Computational Indistinguishability Amplification: Tight Product Theorems for System Composition*

Ueli Maurer Stefano Tessaro

Department of Computer Science
ETH Zurich
8092 Zurich, Switzerland
{maurer,tessaros}@inf.ethz.ch

Abstract

Computational indistinguishability amplification is the problem of strengthening cryptographic primitives whose security is defined by bounding the distinguishing advantage of an efficient distinguisher. Examples include pseudorandom generators (PRGs), pseudorandom functions (PRFs), and pseudorandom permutations (PRPs).

The literature on computational indistinguishability amplification consists only of few isolated results. Yao's XOR-lemma implies, by a hybrid argument, that no efficient distinguisher has advantage better than (roughly) $n2^{m-1}\delta^m$ in distinguishing the XOR of m independent n-bit PRG outputs $S_1, \ldots, S_m$ from uniform randomness if no efficient distinguisher has advantage more than δ in distinguishing S_i from a uniform n-bit string. The factor 2^{m-1} allows for security amplification only if $\delta < \frac{1}{2}$: For the case of PRFs, a random-offset XOR-construction of Myers was the first result to achieve *strong* security amplification, i.e., also for $\frac{1}{2} \leq \delta < 1$.

This paper proposes a systematic treatment of computational indistinguishability amplification. We generalize and improve the above product theorem for the XOR of PRGs along five axes. First, we prove the *tight* information-theoretic bound $2^{m-1}\delta^m$ (without factor n) also for the computational setting. Second, we prove results for *interactive* systems (e.g. PRFs or PRPs). Third, we consider the general class of *neutralizing combination constructions*, not just XOR. As an application, this yields the first indistinguishability amplification results for the cascade of PRPs (i.e., block ciphers) converting a weak PRP into an arbitrarily strong PRP, both for single-sided and two-sided queries. Fourth, *strong* security amplification is achieved for a subclass of neutralizing constructions which includes as a special case the construction of Myers. As an application we obtain highly practical optimal security amplification for block ciphers, simply by adding random offsets at the input and output of the cascade. Fifth, we show strong security amplification also for *weakened assumptions* like security against random-input (as opposed to chosen-input) attacks.

A key technique is a generalization of Yao's XOR-lemma to (interactive) systems, which is of independent interest.

*An extended abstract of this paper appears in the proceedings of CRYPTO 2009. This is the full version.

1 Introduction

1.1 Security Amplification

The security of all computationally secure cryptographic systems, even of those called "provably secure" in the literature, relies on unproven assumptions about the underlying cryptographic primitives. Typical assumptions are that a certain construction is a one-way function (OWF), a collision-resistant hash function, a pseudorandom generator (PRG), a pseudorandom function (PRF), a pseudorandom permutation (PRP), etc. To weaken these assumptions is both a fundamental challenge in the theory of cryptography and a major goal for the cautious and prudent design of practical cryptographic systems. Many reductions of strong primitives to weak primitives are known. For example, one of the outstanding results is the construction of a PRG from any OWF [14]. However, this reduction, like many other reductions, is highly inefficient and, while of high theoretical value, not of practical relevance.

A specific way to weaken an assumption is to require only that the security property is mildly true. For instance, a δ-OWF can be efficiently inverted with probability at most δ (rather than a negligible quantity for a regular OWF). Similarly, for a δ-PRG no efficient distinguisher has an advantage more than δ in distinguishing its output from a uniform random string. The corresponding definitions of a δ-PRF, a δ-PRP, etc., are straight-forward. Such a weakened assumption is more likely to be true. For example, it is more conservative to only assume that AES is a 0.99-PRP rather than a fully secure PRP.

The natural question is whether several weak primitives can be efficiently combined to obtain a stronger version of the primitive, ideally one with the full-fledged security property.[1] This is called *security amplification*, in some cases hardness amplification. The classical result on security amplification due to Yao [37] is that the parallel composition of m δ-OWFs (where $\delta < 1$) results in a $(\delta^m + \nu)$-OWF, where ν is some negligible quantity and δ^m can be made negligible for large enough m. Security amplifications of a wide range of cryptographic primitives has subsequently been considered, including for example regular OWFs and OWPs [10, 12], two-party protocols [1, 31, 32, 36, 13], key-agreement and public-key encryption [7, 17, 18], collision-resistant hash functions [4], and watermarking schemes [19].[2]

The term *indistinguishability amplification* refers to security amplification when the relevant security quantity is the *distinguishing advantage* for the best distinguisher from a certain class of distinguishers, typically the class of efficient distinguishers.

1.2 The XOR-Lemma and Amplification for PRGs

Before we discuss the XOR-lemma, let us compare the prediction advantage and the distinguishing advantage of a biased bit, in an information-theoretic setting, i.e., allowing arbitrary computing power. A bit with bias ϵ takes on the two values with probabilities $\frac{1}{2} - \epsilon$ and $\frac{1}{2} + \epsilon$. When such a bit must be guessed, one would choose the more likely value and be correct with probability $\frac{1}{2} + \epsilon$. To calibrate the guessing advantage, between 0 (when $\epsilon = 0$) and 1 (when the bit is fixed, i.e., $\epsilon = \frac{1}{2}$), one defines the advantage to be 2ϵ. In contrast, the distinguishing advantage is defined as

[1] Typically one considers several independent instantiations of the *same* weak primitive, but most results actually hold for several different instantiations.

[2] So-called combiners [15] are another method for relaxing security assumptions: They guarantee that a construction involving several instantiations of a primitive is (fully) secure if at least one (or several, but not all) of them are (fully) secure. However, they do not amplify the security of the underlying primitives.

ϵ (with no factor 2) since it is naturally defined for general random variables (not only bits) as the distance of the probability distribution from the uniform one.

As an example, consider two independent bits with biases ϵ_1 and ϵ_2. It is easy to see that the bias of the XOR is $2\epsilon_1\epsilon_2$. For instance, the XOR of a 0.1-biased bit (40/60) and a 0.2-biased bit (30/70) is a 0.04-biased bit (46/54), where $0.04 = 2 \cdot 0.01 \cdot 0.02$. More generally, the bias of the XOR of m bits is 2^{m-1} times the product of the biases. For the XOR of m bit-strings $S_1, \ldots, S_m$ of length n, where S_i has distance δ_i from a uniform n-bit string, the distance from uniform of the XOR of the strings, $S_1 \oplus S_2 \oplus \cdots \oplus S_m$, is bounded by $2^{m-1} \prod_{i=1}^{m} \delta_i$. This bound is tight, as for example the case $n = 1$ discussed above illustrates.

Let us now move to the computational setting, i.e., to Yao's XOR-lemma [37, 11], which is much more involved and is another seminal security amplification result. One typically considers a predicate $B(x)$ of the input of a OWF f which is hard to guess when given the output $f(x)$, for uniformly chosen x. But the setting of the XOR-lemma is more general. It states[3] that if for bits $B_1, \ldots, B_m$ the advantage in guessing B_i given some correlated information X_i is at most α_i for any algorithm with complexity t', then no algorithm with complexity t has advantage more than $\prod_{i=1}^{m} \alpha_i + \gamma$ in guessing their XOR-sum, i.e., $B_1 \oplus \cdots \oplus B_m$, given $X_1, \ldots, X_m$, where γ can be made arbitrarily small, at the cost of making t smaller with respect to t'.[4] In terms of distinguishing advantages δ_i, the bound is $2^{m-1} \prod_{i=1}^{m} \delta_i + \gamma$ (for the reasons described above).

Moreover, a standard hybrid argument, to use the unpredictability of bits to prove the indistinguishability of bit-strings, implies an indistinguishability amplification result for PRGs. Consider m independent PRG outputs, $S_1, \ldots, S_m$, each an n-bit string. If no distinguisher with complexity t' has advantage more than δ_i in distinguishing S_i from a uniform random n-bit string, then no distinguisher with complexity (roughly) t has advantage more than $n(2^{m-1} \prod_{i=1}^{m} \delta_i + \gamma)$ in distinguishing $S_1 \oplus S_2 \oplus \cdots \oplus S_m$ from a uniform random n-bit string.[5] The factor n comes from the hybrid argument over the individual bits of the bit-string.

As explained, the factor 2^{m-1} is unavoidable, since it holds even in the information-theoretic setting. Unfortunately, it means that an amplification can be achieved only if the component constructions are better than $\frac{1}{2}$-secure, i.e., if $\delta_i < \frac{1}{2}$.

