C}$ and $\mathbf{S}_1$ faithfully. Let j denote the number of queries issued to $\mathbf{S}_2$ in t, let u^j and v^j denote these queries and the corresponding answers, respectively. For the described player $\mathbf{D}'$, we have

$$\begin{aligned}
\mathsf{P}^{\mathbf{DS}}_{B_{k''}|TF_k^1}(t) &= \mathsf{P}^{\mathbf{D}'\mathbf{S}_2}_{B_{k''}|U^jV^j\overline{B_j}}(u^j, v^j) \\
&\leq \max_{\mathbf{D}} \mathsf{P}^{\mathbf{DS}_2}_{B_{k''}|U^jV^j\overline{B_j}}(u^j, v^j) \\
&= \nu_{k''-j}(\mathbf{S}[u^j, v^j 0^j]) \\
&\leq \lambda_{k''}(\mathbf{S}_2),
\end{aligned}$$

and since $\sum_t \mathsf{P}^{\mathbf{DS}}_{T|F_k^1}(t) = 1$, from (6) we have $\mathsf{P}^{\mathbf{DS}}(B_{k''}|F_k^1) \leq \lambda_{k''}(\mathbf{S}_2)$. The same argument gives us a symmetric bound for $\mathsf{P}^{\mathbf{DS}}(A_{k'}|F_k^2)$ and concludes the proof. □

4 Free-Start Distinguishing

The notion of winning a game with a free start, captured by the quantity $\lambda_k(\mathbf{S})$, has a counterpart in the language of systems indistinguishability, which we now define formally.

Definition 7. *For any random systems* $\mathbf{S}$ *and* $\mathbf{T}$*, we define the* free-start distinguishing advantage of $\mathbf{S}$ and $\mathbf{T}$ *to be*

$$\Lambda_k(\mathbf{S}, \mathbf{T}) := \max_{j, x^j, y^j} \Delta_{k-j}(\mathbf{S}[x^j, y^j], \mathbf{T}[x^j, y^j]),$$

where the maximization goes over all $j \in \{0, \ldots, k-1\}$ *and all* x^j, y^j *such that the systems on the right side are defined.*

Informally, suppose that the distinguisher is allowed to choose an arbitrary transcript $x^j y^j$ compatible with both the systems it is supposed to distinguish, project them to the states described by this transcript and then try to distinguish the resulting systems with the remaining $k-j$ queries. Then the quantity $\Lambda_k(\mathbf{S}, \mathbf{T})$ denotes the optimal advantage it can achieve.

To demonstrate the relationship between λ_k and Λ_k, we exploit the connection between distinguishing two systems and winning an appropriately defined game described in [6]. Let us consider the setting with a real system $\mathbf{F}$ (e.g. a random function) and an ideal system $\mathbf{I}$ (e.g. a uniform random function). Using Lemma 2 (and, in particular, condition (5)), we can add MBOs to the systems $\mathbf{F}$ and $\mathbf{I}$ to obtain systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$ such that $\nu_k(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle) = \Delta_k(\mathbf{F}, \mathbf{I})$ and the systems behave identically as long as the MBO is not provoked. Since provoking this MBO corresponds to distinguishing the systems, one can expect $\nu_{k-j}(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle[x^j, y^j 0^j])$ to be related to the advantage in distinguishing $\mathbf{F}$ and $\mathbf{I}$ projected to the state described by the transcript $x^j y^j$ on the remaining $k-j$ queries. In the following, we capture this intuition.

Lemma 4. *Let* $\mathbf{F}$ *and* $\mathbf{I}$ *be two random systems, let* $\hat{\mathbf{F}}$, $\hat{\mathbf{I}}$ *be the systems obtained from* $\mathbf{F}$, $\mathbf{I}$ *by adding the MBOs according to Lemma 2 and condition (5). Then we have*

$$\nu_k(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle[\overline{x}^j, \overline{y}^j 0^j]) = \Delta_k(\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]^-)$$

for any $\overline{x}^j$,$\overline{y}^j$ *such that the system on the left side is defined.*

Proof. First note that $\nu_k(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle[\overline{x}^j, \overline{y}^j 0^j]) = \nu_k(\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j])$, hence it suffices to prove $\nu_k(\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]) = \Delta_k(\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]^-)$. We prove this claim by showing that the MBO of $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]$, originally defined to capture the differences between $\mathbf{F}$ and $\mathbf{I}$, keeps the properties guaranteed by Lemma 2 also with respect to the systems $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$ and $\hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]^-$. We achieve this by showing that the system $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]$ satisfies the condition (5) with respect to the systems $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$ and $\hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]^-$. Seeing this, the claim follows from Lemma 2.