1.3 Natural Questions and Previous Results

The above discussion suggests a number of natural questions:

(1) Can the factor n in the bound for the XOR of PRGs be eliminated, to obtain a tight bound, namely the equivalent of the information-theoretic counterpart?

(2) Can the result be extended to the XOR of PRFs, i.e., primitives for which the security is defined by an *interactive* game, not by the (static) indistinguishability of random variables?

(3) If the answer is "yes", can such a result be extended to other constructions, most importantly the cascade of PRPs?

[3] In fact, one needs a "tight" version of the XOR-lemma for this statement to hold, such as the one by Levin [22, 11], or one obtained from a tight hard-core lemma (e.g. [17]) via the techniques of [20].

[4] As usual in complexity-theoretic hardness amplification, we experience an *unavoidable* [33] trade-off between the choice of γ (the tightness of the bound) and the complexity of the reduction.

[5] It is not clear to us whether this fact has been published, or is unpublished but well-known folklore, or not so well-known (see also [6] for a similar statement about security amplification for the XOR of PRGs).

(4) Can the factor 2^{m-1} be eliminated so that security amplification from arbitrarily weak components can be achieved?

We will answer all these questions positively.

In the information-theoretic setting, questions 2 and 3 were answered by Maurer, Pietrzak, and Renner [26], whose abstract approach we follow, and the special case of permutations had previously been solved by Vaudenay [34, 35]. In contrast, there are only a few isolated results on *computational* indistinguishability amplification, which we now discuss. Myers [29] was the first to consider security amplification for PRFs. Interestingly, he did not solve question 2 above, which remained open, but he actually solved part of question 4. More precisely, he showed for the XOR of PRFs, with the modification that for each PRF a random (secret) offset is XORed to the input, that the stronger bound (without the factor 2^{m-1}) can be achieved. However, his treatment is specific for his construction and does not extend to other settings like the cascade of PRPs. Dodis et al. [6] addressed question 2 and gave a positive answer using techniques originating from the setting of hardness amplification of weakly verifiable puzzles [3, 21]. However, their focus is on general interactive cryptographic primitives, including for example message authentication codes (MACs), and the resulting bound for the case of PRFs depends on the number of queries the distinguisher is allowed to ask and is thus far from the (tight) information-theoretic bound shown in [26].

Little is known about the cascade of weak PRPs, which is perhaps the case of highest practical interest as it addresses security amplification for block ciphers.[6] Luby and Rackoff [23] proved an amplification result for the cascade of *two* weak PRPs. This result was extended by Myers [28] to the cascade of a small number of PRPs, but he notes that this result falls short of constructing a (regular) PRP from a weak PRP and states this as an open problem, which we solve.

1.4 Contributions of this Paper

In our attempt at solving the different open questions explained above, we take a very general approach, not targeted at specific constructions. The goal is to develop a deeper and more general understanding and to prove results of a generality that can be useful for other applications.

A first result is a generalization of the XOR-lemma to interactive systems. If a system (as opposed to a random variable for the standard XOR-lemma) of a general type depends on a bit, and no efficient algorithm with access to the system can predict the bit better than with a certain advantage, then the advantage in predicting the XOR of several such bits is bounded by the product of the individual advantages, even if the predictor has complete and arbitrary independent access to all the involved systems.

The XOR of strings or (of the output) of systems, as well as the cascade of systems implementing permutations, are both special cases of a more general concept which was called *neutralizing construction* in [26]. Intuitively, a construction involving several component systems is neutralizing if it is equivalent to an ideal system whenever one component is ideal. For example, the XOR of several PRFs is equivalent to a truly random function if (any) one of the PRFs is replaced by a truly random function.

[6] Cascades of block ciphers were considered by Even and Goldreich [8] and Maurer and Massey [25], but those results only prove that the cascade is as secure as the strongest component (with no amplification), i.e., that the cascade is a combiner for encryption. Bellare and Rogaway [2] showed a certain security amplification (of a different type) for cascade encryption in the ideal cipher model, which is a purely information-theoretic consideration.

We prove two tight general product theorems. The first theorem relies on the XOR-lemma and shows that for *all* neutralizing constructions the distinguishing advantage of the combined system is 2^{m-1} times the product of the individual advantages, which is optimal. The second theorem gets rid of the factor 2^{m-1} by considering a special class of *randomized* neutralizing constructions. The applications mentioned in the abstract and the previous sections follow directly from these general theorems.[7] In particular, one application is a highly practical construction for optimal security amplification for block ciphers, simply by adding random offsets at the input and output of the cascade.

This paper adopts a *concrete* approach, i.e. we do not use asymptotics and statements are inherently non-uniform. Still, all results can be extended to the uniform setting by using standard techniques, and we informally comment on the necessary changes in order to derive a uniform version of the given statements.

1.5 Outline of This Paper

This paper is organized as follows. Section 2 introduces the tools needed throughout this paper. In particular, we review the notions of interactive systems and constructions, and provide a formal definition of the central concept of a neutralizing construction.

Section 3 presents the first main result of this paper, i.e., a generalization of Yao's XOR-lemma to the setting of interactive systems. This result is subsequently used in Section 4 to derive a general product theorem for arbitrary neutralizing constructions. We also present applications of this result, including the first security amplification result for the cascade of (possibly two-sided) PRPs, as well as improved bounds for the XOR of PRFs.

Moreover, Section 5 presents a strong product theorem for a general class of randomized neutralizing constructions. Finally, we discuss some applications of this result. In particular, we prove strong security amplification result for PRPs, as well as strong security amplification for PRFs which are secure when evaluated at random inputs.

2 Preliminaries

2.1 Notational Preliminaries and Computational Model

Throughout this paper, we use calligraphic letters $\mathcal{X}, \mathcal{Y}, \ldots$ to denote sets, upper-case letters $X, Y, \ldots$ to denote random variables, and lower-case letters $x, y, \ldots$ denote the values they take. Moreover, $\mathsf{P}[\mathcal{A}]$ denotes the probability of an event $\mathcal{A}$, while we use the shorthands $\mathsf{P}_X(x) := \mathsf{P}[X = x]$, $\mathsf{P}_{X|Y}(x, y) := \mathsf{P}[X = x|Y = y]$, $\mathsf{P}_{X\mathcal{A}|Y\mathcal{B}}(x, y) := \mathsf{P}[\mathcal{A} \wedge X = x|\mathcal{B} \wedge Y = y]$, etc. Also, P_X, $\mathsf{P}_{X|Y}$, $\mathsf{P}_{\mathcal{A}X|\mathcal{B}Y}$ denote the corresponding (conditional) probability distributions,[8] and $x \leftarrow \mathsf{P}_X$ is the action of sampling a concrete value x according to the distribution P_X. Finally, $\mathsf{E}[X]$ is the expected value of the (real-valued) random variable X.

We consider *interactive* randomized stateful algorithms in some a-priori fixed (but otherwise unspecified) RAM model of computation. In particular, such an algorithm keeps a state (consisting, say, of the contents of the memory space it employs), and answers each query depending on the

[7] For each application of the second theorem, one also needs an information-theoretic indistinguishability proof based on the conditional equivalence of two systems, conditioned on an event that must be proved to be unlikely to occur.

[8] In particular, $\mathsf{P}_{X|Y}$ and $\mathsf{P}_{\mathcal{A}X|\mathcal{B}Y}$ take *two* arguments corresponding to all possible values of X and Y, respectively.

input of this query, some coin flips, the current state (which is possibly updated), and (possibly) one or more queries to an underlying system. It is also convenient to denote by $A[\sigma]$ the algorithm obtained by *setting* the state of A to σ (provided σ is a compatible state), and then behaving according to A's description. Additionally, we say that the algorithm A has *time complexity* t_A (where t_A is a function $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$) if the sum of the length of the description of A and the total number of steps of A is at most $t_A(q, s)$ for all sequences of q queries, all compatible initial states with size s, and all compatible interactions with an underlying system. We use the shorthand $t_A(q) := t_A(q, 0)$.

2.2 Discrete Systems and Constructions

Discrete Systems, Constructions, and Distinguishers. This paper deals with the general notion of a (single-interface) *discrete system* $\mathbf{F}$ taking inputs $X_1, X_2, \ldots$ and returning outputs $Y_1, Y_2, \ldots$, where the i-th output Y_i depends (probabilistically) on the first i inputs $X^i = [X_1, \ldots, X_i]$ as well as on all previous $i-1$ outputs $Y^{i-1} = [Y_1, \ldots, Y_{i-1}]$. (If all inputs and outputs are in sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, we call $\mathbf{F}$ an $(\mathcal{X}, \mathcal{Y})$-*system*.) Its input-output behavior is minimally described (see e.g. [24]) by the (infinite) sequence of conditional probability distributions $\mathsf{p}^{\mathbf{F}}_{Y_i|X^iY^{i-1}}$ (for all $i \geq 1$).[9] In general, we use the name "system" (as well as $\mathbf{F}$) interchangeably to denote both the input-output behavior determined by conditional probability distributions and an actual discrete system realizing this behavior. It thus makes sense to say that two systems $\mathbf{F}$ and $\mathbf{G}$ are *equivalent* (denoted $\mathbf{F} \equiv \mathbf{G}$) if they have the same input-output behavior. A *random variable* X is the simplest type of system, which answers its first query with the value X (and ignoring any given input), and ignores any further query.[10]

With $\mathbf{C}(\cdot)$ we denote a *construction* invoking one or more underlying compatible *subsystems*, whereas $\mathbf{C}(\mathbf{F})$, $\mathbf{C}(\mathbf{F}, \mathbf{G})$, etc denote the systems obtained when $\mathbf{C}$ is instantiated with $\mathbf{F}$ (and $\mathbf{G}$). The shorthand $\mathbf{C}(\mathbf{F}, \cdot)$ indicates the construction that behaves as $\mathbf{C}(\mathbf{F}, \mathbf{G})$ given access to the subsystem $\mathbf{G}$. (All notations extend naturally to constructions with more than two subsystems.) A *distinguisher* $\mathbf{D}$ is a system interacting with another system $\mathbf{F}$ giving inputs $X_1, X_2, \ldots$ and obtaining outputs $Y_1, Y_2, \ldots$, outputting a decision bit after a certain number q of queries depending on the transcript (X^q, Y^q): In particular, we denote as $\mathsf{P}[\mathbf{D}(\mathbf{F}) = 1]$ the probability that it outputs 1.

In a broader context (such as within the pseudocode description of a random experiment), the notation $\mathbf{D}(\mathbf{F})$ (or $\mathbf{D}(\mathbf{C}(\mathbf{F}))$) is the binary random variable which consists of the binary output of an *independent* copy of $\mathbf{D}$ run on a fresh independent instance of $\mathbf{F}$ (or $\mathbf{C}(\mathbf{F})$). However, if $\mathbf{F}$ is an understood given system (for instance in the description of a reduction), then the bit $\mathbf{D}(\mathbf{F})$ (or $\mathbf{D}(\mathbf{C}(\mathbf{F}))$) is sampled by letting a new independent copy of $\mathbf{D}$ (and possibly $\mathbf{C}(\cdot)$) interact with the *same* instance of the given system.

We say that an interactive algorithm A implements a system $\mathbf{F}$ or a construction $\mathbf{C}(\cdot)$ if it has the same input-output behavior as $\mathbf{F}$ and $\mathbf{C}(\cdot)$, respectively. In particular, we use A (rather than $\mathbf{F}$) whenever we want to stress that we use the particular implementation A of $\mathbf{F}$.

[9] We use lower-case p, rather than P, in order to stress that these *conditional* probability distributions alone do not define a random experiment.

[10] However, in order to render the presentation of this paper more homogenous, we usually take the convention that the random variable answers infinely many queries by returning the *same* value as in its first query.

Distinguishing Advantages. The *distinguishing advantage of a distinguisher* $\mathbf{D}$ *in distinguishing two systems* $\mathbf{F}$ *and* $\mathbf{G}$ is the quantity

$$\Delta^{\mathbf{D}}(\mathbf{F},\mathbf{G}) := \left|\mathsf{P}[\mathbf{D}(\mathbf{F})=1] - \mathsf{P}[\mathbf{D}(\mathbf{G})=1]\right|.$$

We denote as $\Delta_t(\mathbf{F},\mathbf{G})$, $\Delta_q(\mathbf{F},\mathbf{G})$, and $\Delta_{t,q}(\mathbf{F},\mathbf{G})$ the best distinguishing advantages $\Delta^{\mathbf{D}}(\mathbf{F},\mathbf{G})$ taken over all distinguishers with time complexity at most t, issuing at most q queries, or both, respectively.[11]

System Composition. Given m systems $\mathbf{F}_1,\ldots,\mathbf{F}_m$, we use the shorthand $\mathbf{F}_1\|\ldots\|\mathbf{F}_m$ to denote their *parallel composition*, i.e., the system allowing parallel concurrent interaction with the (independent) m systems.[12] Moreover, for $(\mathcal{X},\mathcal{Y})$-systems $\mathbf{F}$ and $\mathbf{G}$, and for a random bit B (with distribution P_B), the system $\langle\mathbf{F},\mathbf{G}\rangle_B$ samples the bit B and acts as the system $\mathbf{F}$ if $B=0$, and $\mathbf{G}$ otherwise. Additionally, for any *quasi-group operation*[13] $\star$ on $\mathcal{Y}$ the $(\mathcal{X},\mathcal{Y})$-system $\mathbf{F}\star\mathbf{G}$ on input x invokes both $\mathbf{F}$ and $\mathbf{G}$ with input x, obtaining y and y', and returns $y\star y'$.[14] Also, for an $(\mathcal{X},\mathcal{Y})$-system $\mathbf{F}$ and a $(\mathcal{Y},\mathcal{Z})$-system $\mathbf{G}$ we denote with $\mathbf{F}\rhd\mathbf{G}$ the *cascade of* $\mathbf{F}$ *and* $\mathbf{G}$, i.e., the system which on input x first invokes $\mathbf{F}$ on this input, and the resulting output is fed into $\mathbf{G}$ to obtain the final output.

Stateless Systems. We say that a system $\mathbf{F}$ is *stateless* if there exists a conditional probability distribution $\mathsf{p}^{\mathbf{F}}_{Y|X}$ such that $\mathsf{p}^{\mathbf{F}}_{Y_i|X^iY^{i-1}}(y_i,x^i,y^{i-1}) = \mathsf{p}^{\mathbf{F}}_{Y|X}(y_i,x_i)$ for all $y_i,x^i=[x_1,\ldots,x_i]$, and $y^{i-1}=[y_1,\ldots,y_{i-1}]$. Moreover, the system $\mathbf{F}$ is *convex-combination stateless* (*cc-stateless*, for short) if there exists a random variable S and a construction $\mathbf{F}(\cdot)$ such that $\mathbf{F}(S)\equiv\mathbf{F}$ (note that we use the same letter to denote both the system itself as well as the corresponding construction), and $\mathbf{F}(s)$ is stateless for *all* values s taken by S. Depending on the context, S may e.g. be a seed, a key, or an internal function table. A non-trivial example of a cc-stateless system is a randomized encryption scheme, which takes a secret key and encrypts each message with independent randomness. Note that $\langle\mathbf{F},\mathbf{G}\rangle_B$ is cc-stateless if both $\mathbf{F},\mathbf{G}$ are cc-stateless.

Random Functions. A *random function* $\mathbf{F}:\mathcal{X}\to\mathcal{Y}$ is an $(\mathcal{X},\mathcal{Y})$-system which answers consistently, i.e. $X_i=X_j$ implies $Y_i=Y_j$. It is called a *random permutation* if additionally $Y_i=Y_j$ implies $X_i=X_j$. For a cc-stateless random function $\mathbf{F}:\mathcal{X}\to\mathcal{Y}$ with $\mathbf{F}\equiv\mathbf{F}(S)$, it is easy to see that the system $\mathbf{F}(s)$ is a (deterministic) function $\mathcal{X}\to\mathcal{Y}$ for all s. (This is sometimes called a *keyed function family*, but we also consider the case where s is huge and is hence not a key.) Special cases are a *uniform random function* (URF) $\mathbf{R}:\mathcal{X}\to\mathcal{Y}$ and a *uniform random permutation* (URP) $\mathbf{P}:\mathcal{X}\to\mathcal{X}$ that realize a uniformly chosen function $\mathcal{X}\to\mathcal{Y}$ and permutation $\mathcal{X}\to\mathcal{X}$, respectively. For simplicity, but with some abuse of notation, we denote as $\mathbf{F}(s,x)$ the evaluation of $\mathbf{F}(s)$ with input x.