Throughout the proof let p denote the probability $\mathsf{p}^{\hat{\mathbf{F}}}_{Y^j A^j | X^j}(\overline{y}^j, 0^j, \overline{x}^j) = \mathsf{p}^{\hat{\mathbf{I}}}_{Y^j A^j | X^j}(\overline{y}^j, 0^j, \overline{x}^j)$ (by the assumptions of the lemma, $p > 0$). We first show that the relevant probabilities describing the behavior of the random system $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]$ (and $\hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]$) correspond to the probabilities describing the original system $\hat{\mathbf{F}}$ (and $\hat{\mathbf{I}}$) scaled by the factor $1/p$. More precisely, we have

$$\begin{aligned}
\mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]}_{Y^i | X^i}(y^i, x^i) &= \sum_{a^i \in \mathrm{ms}(i)} \mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]}_{Y^i A^i | X^i}(y^i, a^i, x^i) \\
&= \frac{1}{p} \cdot \sum_{a^i \in \mathrm{ms}(i)} \mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i} A^{j+i} | X^{j+i}}(\overline{y}^j y^i, 0^j a^i, \overline{x}^j x^i) \\
&= \frac{1}{p} \cdot \mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i} A^j | X^{j+i}}(\overline{y}^j y^i, 0^j, \overline{x}^j x^i)
\end{aligned}$$

and similarly $\mathsf{p}^{\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]}_{Y^i|X^i}(y^i,x^i) = \frac{1}{p}\cdot \mathsf{p}^{\hat{\mathbf{I}}}_{Y^{j+i}A^j|X^{j+i}}(\overline{y}^j y^i, 0^j, \overline{x}^j x^i)$. We can use this to express the quantity $m^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-,\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{x^i,y^i}$ as

$$\begin{aligned} m^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-,\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{x^i,y^i} &= \min\left\{\mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{Y^i|X^i}(y^i,x^i), \mathsf{p}^{\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{Y^i|X^i}(y^i,x^i)\right\} \\ &= \frac{1}{p}\cdot \min\Big\{\mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i}A^j|X^{j+i}}(\overline{y}^j y^i, 0^j, \overline{x}^j x^i), \\ &\qquad\qquad \mathsf{p}^{\hat{\mathbf{I}}}_{Y^{j+i}A^j|X^{j+i}}(\overline{y}^j y^i, 0^j, \overline{x}^j x^i)\Big\} \\ &= \frac{1}{p}\cdot \mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i}A^{j+i}|X^{j+i}}(\overline{y}^j y^i, 0^{j+i}, \overline{x}^j x^i) \qquad (7) \\ &= \frac{1}{p}\cdot m^{\hat{\mathbf{F}},\hat{\mathbf{I}}}_{\overline{x}^j x^i,\overline{y}^j y^i}. \end{aligned}$$

To justify the step (7), note that from the condition (5), which is satisfied for $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$, we have $\mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i}A^{j+i}|X^{j+i}}(\overline{y}^j y^i, 0^{j+i}, \overline{x}^j x^i) = \mathsf{p}^{\hat{\mathbf{I}}}_{Y^{j+i}A^{j+i}|X^{j+i}}(\overline{y}^j y^i, 0^{j+i}, \overline{x}^j x^i)$ and also $\mathsf{p}_{Y^{j+i}A^j|X^{j+i}}(\overline{y}^j y^i, 0^j, \overline{x}^j x^i) = \mathsf{p}_{Y^{j+i}A^{j+i}|X^{j+i}}(\overline{y}^j y^i, 0^{j+i}, \overline{x}^j x^i)$ for at least one of the systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$.

Now we can verify that the condition (5) is satisfied also for the system $\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]$ with respect to the systems $\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-$ and $\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-$. For the first equation of (5), we have

$$\begin{aligned} \mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]}_{Y^i A_i|X^i}(y^i,0,x^i) &= \frac{1}{p}\cdot \mathsf{p}^{\hat{\mathbf{F}}}_{Y^{j+i}A_{j+i}|X^{j+i}}(\overline{y}^j y^i, 0, \overline{x}^j x^i) \\ &= \frac{1}{p}\cdot m^{\hat{\mathbf{F}},\hat{\mathbf{I}}}_{\overline{x}^j x^i,\overline{y}^j y^i} = m^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-,\hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{x^i,y^i} \end{aligned}$$

and since clearly $\mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]}_{Y^i|X^i}(y^i,x^i) = \mathsf{p}^{\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-}_{Y^i|X^i}(y^i,x^i)$, the second equation of (5) is satisfied as well. Therefore, by Lemma 2(iv), we have $\nu_k(\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]) = \Delta_k(\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-)$. □

Lemma 4 involves the systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$ projected to a specific state, but it is more desirable to consider the original systems $\mathbf{F}$ and $\mathbf{I}$ instead. This is achieved by the following lemma.

Lemma 5. *In the setting described in Lemma 4, we have*

$$\Delta_k(\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-) \le \Delta_k(\mathbf{F}[\overline{x}^j,\overline{y}^j], \mathbf{I}[\overline{x}^j,\overline{y}^j])$$

for any $\overline{x}^j$,$\overline{y}^j$ such that the systems on the left side are defined.

Proof. To prove the lemma, we show that for any distinguisher $\mathbf{D}$ we have $\delta^{\mathbf{D}}_k(\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-) \le \delta^{\mathbf{D}}_k(\mathbf{F}[\overline{x}^j,\overline{y}^j], \mathbf{I}[\overline{x}^j,\overline{y}^j])$. Without loss of generality, let us assume $\mathsf{p}^{\mathbf{F}}_{Y^j|X^j}(\overline{y}^j,\overline{x}^j) \ge \mathsf{p}^{\mathbf{I}}_{Y^j|X^j}(\overline{y}^j,\overline{x}^j)$, otherwise the proof would

be symmetric. This assumption implies $\hat{\mathbf{I}}[\overline{x}^j, \overline{y}^j 0^j]^- \equiv \mathbf{I}[\overline{x}^j, \overline{y}^j]$, hence it suffices to prove

$$\delta_k^{\mathbf{D}}(\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-, \mathbf{I}[\overline{x}^j, \overline{y}^j]) \leq \delta_k^{\mathbf{D}}(\mathbf{F}[\overline{x}^j, \overline{y}^j], \mathbf{I}[\overline{x}^j, \overline{y}^j]).$$

Using (4) to express both sides of this inequality, we see that we only need to prove that for all $x^k \in \mathcal{X}^k$ and $y^k \in \mathcal{Y}^k$,

$$\mathsf{p}_{Y^k|X^k}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-}(y^k, x^k) < \mathsf{p}_{Y^k|X^k}^{\mathbf{I}[\overline{x}^j, \overline{y}^j]}(y^k, x^k) \Rightarrow \mathsf{p}_{Y^k|X^k}^{\mathbf{F}[\overline{x}^j, \overline{y}^j]}(y^k, x^k) \leq \mathsf{p}_{Y^k|X^k}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-}(y^k, x^k). \tag{8}$$

In the systems $\mathbf{I}[\overline{x}^j, \overline{y}^j]$, $\mathbf{F}[\overline{x}^j, \overline{y}^j]$ and $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$, the conditional distributions $\mathsf{p}_{Y^k|X^k}(y^k, x^k)$ are given by the following expressions, respectively:

$$\mathsf{p}_{Y^k|X^k}^{\mathbf{I}[\overline{x}^j, \overline{y}^j]}(y^k, x^k) = \frac{\mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{I}}(\overline{y}^j y^k, \overline{x}^j x^k)}{\mathsf{p}_{Y^j|X^j}^{\mathbf{I}}(\overline{y}^j, \overline{x}^j)} \tag{9}$$