Informally, in an asymptotic setting, it is convenient to say that an efficient $\mathbf{F}(\cdot)$ is a δ-*pseudorandom function* (PRF) if $\Delta_{t,q}(\mathbf{F}(S),\mathbf{R})\le\delta+\mathsf{negl}$ for a (short) key S, a URF $\mathbf{R}$, all polynomially bounded t and q, and some negligible[15] function negl. Analogously, if an efficient

[11] Despite notational overloading, the considered advantage notion will be always clear from the context.

[12] The systems do not interact with each other, and each query to the parallel composition is addressed to one of the systems.

[13] That is, given $a,c\in\mathcal{Y}$ (or $b,c\in\mathcal{Y}$) there exists a unique b (a) such that $a\star b=c$. An example is bit-wise XOR $\oplus$ for $\mathcal{Y}=\{0,1\}^n$, but any group operation is a quasi-group operation as well.

[14] We denote as $\mathbf{F}_1\star\cdots\star\mathbf{F}_m$ the system $(\cdots((\mathbf{F}_1\star\mathbf{F}_2)\star\mathbf{F}_3)\cdots)\star\mathbf{F}_m$.

[15] Recall that a function $\nu:\mathbb{N}\to\mathbb{R}_{\ge0}$ is *negligible* if it vanishes faster than the inverse of any polynomial.

$\mathbf{Q}(\cdot)$ implements a permutation for all keys, it is called a *δ-pseudorandom permutation (PRP)* if $\Delta_{t,q}(\mathbf{Q}(S), \mathbf{P}) \leq \delta + \mathsf{negl}$ for a URP $\mathbf{P}$ and for all polynomially bounded t and q.

The inverse $\mathbf{Q}^{-1}$ of a cc-stateless random permutation $\mathbf{Q}$ is well-defined, and $\langle\mathbf{Q}\rangle$ is the system accepting *forward queries* $(x, +)$ (answered by $\mathbf{Q}(s, x)$ on key s) and *backward queries* $(y, -)$ (answered as $\mathbf{Q}^{-1}(s, y)$). In particular $\langle\mathbf{Q}\rangle \rhd \langle\mathbf{Q}'\rangle$ stands for the system $\langle\mathbf{Q} \rhd \mathbf{Q}'\rangle$. An efficient $\mathbf{Q}(\cdot)$ is called a *δ-two-sided PRP*[16] if $\Delta_{t,q}(\langle\mathbf{Q}(S)\rangle, \langle\mathbf{P}\rangle) \leq \epsilon + \mathsf{negl}$ for all polynomial q and t. (Of course, one assumes that backward queries can be computed efficiently given s.)

Neutralizing Constructions. A construction $\mathbf{C}(\cdot)$ is called *neutralizing* [26] for system classes $\mathcal{F}_1, \ldots, \mathcal{F}_m$ and ideal systems $\mathbf{I}_1 \in \mathcal{F}_1, \ldots, \mathbf{I}_m \in \mathcal{F}_m$, if for all $\mathbf{S}_i \in \mathcal{F}_i$ $(i = 1, \ldots, m)$ we have $\mathbf{C}(\mathbf{S}_1, \ldots, \mathbf{S}_m) \equiv \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)$ whenever there exists some i with $\mathbf{S}_i = \mathbf{I}_i$.[17] In particular, neutralizing constructions capture the notion of a *combiner* [15] for computational indistinguishability properties: Whenever at least one system $\mathbf{S}_i \in \mathcal{F}_i$ is computationally indistinguishable from $\mathbf{I}_i$, then $\mathbf{C}(\mathbf{S}_1, \ldots, \mathbf{S}_m)$ is computationally indistinguishable from $\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)$. Typical examples for the classes $\mathcal{F}_i$ are the classes of random variables, random functions, and random permutations. Without loss of generality, we assume that each such system class $\mathcal{F}$ satisfies the natural property that if $\mathbf{F} \in \mathcal{F}$ then fixing (even partially) the random choices in some implementation of $\mathbf{F}$ yields a system in $\mathcal{F}$ as well. In particular, it is convenient to say that $\mathbf{C}(\cdot)$ is neutralizing for $\mathbf{F}_1, \ldots, \mathbf{F}_m$ and $\mathbf{I}_1, \ldots, \mathbf{I}_m$ if it is neutralizing for some classes $\mathcal{F}_1, \ldots, \mathcal{F}_m$ with $\mathbf{F}_1, \mathbf{I}_1 \in \mathcal{F}_1, \ldots, \mathbf{F}_m, \mathbf{I}_m \in \mathcal{F}_m$.

Every quasi-group operation $\star$ on a set $\mathcal{Y}$ induces a construction $\mathbf{C}(\cdot)$ such that $\mathbf{C}(\mathbf{F}, \mathbf{G}) := \mathbf{F} \star \mathbf{G}$ which is neutralizing for any two random functions $\mathbf{F}, \mathbf{G} : \mathcal{X} \to \mathcal{Y}$ and ideal systems $\mathbf{I}, \mathbf{J}$ being independent URFs. In particular, $\mathbf{I} \star \mathbf{J}$ is also a URF. As a special case, this result holds for random variables X, Y over $\mathcal{Y}$, the ideal systems being uniform random elements of $\mathcal{Y}$. Moreover, the cascade operator $\rhd$ induces a construction $\mathbf{C}'(\cdot)$ with $\mathbf{C}'(\mathbf{Q}_1, \mathbf{Q}_2) := \mathbf{Q}_1 \rhd \mathbf{Q}_2$ which is neutralizing for any two cc-stateless random permutations $\mathbf{Q}_1, \mathbf{Q}_2 : \mathcal{X} \to \mathcal{X}$ (in fact $\mathbf{Q}_1$ can possibly be stateful) where the ideal systems $\mathbf{I}, \mathbf{J}$ are both URPs $\mathcal{X} \to \mathcal{X}$. In particular, $\mathbf{I} \rhd \mathbf{J}$ is also a URP. If $\mathbf{Q}_1$ is cc-stateless, then the same result holds even in the two-sided case for $\langle\mathbf{Q}_1\rangle$ and $\langle\mathbf{Q}_2\rangle$ (with ideal system $\langle\mathbf{P}\rangle$ for a URP $\mathbf{P}$). Both constructions extend naturally to an arbitrary number of subsystems.

2.3 Indistinguishability Proofs

This section provides a self-contained introduction to some specific tools from the random systems framework [24, 26] which are used in the second part of this paper.

Monotone Event Sequences and Indistinguishability. Given a system $\mathbf{F}$, a *monotone event sequence (MES)* $\mathcal{A} = A_0, A_1, \ldots$ on $\mathbf{F}$ is a sequence of events where A_i is defined after the i'th query has been answered by $\mathbf{F}$ and such that A_0 initially holds (i.e., before the first query is issued). Furthermore, if A_i does not hold for some $i > 0$ (i.e. the complement $\overline{A_i}$ holds), then A_j does not hold for all $j \geq i$.

Definition 1. Let $\mathbf{F}, \mathbf{G}$ be systems, and let $\mathcal{A}$ be a MES on $\mathbf{F}$. We write $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}$ if for all $i \geq 1$ and for all y_i, x^i, y^{i-1},

$$\mathsf{p}^{\mathbf{F}}_{Y_i|X^iY^{i-1}A_i}(y_i, x^i, y^{i-1}) = \mathsf{p}^{\mathbf{G}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1}).$$

[16] In the literature the name *strong* PRP is commonly used, but this term is slightly confusing in the context of this paper.

[17] This definition differs from the one given in [26] and in the proceedings version of this paper, but captures more naturally the concept of a neutralizing construction as a combiner.

Furthermore, if the MES $\mathcal{B}$ is additionally defined on $\mathbf{G}$, we write $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}|\mathcal{B}$ if for all $i \geq 1$ and for all y_i, x^i, y^{i-1},

$$\mathsf{p}^{\mathbf{F}}_{Y_i|X^iY^{i-1}A_{i-1}}(y_i, x^i, y^{i-1}) = \mathsf{p}^{\mathbf{G}}_{Y_i|X^iY^{i-1}B_{i-1}}(y_i, x^i, y^{i-1}).$$

The shorthand $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$ stands for the probability that the distinguisher $\mathbf{D}$ makes $\mathcal{A}$ fail within q queries while interacting with $\mathbf{F}$. The following lemma [24, 26] shows the connection between the probability of making a MES fail and the distinguishing advantage.[18]

Lemma 1. *Let $\mathbf{F}$ and $\mathbf{G}$ be systems, let $\mathbf{D}$ be a q-query distinguisher, and let $\mathcal{A}$ and $\mathcal{B}$ be monotone event sequences on $\mathbf{F}$ and $\mathbf{G}$, respectively.*

(i) If $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}$, then $\Delta^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) \leq \nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$.