$$\mathsf{p}_{Y^k|X^k}^{\mathbf{F}[\overline{x}^j, \overline{y}^j]}(y^k, x^k) = \frac{\mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{F}}(\overline{y}^j y^k, \overline{x}^j x^k)}{\mathsf{p}_{Y^j|X^j}^{\mathbf{F}}(\overline{y}^j, \overline{x}^j)} \tag{10}$$

$$\mathsf{p}_{Y^k|X^k}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-}(y^k, x^k) = \frac{\mathsf{p}_{Y^{j+k}A^j|X^{j+k}}^{\hat{\mathbf{F}}}(\overline{y}^j y^k, 0^j, \overline{x}^j x^k)}{\mathsf{p}_{Y^j A^j|X^j}^{\hat{\mathbf{F}}}(\overline{y}^j, 0^j, \overline{x}^j)} \tag{11}$$

Informally, the conditional distributions $\mathsf{p}_{Y^k|X^k}$ of the systems $\mathbf{I}[\overline{x}^j, \overline{y}^j]$, $\mathbf{F}[\overline{x}^j, \overline{y}^j]$ and $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$ are again related to the conditional distributions $\mathsf{p}_{Y^{j+k}|X^{j+k}}$ of the original systems ($\mathbf{I}$, $\mathbf{F}$, and $\hat{\mathbf{F}}$ with $A_j = 0$, respectively) by some scaling factors (the denominators in the above equations). The factor turns out to be the same for $\mathbf{I}[\overline{x}^j, \overline{y}^j]$ and $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$, however for $\mathbf{F}[\overline{x}^j, \overline{y}^j]$ it may be different. This results into a different scaling of the distributions for $\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-$ and $\mathbf{F}[\overline{x}^j, \overline{y}^j]$ and allows us to show that (8) is indeed satisfied. A more detailed argument follows.

Let us fix x^k and y^k such that $\mathsf{p}_{Y^k|X^k}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-}(y^k, x^k) < \mathsf{p}_{Y^k|X^k}^{\mathbf{I}[\overline{x}^j, \overline{y}^j]}(y^k, x^k)$. By the definition of A_i we have $\mathsf{p}_{Y^j|X^j}^{\mathbf{I}}(\overline{y}^j, \overline{x}^j) = \mathsf{p}_{Y^j A^j|X^j}^{\hat{\mathbf{F}}}(\overline{y}^j, 0^j, \overline{x}^j)$, hence by comparing the equations (9) and (11) we get $\mathsf{p}_{Y^{j+k}A^j|X^{j+k}}^{\hat{\mathbf{F}}}(\overline{y}^j y^k, 0^j, \overline{x}^j x^k) < \mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{I}}(\overline{y}^j y^k, \overline{x}^j x^k)$. This in turn implies $\mathsf{p}_{Y^{j+k}A^{j+k}|X^{j+k}}^{\hat{\mathbf{F}}}(\overline{y}^j y^k, 0^{j+k}, \overline{x}^j x^k) < \mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{I}}(\overline{y}^j y^k, \overline{x}^j x^k)$. Now, recalling that the MBO A_i is defined to satisfy the properties (5), we see that $\mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{F}}(\overline{y}^j y^k, \overline{x}^j x^k) < \mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{I}}(\overline{y}^j y^k, \overline{x}^j x^k)$ and therefore also $\mathsf{p}_{Y^{j+k}A^{j+k}|X^{j+k}}^{\hat{\mathbf{F}}}(\overline{y}^j y^k, 0^{j+k}, \overline{x}^j x^k) = \mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{F}}(\overline{y}^j y^k, \overline{x}^j x^k)$. This in turn implies $\mathsf{p}_{Y^{j+k}A^j|X^{j+k}}^{\hat{\mathbf{F}}}(\overline{y}^j y^k, 0^j, \overline{x}^j x^k) = \mathsf{p}_{Y^{j+k}|X^{j+k}}^{\mathbf{F}}(\overline{y}^j y^k, \overline{x}^j x^k)$, hence the numerators in (10) and (11) are the same. The denominators are easy to compare, it obviously holds $\mathsf{p}_{Y^j|X^j}^{\mathbf{F}}(\overline{y}^j, \overline{x}^j) \geq \mathsf{p}_{Y^j A^j|X^j}^{\hat{\mathbf{F}}}(\overline{y}^j, 0^j, \overline{x}^j)$, hence from (10) and (11) we obtain $\mathsf{p}_{Y^k|X^k}^{\mathbf{F}[\overline{x}^j, \overline{y}^j]}(y^k, x^k) \leq \mathsf{p}_{Y^k|X^k}^{\hat{\mathbf{F}}[\overline{x}^j, \overline{y}^j 0^j]^-}(y^k, x^k)$, completing the proof of (8). □

Note that combining the technical Lemmas 4 and 5 gives us

$$\lambda_k(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle) = \max_{j,\overline{x}^j,\overline{y}^j} \Delta_{k-j}\left(\hat{\mathbf{F}}[\overline{x}^j,\overline{y}^j 0^j]^-, \hat{\mathbf{I}}[\overline{x}^j,\overline{y}^j 0^j]^-\right) \leq \Lambda_k(\mathbf{F},\mathbf{I}) \qquad (12)$$

for the systems described above.

5 Connection to Indistinguishability Amplification

We are now ready to prove our main theorem. First we define some intuitive notation: by $\mathcal{D}\mathbf{C}(\cdot,\mathbf{J})$ we denote the class of distinguishers obtained by connecting any distinguisher to $\mathbf{C}(\cdot,\mathbf{J})$ and placing the system to be distinguished as the first subsystem. The class of distinguishers $\mathcal{D}\mathbf{C}(\mathbf{I},\cdot)$ is defined analogously.

Theorem 1. *Let $\mathbf{C}(\cdot,\cdot)$ be a neutralizing construction for the pairs $(\mathbf{F},\mathbf{I})$ and $(\mathbf{G},\mathbf{J})$ of systems. Let $\mathbf{Q}$ denote the system $\mathbf{C}(\mathbf{I},\mathbf{J})$. Then, for all k,*

$$\Delta_k(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{Q}) \leq 2\left(\delta_{k'}^{\mathcal{D}\mathbf{C}(\cdot,\mathbf{J})}(\mathbf{F},\mathbf{I})\cdot\Lambda_{k''}(\mathbf{G},\mathbf{J}) + \delta_{k''}^{\mathcal{D}\mathbf{C}(\mathbf{I},\cdot)}(\mathbf{G},\mathbf{J})\cdot\Lambda_{k'}(\mathbf{F},\mathbf{I})\right).$$

Proof. We use the technique from the proof of Theorem 1 in [6] to transform the task of distinguishing $\mathbf{C}(\mathbf{F},\mathbf{G})$ from $\mathbf{Q}$ to the task of provoking the MBO of the system $\mathbf{S} := \hat{\mathbf{C}}(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle_{Z_1},\langle\hat{\mathbf{G}}/\hat{\mathbf{J}}\rangle_{Z_2})$, where $\hat{\mathbf{F}}$, $\hat{\mathbf{I}}$ and $\hat{\mathbf{G}}$, $\hat{\mathbf{J}}$ are obtained using Lemma 2 from $\mathbf{F}$, $\mathbf{I}$ and $\mathbf{G}$, $\mathbf{J}$, respectively; and $\hat{\mathbf{C}}$ is the same construction as $\mathbf{C}$ except that it also has an MBO, which is defined as the AND of the two internal MBOs. Then we use a different approach to bound the value $\nu_k(\mathbf{S})$, exploiting the concept of free-start distinguishing.

First, by Lemma 1 (ii) we have $\Delta_k(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{Q}) = 2\cdot\Delta_k(\langle\mathbf{C}(\mathbf{F},\mathbf{G})/\mathbf{Q}\rangle_Z,\mathbf{Q})$ and by Lemma 1 (i) $\Delta_k(\langle\mathbf{C}(\mathbf{F},\mathbf{G})/\mathbf{Q}\rangle_Z,\mathbf{Q})$ is the optimal advantage in guessing the uniform random bit Z' in the system $\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G})/\mathbf{Q}\rangle_Z/\mathbf{Q}\rangle_{Z'}$. However, thanks to the neutralizing property of $\mathbf{C}(\cdot,\cdot)$. it can be easily verified that $\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G})/\mathbf{Q}\rangle_Z/\mathbf{Q}\rangle_{Z'} \equiv \mathbf{C}(\langle\mathbf{F}/\mathbf{I}\rangle_{Z_1},\langle\mathbf{G}/\mathbf{J}\rangle_{Z_2})$ for independent uniformly random bits $Z_1 := Z$ and $Z_2 := Z\oplus Z'$. Hence, $\Delta_k(\langle\mathbf{C}(\mathbf{F},\mathbf{G})/\mathbf{Q}\rangle_Z,\mathbf{Q})$ is also the optimal advantage in guessing the bit $Z' = Z_1\oplus Z_2$ in $\mathbf{C}(\langle\mathbf{F}/\mathbf{I}\rangle_{Z_1},\langle\mathbf{G}/\mathbf{J}\rangle_{Z_2})$.