(ii) If $\mathbf{F}|\mathcal{A} \equiv \mathbf{G}|\mathcal{B}$ and $\mathsf{p}^{\mathbf{F}}_{A_i|X^iY^{i-1}A_{i-1}}(x^i, y^{i-1}) \leq \mathsf{p}^{\mathbf{G}}_{B_i|X^iY^{i-1}B_{i-1}}(x^i, y^{i-1})$ for all (compatible) x^i, y^{i-1}, then $\Delta^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) \leq \nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$.

Upper Bounding $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$. We present two lemmas which simplify the problem of upper bounding $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$. The first one is, to our knowledge, new and considers the setting where a MES $\mathcal{A} = A_0, A_1, \ldots$ is defined on a system $\mathbf{F} \equiv \mathbf{F}(S)$ (i.e., where S is some internal variable of $\mathbf{F}$) such that the behavior of $\mathbf{F}$ is independent of S (i.e., $\mathbf{F}(s) \equiv \mathbf{F}$ for all values s taken by S). We show that if $\mathcal{A}$ only depends on the input-output behavior and the value of S, we can equivalently consider a game where the distinguisher $\mathbf{D}$ interacts with $\mathbf{F}$, generating a transcript X^qY^q, and only subsequently we sample an independent S, and check whether (X^q, Y^q, S) implies $\overline{A_q}$.

Lemma 2. *Let $\mathbf{F} = \mathbf{F}(S)$ be such that $\mathbf{F}(s) \equiv \mathbf{F}$ for all s. Let $\mathcal{A}$ be a MES such that for all $i \geq 1$ there exists a set $\mathcal{B}_i$ with the property that $\overline{A_i}$ holds if and only if $(X^i, Y^i, S) \in \mathcal{B}_i$. Then we have*

$$\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q}) \leq \max_{x^q, y^q} \mathsf{P}_S[(x^q, y^q, S) \in \mathcal{B}_q],$$

where the maximum is taken over all x^q, y^q which are compatible with an interaction of $\mathbf{D}$ with $\mathbf{F}$, and the probability on the right-hand side is taken over the choice of S.

Proof. Let $\mathbf{D}$ be a given distinguisher interacting with $\mathbf{F}$. Then we have (with $\mathsf{P}^{\mathbf{D}\,\mathbf{F}}$ being probabilities in the random experiment where $\mathbf{D}$ interacts with $\mathbf{F}$, and $\mathsf{p}^{\mathbf{F}}$ being (conditional) probabilities defined by the input-output behavior of $\mathbf{F}$)

$$\begin{aligned}
\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q}) = \sum_{x^q, y^q, s} \mathsf{P}^{\mathbf{D}\mathbf{F}}_{X^qY^qS\overline{A}_q}(x^q, y^q, s) &= \sum_{x^q, y^q, s} \mathsf{P}_S(s) \cdot \mathsf{P}^{\mathbf{D}\,\mathbf{F}}_{X^qY^q|S}(x^q, y^q, s) \cdot \mathsf{p}^{\mathbf{F}}_{\overline{A_q}|X^qY^qS}(x^q, y^q, s) \\
&= \sum_{x^q, y^q, s} \mathsf{P}_S(s) \cdot \mathsf{P}^{\mathbf{D}\,\mathbf{F}}_{X^qY^q}(x^q, y^q) \cdot \mathsf{p}^{\mathbf{F}}_{\overline{A_q}|X^qY^qS}(x^q, y^q, s) \\
&= \sum_{x^q, y^q} \mathsf{P}^{\mathbf{D}\,\mathbf{F}}_{X^qY^q}(x^q, y^q) \cdot \underbrace{\sum_{s} \mathsf{P}_S(s) \cdot \mathsf{p}^{\mathbf{F}}_{\overline{A_q}|X^qY^qS}(x^q, y^q, s)}_{=\mathsf{P}_S[(x^q, y^q, S) \in \mathcal{B}_q]},
\end{aligned}$$

and the claim follows from the fact that the maximum is at least as large as the average. □

[18] In fact [24] states these results in terms of $\Delta_q(\mathbf{F}, \mathbf{G})$ and the *best* $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$ for a q-query $\mathbf{D}$, but it is easy to see that this more concrete version also holds. (This allows us to apply the results to special subclasses of distinguishers.)

The following lemma [24] gives a simple condition under which it suffices to consider non-adaptive strategies to upper bound $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q})$.

Lemma 3. *Let* $\mathbf{F}$ *be a system with a MES* $\mathcal{A}$ *and assume that*

$$\mathsf{p}^{\mathbf{F}}_{A_i|X^iY^{i-1}A_{i-1}}(x^i, y^{i-1}) = \mathsf{p}^{\mathbf{F}}_{A_i|X^iA_{i-1}}(x^i)$$

for all sequences of inputs x^i *and all possible sequences of* $i-1$ *outputs* y^{i-1}*. Then, we have* $\nu^{\mathbf{D}}(\mathbf{F}, \overline{A_q}) \leq \max_{x^q} \mathsf{p}^{\mathbf{F}}_{\overline{A_q}|x^q}$*, where the maximum is taken over all sequences* x^q *of inputs which can ever be output by* $\mathbf{D}$*.*

3 The Generalized XOR-Lemma

3.1 System-Bit Pairs and Theorem Statement

System-Bit Pairs. A *system-bit pair* is a system of the form $(\mathbf{F}, B)$, where $B \in \{0,1\}$ is a bit value, which is (generally) correlated with the system $\mathbf{F}$. This can formally be described by the distribution P_B of B and the two systems $\mathbf{F}_0$ and $\mathbf{F}_1$ conditioned on the value taken by B, i.e. $(\mathbf{F}, B) = (\langle \mathbf{F}_0, \mathbf{F}_1 \rangle_B, B)$. An example of a possible system-bit pair is a URF $\mathbf{R} : \{0,1\}^m \to \{0,1\}$ and the parity of its function table.

The following quantity characterizes the performance of an adversary[19] $\mathbf{A}$ in guessing the bit B when given access to $\mathbf{F}$ only.

Definition 2. The *guessing advantage of an adversary* $\mathbf{A}$ *in guessing* B for a system-bit pair $(\mathbf{F}, B)$ is the quantity

$$\Gamma^{\mathbf{A}}(\mathbf{F}, B) := 2 \cdot \mathsf{P}[\mathbf{A}(\mathbf{F}) = B] - 1.$$

Additionally, we denote as $\Gamma_{t,q}(\mathbf{F}, B)$ the maximal guessing advantage $\Gamma^{\mathbf{A}}(\mathbf{F}, B)$ taken over all q-query adversaries $\mathbf{A}$ with complexity at most t.

Note that $\Gamma^{\mathbf{A}}(\mathbf{F}, B) \in [-1, 1]$, where 1 means that $\mathbf{A}$ is able to perfectly predict B by interacting with $\mathbf{F}$, while -1 means that $\mathbf{A}$ is never correct.[20] The following connection between the guessing and the distinguishing advantages is well known (cf. e.g. [26]).

Lemma 4. *For all* $\mathbf{F}$*,* $\mathbf{G}$*, and* $\mathbf{D}$*, and a uniform bit* B*,*

$$\Delta^{\mathbf{D}}(\mathbf{F}, \mathbf{G}) = \left| \Gamma^{\mathbf{D}}(\langle \mathbf{F}, \mathbf{G} \rangle_B, B) \right|.$$

Finally, note that if $(\mathbf{G}_i, B_i)$ is cc-stateless, then $\mathbf{G}_i$ is cc-stateless, but the converse is not always true: Consider the system-bit pair $(\mathbf{F}, B)$ such that B is a uniform random bit, and $\mathbf{F}$ has one-bit inputs and outputs: It answers the first query x_1 with a random bit y_1, and the second query x_2 is answered by $x_1 \oplus B$. All remaining queries are answered by independent random bits. Clearly, $\mathbf{F}$ by itself is (cc-)stateless (all answers are random and independent), but $(\mathbf{F}, B)$ is not cc-stateless since $y_2 := x_1 \oplus B$ must hold.

[19] We stress that distinguishers and adversaries are objects of the same type. The name adversary is used to stress the fact that we are not exclusively considering a distinguishing scenario.

[20] In particular, flipping the output bit of such an $\mathbf{A}$ yields one which is always correct.