We can now extend the systems $\mathbf{F}$ and $\mathbf{I}$ by adding MBOs satisfying the equations (5) to obtain the systems $\hat{\mathbf{F}}$ and $\hat{\mathbf{I}}$ with the properties guaranteed by Lemma 2. Similarly, we can extend $\mathbf{G}$ and $\mathbf{J}$ and obtain the systems $\hat{\mathbf{G}}$ and $\hat{\mathbf{J}}$. Since the MBO in $\mathbf{S}$ can always be ignored, the task of guessing $Z_1\oplus Z_2$ can only be easier in $\mathbf{S}$ compared to $\mathbf{C}(\langle\mathbf{F}/\mathbf{I}\rangle_{Z_1},\langle\mathbf{G}/\mathbf{J}\rangle_{Z_2})$. However, as long as one of the MBOs in the subsystems of $\mathbf{S}$ is 0, the advantage in guessing the corresponding bit Z_i is 0 and hence also the advantage in guessing $Z_1\oplus Z_2$ is 0. Therefore the latter advantage can be upper-bounded by $\nu_k(\mathbf{S})$.

Using Lemma 3, for any distinguisher $\mathbf{D}$ we have

$$\nu_k^{\mathbf{D}}(\mathbf{S}) \leq \mathsf{P}^{\mathbf{DS}}(F_k^1)\cdot\lambda_{k''}(\langle\hat{\mathbf{G}}/\hat{\mathbf{J}}\rangle) + \mathsf{P}^{\mathbf{DS}}(F_k^2)\cdot\lambda_{k'}(\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle).$$

Let us first bound the term $\mathsf{P}^{\mathbf{DS}}(F_k^1)$. Since $\langle\hat{\mathbf{F}}/\hat{\mathbf{I}}\rangle^{\dashv} \equiv \hat{\mathbf{F}}^{\dashv}$ and $\langle\hat{\mathbf{G}}/\hat{\mathbf{J}}\rangle^{\dashv} \equiv \hat{\mathbf{J}}^{\dashv}$, we have $\mathsf{P}^{\mathbf{DS}}(F_k^1) = \mathsf{P}^{\mathbf{D}\hat{\mathbf{C}}(\hat{\mathbf{F}},\hat{\mathbf{J}})}(F_k^1)$. Moreover, $\mathsf{P}^{\mathbf{D}\hat{\mathbf{C}}(\hat{\mathbf{F}},\hat{\mathbf{J}})}(F_k^1) \leq \nu_k^{\mathbf{D}}(\mathbf{C}(\hat{\mathbf{F}},\mathbf{J}))$

since on the left side, we only consider the MBO of $\hat{\mathbf{F}}$ being provoked first, while on the right side is the probability of it being provoked at any time. Obviously $\nu_k^{\mathbf{D}}(\mathbf{C}(\hat{\mathbf{F}}, \mathbf{J})) \leq \nu_{k'}^{\mathbf{DC}(\cdot,\mathbf{J})}(\hat{\mathbf{F}})$ and by Lemma 2 we have $\nu_{k'}^{\mathbf{DC}(\cdot,\mathbf{J})}(\hat{\mathbf{F}}) = \delta_{k'}^{\mathbf{DC}(\cdot,\mathbf{J})}(\mathbf{F}, \mathbf{I})$. By a symmetric reasoning we obtain $\mathsf{P}^{\mathbf{DS}}(F_k^2) \leq \delta_{k''}^{\mathbf{DC}(\mathbf{I},\cdot)}(\mathbf{G}, \mathbf{J})$.

Finally, using (12) we obtain the bounds $\lambda_{k''}(\langle \hat{\mathbf{G}}/\hat{\mathbf{J}} \rangle) \leq \Lambda_{k''}(\mathbf{G}, \mathbf{J})$ and $\lambda_{k'}(\langle \hat{\mathbf{F}}/\hat{\mathbf{I}} \rangle) \leq \Lambda_{k'}(\mathbf{F}, \mathbf{I})$, which together conclude the proof. □

For the two particular neutralizing constructions that motivate our analysis, we obtain the following corollaries.

Corollary 1. *Let* $\mathbf{F}$ *and* $\mathbf{G}$ *be* $(\mathcal{X}, \mathcal{Y})$*-random functions, let* $\star$ *be a quasi-group operation on* $\mathcal{Y}$*. Then, for all* k*,*

$$\Delta_k(\mathbf{F} \star \mathbf{G}, \mathbf{R}) \leq 2\left(\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{R}) \cdot \Lambda_k(\mathbf{G}, \mathbf{R}) + \Delta_k^{\mathsf{NA}}(\mathbf{G}, \mathbf{R}) \cdot \Lambda_k(\mathbf{F}, \mathbf{R})\right).$$

Proof. Applying Theorem 1 to the neutralizing construction $\mathbf{F} \star \mathbf{G}$, it only remains to prove that $\mathcal{D}(\cdot \star \mathbf{R})$ corresponds to the class of non-adaptive distinguishers. This is indeed the case, since any distinguisher will only receive random outputs from $\mathbf{F} \star \mathbf{R}$. It could simulate these outputs itself, ignoring the actual outputs, thus operating non-adaptively. The same holds for the class of distinguishers $\mathcal{D}(\mathbf{R} \star \cdot)$. Recalling that $\delta_k^{\mathsf{NA}}(\mathbf{S}, \mathbf{T}) = \Delta_k^{\mathsf{NA}}(\mathbf{S}, \mathbf{T})$ for any systems $\mathbf{S}$, $\mathbf{T}$ completes the proof. □

Corollary 2. *Let* $\mathbf{F}$ *and* $\mathbf{G}$ *be* $(\mathcal{X}, \mathcal{X})$*-random permutations, let* $\mathbf{G}$ *be cc-stateless. Then, for all* k*,*

$$\Delta_k(\mathbf{F} \triangleright \mathbf{G}, \mathbf{P}) \leq 2\left(\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{P}) \cdot \Lambda_k(\mathbf{G}, \mathbf{P}) + \Delta_k^{\mathsf{RI}}(\mathbf{G}, \mathbf{P}) \cdot \Lambda_k(\mathbf{F}, \mathbf{P})\right).$$

Proof. Again, when applying Theorem 1 to the neutralizing construction $\mathbf{F} \triangleright \mathbf{G}$, we need to justify that the distinguisher classes $\mathcal{D}(\cdot \triangleright \mathbf{P})$ and $\mathcal{D}(\mathbf{P} \triangleright \cdot)$ correspond to NA and RI, respectively. In the first case, the distinguisher only receives random outputs, so it can again simulate them itself and hence corresponds to a non-adaptive distinguisher. In the second case, the distinguisher $\mathbf{D}(\mathbf{P} \triangleright \cdot)$ can only provide random inputs to the distinguished system, with the possibility of repeating an input. However, since both $\mathbf{G}$ and $\mathbf{P}$ are cc-stateless permutations, repeated inputs will only produce repeated outputs and hence cannot help the distinguisher. □

6 Conclusion and Further Research

Our main theorem unifies the claims of both Theorem 1 and Theorem 2 in [6] under reasonable assumptions. To see this, let us focus for example on the natural case of random functions, assuming $\mathbf{F} \equiv \mathbf{G}$ and $\mathbf{I} \equiv \mathbf{J} \equiv \mathbf{R}$. Our theorem gives a better bound than Theorem 2 in [6] as long as $\Lambda_k(\mathbf{F}, \mathbf{R}) < 1/2$. It also improves the bound from Theorem 1 in [6] as long as

$$\frac{\Lambda_k(\mathbf{F}, \mathbf{R})}{\Delta_k(\mathbf{F}, \mathbf{R})} < \frac{1}{2} \cdot \frac{\Delta_k(\mathbf{F}, \mathbf{R})}{\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{R})}.$$

This means, loosely speaking, that the improvement occurs as long as the ratio of advantage gained from the free choice of state is smaller than the ratio of advantage gained from extending the distinguisher class.