THE XOR-LEMMA. Given m system-bit pairs $(\mathbf{G}_1, B_1), \ldots, (\mathbf{G}_m, B_m)$, we are interested in the advantage $\Gamma_{t,q_1,\ldots,q_m}(\mathbf{G}_1 \| \cdots \| \mathbf{G}_m, B_1 \oplus \cdots \oplus B_m)$ of guessing the bit $B_1 \oplus \cdots \oplus B_m$ given *parallel* access to the systems $\mathbf{G}_1, \ldots, \mathbf{G}_m$, where at most q_i queries to each system $\mathbf{G}_i$ are allowed. That is, we consider the most general attack where the adversary can query each subsystem $\mathbf{G}_i$ individually at most q_i times, *adaptively* depending on the answers of queries to other subsystems. We show that the advantage is upper bounded by the *product* of the individual advantages $\Gamma_{t',q'}(\mathbf{G}_i, B_i)$ for $i = 1, \ldots, m$ (for appropriate t', q'), with an extra additive term $\gamma > 0$ which can be made arbitrarily small (but influences the efficiency of the reduction). The result holds provided that all but one of the system-bit pairs are cc-stateless. Our result generalizes the original XOR-lemma by Yao [37, 11], which considered the special case of system-bit pairs (X_i, B_i), where X_i is a random variable.

We stress that our result only requires the ability to efficiently implement the cc-stateless system-bit pairs $(\mathbf{G}_i, B_i) = (\mathbf{G}_i(S), B_i(S))$. This may be possible, for instance by using a *stateful* algorithm, even if $\mathbf{G}(\cdot)$ and $B(\cdot)$ themselves are not efficiently computable: In fact, S may even be exponentially large. As an example, the aforementioned system-bit pair $(\mathbf{R}, B)$, where $\mathbf{R} : \{0,1\}^n \to \{0,1\}$ is a URF, and B is the parity of its function table, is clearly cc-stateless, and can efficiently be implemented by first sampling a random B, and then answering queries to $\mathbf{R}$ with independent random bits, with the exception of the last one, which is answered so that the parity equals B.

In the following, we define the quantity $\varphi := 2\left(\frac{24m}{\gamma}\right)^2 \cdot \ln\left(\frac{7m}{\gamma}\right)$ for understood m and γ. Also, t_{G_i} and s_{G_i} are the time and space[21] complexities of some implementation G_i of the system $\mathbf{G}_i$, whereas $t_{(G_i,B_i)}$ is the time-complexity of an implementation of the pair $(\mathbf{G}_i, B_i)$. (Note that an efficient implementation of the latter implies one for the former, but we allow for this distinction.) For all i, we denote $l_i := s_{G_i}(q_i \cdot \varphi)$ and $l_{<i} := \sum_{j=1}^{i-1} l_j$ (for understood $q_1, \ldots, q_{m-1}$).

Theorem 5 (XOR-Lemma)**.** *Let $(\mathbf{G}_1, B_1), \ldots, (\mathbf{G}_{m-1}, B_{m-1})$ be cc-stateless system-bit pairs, and let $(\mathbf{G}_m, B_m)$ be an arbitrary system-bit pair. For all t, $q_1, \ldots, q_m$, $\gamma > 0$,*

$$\Gamma_{t,q_1,\ldots,q_m}(\mathbf{G}_1 \| \ldots \| \mathbf{G}_m, B_1 \oplus \cdots \oplus B_m) \leq \prod_{i=1}^{m} \Gamma_{t'_i,q'_i}(\mathbf{G}_i, B_i) + \gamma,$$

where $t'_i := l_{<i} + \varphi \cdot \left[t + \mathcal{O}\left(\sum_{j=1}^{i-1} t_{G_j}(q_j, l_j) + \sum_{j=i+1}^{m} t_{(G_j,B_j)}(q_j)\right)\right]$ and $q'_i := \varphi \cdot q_i$ for $i = 1, \ldots, m-1$, whereas $t_m := l_{<m} + t + \mathcal{O}\left(\sum_{j=1}^{m-1} t_{G_j}(q_j, l_j)\right)$ and $q'_m := q_m$.

The asymmetry of our proof technique allows $(\mathbf{G}_m, B_m)$ to be fully stateful.[22] Furthermore, both t'_m and q'_m are much smaller then the corresponding terms t'_i and q'_i for $i = 1, \ldots, m-1$. The following paragraph provides a proof sketch for the case $m = 2$. The full proof is deferred to Section 3.2.

PROOF IDEA FOR $m = 2$. The proof follows similar lines as Levin's proof of the XOR-lemma [22, 11], but with some major differences due to the peculiarities of reactive systems. For simplicity, we let $(\mathbf{G}_1, B_1) = (\mathbf{F}, B)$ and $(\mathbf{G}_2, B_2) = (\mathbf{G}, C)$. Let $\mathbf{A}$ be an adversary with $\Gamma^{\mathbf{A}}(\mathbf{F} \| \mathbf{G}, B \oplus C) >$

[21] i.e. the maximal size of the state after the given number of queries

[22] An incomparable generalization of the XOR-lemma for stateful interactive systems was proposed by Halevi and Rabin [13]. However, it relies on *sequential* (rather than parallel) access to the systems $\mathbf{G}_1, \ldots, \mathbf{G}_m$, which is not sufficient for the applications of this paper.

$\delta \cdot \epsilon + \gamma$. We show that either there exists an adversary $\mathbf{A}'$ such that $\Gamma^{\mathbf{A}'}(\mathbf{F}, B) > \delta$ or there exists an adversary $\mathbf{A}''$ such that $\Gamma^{\mathbf{A}''}(\mathbf{G}, C) > \epsilon$, contradicting the assumed hardness of $(\mathbf{F}, B)$ and/or $(\mathbf{G}, C)$. The time complexities of $\mathbf{A}'$ and $\mathbf{A}''$ are strictly related to that of $\mathbf{A}$. Recall that the pair $(\mathbf{F}, B) = (\mathbf{F}(S), B(S))$ is cc-stateless, and for all values s taken by the random variable S we define

$$\alpha_1(s) := \Gamma^{\mathbf{A}}(\mathbf{F}(s)\|\mathbf{G}, 1 \oplus C) \text{ and } \alpha(s) := \Gamma^{\mathbf{A}}(\mathbf{F}(s)\|\mathbf{G}, B(s) \oplus C).$$

By definition, $\mathsf{E}[\alpha(S)] > \delta \cdot \epsilon + \gamma$. Moreover $\alpha(s) = \alpha_1(s)$ if $B(s) = 1$, and $\alpha(s) = -\alpha_1(s)$ otherwise. This implies that $\alpha_1(S)$ has good correlation with $B(S)$, as an adversary $\mathbf{A}'$ outputting 1 with probability $\frac{1}{2} + \frac{\alpha_1(s)}{2}$ (when given access to $\mathbf{F}(s)$) has advantage at least $\delta \cdot \epsilon + \gamma$. If $|\alpha_1(s)| = |\alpha(s)| \leq \epsilon$ holds for all s, then the advantage can be amplified to be larger than δ by outputting 1 with probability $\frac{1}{2} + \frac{\alpha_1(s)}{2\epsilon}$. Of course, $\mathbf{A}'$ does not know $\alpha_1(s)$, but a statistical estimate can be obtained by repeated interaction with $\mathbf{F}(s)$, as it is stateless: The term γ compensates the possible estimation error.

Note that the existence of a single value s with the property that $|\alpha_1(s)| > \epsilon$ implies that there exists a bit b such that the adversary $\mathbf{A}'' := \mathbf{A}(\mathbf{F}(s)\|\cdot) \oplus b$ has advantage larger than ϵ in guessing C from $\mathbf{G}$, i.e., $\mathbf{A}''$ is the adversary that simulates the execution of $\mathbf{A}$ with the parallel composition of $\mathbf{F}(s)$ and the given system $\mathbf{G}$, and outputs $\mathbf{A}$'s output XORed with b. But such an adversary $\mathbf{A}''$ is not necessarily efficient because an efficient implementation of $\mathbf{F}(s)$ may not exist. To overcome this problem, we show that for the above adversary $\mathbf{A}'$ to succeed, it is sufficient that the probability over the choice of S that $|\alpha_1(S)| > \epsilon + \gamma/4$ is smaller than $\gamma/4$. Furthermore, if this probability is at least $\gamma/4$, a probabilistic argument yields a (sufficiently) small state σ for the (efficient) implementation F of $\mathbf{F}$ and a (fixed) bit b such that the *efficient* adversary $\mathbf{A}'' := \mathbf{A}(F[\sigma]\|\cdot) \oplus b$ achieves advantage at least ϵ.