This improvement is significant for any random function $\mathbf{F}$ that satisfies the conditions

$$\Delta_k^{\mathsf{NA}}(\mathbf{F}, \mathbf{R}) \ll \Delta_k(\mathbf{F}, \mathbf{R}) \approx \Lambda_k(\mathbf{F}, \mathbf{R}) \ll 1.$$

As an example, consider the simple cc-stateless random function $\mathbf{F}\colon \{0,1\}^n \to \{0,1\}^n$ that behaves as follows: with probability $2^{-n/2}$ it satisfies the (adaptively verifiable) condition $\mathbf{F}(\mathbf{F}(0)) = 0$ and the remaining values (including $\mathbf{F}(0)$) are chosen uniformly at random, in the rest of the cases (with probability $1 - 2^{-n/2}$) $\mathbf{F}$ behaves exactly like $\mathbf{R}$.

In general, a small $\Delta_k(\mathbf{F}, \mathbf{R})$ does not necessarily imply a small $\Lambda_k(\mathbf{F}, \mathbf{R})$, since it is easy to construct a counterexample where some specific initial transcript leads to a behavior that is easy to distinguish from the ideal system. However, a small value of $\Lambda_k(\mathbf{F}, \mathbf{R})$ may be considered a desirable requirement for a good quasi-random function.

Although it is not difficult to define the concept of free-start distinguishing in the computational setting, our main result does not translate to this setting. This is because such a translation would imply that for example composition of non-adaptively secure pseudo-random permutations is adaptively secure, which would contradict the results in [10] under standard assumptions. Therefore, the implications of our result for the computational setting remain an open question.

Acknowledgements. This research was partially supported by the Swiss National Science Foundation (SNF) project no. 200020-113700/1 and by the grants VEGA 1/0266/09 and UK/385/2009.

References

1. Dodis, Y., Impagliazzo, R., Jaiswal, R., Kabanets, V.: Security Amplification for Interactive Cryptographic Primitives. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 128–145. Springer, Heidelberg (2009)
2. Luby, M., Rackoff, C.: Pseudo-random Permutation Generators and Cryptographic Composition. In: STOC 1986, pp. 356–363 (1986)
3. Maurer, U.: Indistinguishability of Random Systems. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 110–132. Springer, Heidelberg (2002)
4. Maurer, U., Oswald, Y.A., Pietrzak, K., Sjödin, J.: Luby-Rackoff Ciphers with Weak Round Functions. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 391–408. Springer, Heidelberg (2006)
5. Maurer, U., Pietrzak, K.: Composition of Random Systems: When Two Weak Make One Strong. In: Naor, M. (ed.) TCC 2004. LNCS, vol. 2951, pp. 410–427. Springer, Heidelberg (2004)
6. Maurer, U., Pietrzak, K., Renner, R.: Indistinguishability Amplification. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 130–149. Springer, Heidelberg (2007)

7. Maurer, U., Tessaro, S.: Computational Indistinguishability Amplification: Tight Product Theorem for System Composition. In: Halevi, S. (ed.) Advances in Cryptology - CRYPTO 2009. LNCS, vol. 5677, pp. 355–373. Springer, Heidelberg (2009)
8. Myers, S.: On the Development of Blockciphers and Pseudo-random Function Generators Using the Composition and XOR Operators, M.Sc. Thesis (1999)
9. Myers, S.: Efficient Amplification of the Security of Weak Pseudo-random Function Generators. Journal of Cryptology 16(1), 1–24 (2003)
10. Pietrzak, K.: Composition Does Not Imply Adaptive Security. In: Shoup, V. (ed.) CRYPTO 2005. LNCS, vol. 3621, pp. 55–65. Springer, Heidelberg (2005)
11. Vaudenay, S.: Decorrelation: A Theory for Block Cipher Security. Journal of Cryptology 16(4), 249–286 (2003)