3.2 Proof of Theorem 5

The Isolation Lemma. The proof of Theorem 5 relies on the so-called *isolation lemma*, which was previously used in a number of hardness amplification results, including Levin's proof of the XOR-lemma [22, 11], Myers' indistinguishability amplification result for PRFs [29], and hardness amplification of weakly-verifiable puzzles [3]. The isolation lemma reduces the situation where an adversary obtains an advantage $\epsilon \cdot \delta$ in guessing the XOR of the bits $B_i \oplus \cdots \oplus B_m$ for $m - i$ system-bit pairs to either an adversary guessing B_i from $\mathbf{G}_i$ with advantage at least δ, or to the situation where an adversary guesses $B_{i+1} \oplus \cdots \oplus B_m$ for the system-bit pairs $(\mathbf{G}_{i+1}, B_{i+1}), \ldots, (\mathbf{G}_m, B_m)$.

In contrast to previous isolation lemmas in the literature, we give a more concrete statement which allows to considerably simplify the concrete analysis of the resulting adversaries in the reduction; this is a key step in order to prove some of the later results of this paper.

For the remainder of this section, we fix interactive algorithms $G_1, \ldots, G_m$ implementing the systems $\mathbf{G}_1, \ldots, \mathbf{G}_m$. Also, for a given parameter γ we define, for an understood parameter $\overline{\gamma} > 0$,

$$\varphi = 2\left(\frac{24}{\overline{\gamma}}\right)^2 \cdot \ln\left(\frac{7}{\overline{\gamma}}\right).$$

Note that the value φ used above corresponds to the special case where $\overline{\gamma} := \gamma/m$ for the understood parameter γ. The values $l_1, l_2, \ldots$ are defined as above (before the statement of Teorem 5) with respect to φ.

Lemma 6 (Isolation Lemma)**.** *Let $i \in \{1, \ldots, m-1\}$, let $\overline{\gamma} > 0$, and let $\mathbf{A}$ be an adversary with complexity t making q_j queries to $\mathbf{G}_j$ for $j = 1, \ldots, m$. Moreover, let $\sigma_1, \ldots, \sigma_{i-1}$ be valid states for $G_1, \ldots, G_{i-1}$, respectively, with $|\sigma_j| \leq l_j$ (for $j = 1, \ldots, i-1$), and let $b_{[1,i-1]} \in \{0,1\}$ be such that*

$$\Gamma^{\mathbf{A}}(G_1[\sigma_1] \| \cdots \| G_{i-1}[\sigma_{i-1}] \| \mathbf{G}_i \| \cdots \| \mathbf{G}_m, b_{[1,i-1]} \oplus B_i \oplus \cdots \oplus B_m) > \epsilon \cdot \delta + \overline{\gamma}.$$

Then, at least one of the following two statements is true:

(i) There exists a valid state σ_i for G_i with $|\sigma_i| \leq l_i$ and a bit $b_{[1,i]} \in \{0,1\}$ such that

$$\Gamma^{\mathbf{A}}(G_1[\sigma_1] \| \cdots \| G_i[\sigma_i] \| \mathbf{G}_{i+1} \| \cdots \| \mathbf{G}_m, b_{[1,i]} \oplus B_{i+1} \oplus \cdots \oplus B_m) > \epsilon;$$

(ii) There exists an adversary $\mathbf{A}_i'$ with running time t_i' making q_i' queries such that $\Gamma^{\mathbf{A}_i'}(\mathbf{G}_i, B_i) > \delta$, and

$$t_i' = l_{<i} + \varphi \cdot \left[t + \mathcal{O}\left(\sum_{j=1}^{i-1} t_{G_j}(q_j, l_j) + \sum_{j=i+1}^{m} t_{(G_j, B_j)}(q_j) \right) \right]$$

and $q_i' = \varphi \cdot q_i$.

The proof of the isolation lemma is postponed to Section 3.3. The next paragraph shows how the full XOR-lemma can be obtained from the isolation lemma. Note that we do not put priority on optimizing values (such as the function φ defined above), but rather on having simpler expressions, which are in particular independent of the upper bounds on the underlying advantages.

From the Isolation Lemma to the XOR-Lemma. In the following, fix some $t, q_1, \ldots, q_m, \gamma > 0$ and set $\overline{\gamma} := \gamma/(m-1)$. Define t_i', q_i' as in the statement of Theorem 5, and let $\delta_i := \Gamma_{t_i', q_i'}(\mathbf{G}_i, B_i)$.

From now on, let $\mathbf{A}$ be the adversary with running time t making $q_1, \ldots, q_m$ queries to the respective systems, and such that

$$\Gamma^{\mathbf{A}}(\mathbf{G}_1 \| \cdots \| \mathbf{G}_m, B_1 \oplus \cdots \oplus B_m) \geq \delta_1 \cdots \delta_m + \gamma.$$

Define $\gamma_i := (m-1-i) \cdot \overline{\gamma}$ for $i = 0, \ldots, m-1$ and consider the statements $\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, b)$ for a bit $b \in \{0,1\}$ and valid states $\sigma_1, \ldots, \sigma_i$ of $G_1, \ldots, G_i$ respectively, with $|\sigma_j| \leq l_j$ for $j = 1, \ldots, i$ which holds if and only if

$$\begin{aligned} \Gamma^{\mathbf{A}}(G_1[\sigma_1] \| \cdots \| G_i[\sigma_i] \| \mathbf{G}_{i+1} \| \cdots \| \mathbf{G}_m, b \oplus B_{i+1} \oplus \cdots \oplus B_m) &\geq \delta_{i+1} \cdots \delta_m + \gamma_i \\ &\geq \delta_{i+1} \cdot (\delta_{i+2} \cdots \delta_m + \gamma_{i+1}) + \overline{\gamma}, \end{aligned}$$

where we have used the facts that $\gamma_i := \gamma_{i-1} + \overline{\gamma}$ and $\delta_{i+1} \leq 1$. In particular, by our assumption $\mathsf{STAT}_0(0)$ holds. Furthermore, given $\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_{i-1}, b)$ holds for some i we apply the Isolation Lemma (Lemma 6). Note that condition (ii) cannot hold, as otherwise there exists an adversary $\mathbf{A}_i'$ such that $\Gamma^{\mathbf{A}_i'}(\mathbf{G}_i, B_i) > \delta_i$, but this contradicts the assumed hardness of $(\mathbf{G}_i, B_i)$, as we have defined δ_i with respect to the maximal running time of a constructed adversary $\mathbf{A}_i'$. Therefore, condition (i) must hold: In other words, we have shown that for all $i = 0, \ldots, m-2$

$$\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, b) \Longrightarrow \exists \sigma_{i+1}, b' : \mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, \sigma_{i+1}, b \oplus b'),$$

where all the states $\sigma_1, \sigma_2, \ldots$ are valid for their respective implementations and are such that $|\sigma_i| \leq l_i$. Iterating the argument we obtain that there exist $\sigma_1, \ldots, \sigma_{m-1}$ and a bit b such that

$$\Gamma^{\mathbf{A}}(G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_i] \| \mathbf{G}_m, b \oplus B_m) > \delta_m + \widetilde{\gamma}_{m-1} = \delta_m.$$

However, we can now consider the adversary $\mathbf{A}' := \mathbf{A}(G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_i] \| \cdot) \oplus b$ which, given access to $\mathbf{G}_m$, simulates $\mathbf{A}$ interacting with $G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_i] \| \mathbf{G}_m$, obtaining output b', and finally outputs $b \oplus b'$. Such an adversary has advantage larger than $> \delta_m$, contradicting the assumed hardness of $(\mathbf{G}_m, B_m)$.

3.3 Proof of the Isolation Lemma

SETUP. Throughout the proof, let $(\mathbf{G}_i(\cdot), B_i(\cdot))$ and $S \in \mathcal{S}$ be such that $(\mathbf{G}_i(S), B_i(S)) \equiv (\mathbf{G}_i, B_i)$. Note that such a representation of $(\mathbf{G}_i, B_i)$ exists since the system-bit pair is assumed to be cc-stateless (as $i \leq m-1$). In particular, $B_i(s)$ depends (without loss of generality) deterministically on the input s. For the remainder of this proof, it is convenient to define $\overline{\gamma}' := \frac{\overline{\gamma}}{4}$, and for the system-bit pairs $(\mathbf{G}_{i+1}, B_{i+1}), \ldots, (\mathbf{G}_m, B_m)$ (and the given states $\sigma_1, \ldots, \sigma_{i-1}$ in the statement of the isolation lemma) we use the notation

$$\begin{aligned} B_{[i+1,m]} &:= B_{i+1} \oplus B_{i+2} \oplus \cdots \oplus B_m, \\ \mathbf{M}(\cdot) &:= G_1[\sigma_1] \| \cdots \| G_{i-1}[\sigma_{i-1}] \| \cdot \| \mathbf{G}_{i+1} \| \cdots \| \mathbf{G}_m. \end{aligned}$$

Note that these are only notational shorthands. In particular, the system $\mathbf{G}_j$ is correlated with the corresponding bit B_j for $j \geq i+1$ as in the system-bit pair $(\mathbf{G}_j, B_j)$. Moreover, we define $\alpha_1, \alpha : \mathcal{S} \to [-1, 1]$ such that for all $s \in \mathcal{S}$

$$\begin{aligned} \alpha_1(s) &:= 2 \cdot \mathsf{P}[\mathbf{A}(\mathbf{M}(\mathbf{G}_i(s))) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]} = 1] - 1, \\ \alpha(s) &:= 2 \cdot \mathsf{P}[\mathbf{A}(\mathbf{M}(\mathbf{G}_i(s))) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]} = B_i(s)] - 1. \end{aligned}$$

By rearranging terms in the definition of α (i.e. by moving $b_{[1,i-1]} \oplus B_{[i+1,m]}$ to the right hand side) we see that $\mathsf{E}[\alpha(S)] > \epsilon \cdot \delta + \overline{\gamma}$ by the assumption of the lemma, and also $\alpha(s) = \alpha_1(s)$ if $B_i(s) = 1$, and $\alpha(s) = -\alpha_1(s)$ if $B_i(s) = 0$.

The remainder of the proof of the isolation lemma is subdivided into the following two technical lemmas.

Lemma 7. *If $\mathsf{P}[|\alpha_1(S)| > \epsilon + \overline{\gamma}'] > \overline{\gamma}'$, then there exists σ_i for G_i with $|\sigma_i| \leq l_i$ and $b_i \in \{0,1\}$ such that*

$$\Gamma^{\mathbf{A}}(\mathbf{M}(G_i[\sigma_i]), b_{[1,i-1]} \oplus b_i \oplus B_{[i+1,m]}) > \epsilon.$$

Lemma 8. *If $\mathsf{P}[|\alpha_1(S)| > \epsilon + \overline{\gamma}'] \leq \overline{\gamma}'$, then there exists an adversary $\mathbf{A}_i'$ with running time t_i' and making q_i' queries such that $\Gamma^{\mathbf{A}_i'}(\mathbf{G}_i, B_i) > \delta$.*

The isolation lemma is implied by these two lemmas, since either $\mathsf{P}[|\alpha_1(S)| > \epsilon + \overline{\gamma}'] > \overline{\gamma}'$ holds (and in this case Lemma 7 implies statement (i)), or $\mathsf{P}[|\alpha_1(S)| > \epsilon + \overline{\gamma}'] \leq \overline{\gamma}'$ holds, which yields statement (ii).

The remainder of this section is devoted to proving the above two lemmas.

Process SAMPLE:

$b := 0$;
$\sigma_0 :=$ initial state of G_i;
for $j := 1$ to φ **do**
  $o_j := \mathbf{A}(\mathbf{M}(G_i[\sigma_{j-1}])) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]}$;
  $\sigma_j :=$ final state of G_i;
$\overline{\alpha_1} := 2 \cdot \frac{1}{\varphi} \cdot \sum_{j=1}^{\varphi} o_j - 1$;
if $\overline{\alpha_1} > \epsilon + (2/3)\overline{\gamma}'$ **then**
  $b := 1$
else if $\overline{\alpha_1} < -\epsilon - (2/3)\overline{\gamma}'$ **then**
  $b := 0$;
output (σ, b).

Process SAMPLE$'$:

$b := 0$
$s \leftarrow \mathsf{P}_{S_i}$;
for $j := 1$ to φ **do**
  $o_j := \mathbf{A}(\mathbf{M}(\mathbf{G}_i(s))) \oplus b_{[1,i-1]} \oplus B^{(j)}_{[i+1,m]}$;
$\overline{\alpha_1} := 2 \cdot \frac{1}{\varphi} \cdot \sum_{j=1}^{\varphi} o_j - 1$;
if $\overline{\alpha_1} > \epsilon + (2/3)\overline{\gamma}'$ **then**
  $b := 1$
else if $\overline{\alpha_1} < -\epsilon - (2/3)\overline{\gamma}'$ **then**
  $b := 0$;
output (s, b).

Figure 1: Sampling processes SAMPLE and SAMPLE$'$ used in the proof of Lemma 7.

Proof of Lemma 7. We define a random process SAMPLE (which is depicted on the left-hand-side of Figure 1) outputting a pair (σ, b), where σ is a valid state for G_i and $b \in \{0,1\}$ is a bit. The process SAMPLE lets φ independent instances of $\overline{\mathbf{A}}(\cdot) := \mathbf{A}(\mathbf{M}(\cdot)) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]}$[23] sequentially interact with the *same* instance of $\mathbf{G}_i$, producing outputs $o_1, \ldots, o_\varphi$ (and denote by $\overline{o}$ their average). The instance of $\mathbf{G}_i$ is simulated using G_i by letting, for $j = 1, \ldots, \varphi$, the j'th independent instance of $\overline{\mathbf{A}}$ interact with $G_i[\sigma_{j-1}]$, and setting σ_j to be the final state of G_i after such interaction (and $\sigma_0 := \perp$ is the empty, initial state for G_i). Also, it sets $\sigma := \sigma_\varphi$. The process SAMPLE computes $\overline{\alpha_1} := 2 \cdot \overline{o} - 1$. If $\overline{\alpha}_1$ is higher than ϵ, we expect $\Gamma^{\mathbf{A}}(\mathbf{M}(G_i[\sigma]), b_{[1,i-1]} \oplus 1 \oplus B_{[i+1,m]}) > \epsilon$ (and the process hence outputs $(\sigma, 1)$), whereas if it is lower than $-\epsilon$, we expect $\Gamma^{\mathbf{A}}(\mathbf{M}(G_i[\sigma]), b_{[1,i-1]} \oplus 0 \oplus B_{[i+1,m]}) > \epsilon$ (and SAMPLE outputs $(\sigma, 0)$).

We consider a random experiment where SAMPLE samples a pair (σ, b), and we subsequently compute $\mathbf{A}(\mathbf{M}(G_i[\sigma])) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]}$. We define the event $\mathcal{T}$ that one of the **if**-statements within process SAMPLE is executed, and we show that under the assumption that $\mathsf{P}[|\alpha_1(S)| > \epsilon + \overline{\gamma}'] > \overline{\gamma}'$ we have

$$\pi_1 := \mathsf{P}[(\sigma, b) \leftarrow \mathsf{SAMPLE} : \mathbf{A}(\mathbf{M}(G_i[\sigma])) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]} = b \,|\, \mathcal{T}] \geq \frac{1+\epsilon}{2}, \tag{1}$$

as well as $\mathsf{P}[\mathcal{T}] > 0$. Note that this is sufficient to obtain the statement of Lemma 7, as we can choose a pair (σ, b) output by SAMPLE conditioned on the event $\mathcal{T}$ (note that this yields a well-defined distribution of pairs (σ, b)) maximizing the above probability, and set $\sigma_i := \sigma$, $b_{[1,i]} := b_{[1,i-1]} \oplus b$. Since $\mathbf{A}$ issues at most q_i queries, and the above process is repeated φ times, we have $|\sigma_i| \leq s_{G_i}(q_i \cdot \varphi) = l_i$.

In order to prove inequality (1), we consider a second sampling process (called SAMPLE$'$ and depicted on the right-hand side of Figure 1) which samples s according to P_S and subsequently computes the values $o_1, \ldots, o_\varphi$ by letting φ independent instances of $\overline{\mathbf{A}}$ interact with $\mathbf{G}_i(s)$. The process finally outputs the pair (s, b) with b being computed as in SAMPLE. In particular, we also denote here as $\mathcal{T}$ the event that one of the two conditions in the **if**-statement is satisfied, and we define

$$\pi_2 := \mathsf{P}[(s, b) \leftarrow \mathsf{SAMPLE}' : \mathbf{A}(\mathbf{M}(\mathbf{G}_i(s))) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]} = b \,|\, \mathcal{T}].$$

[23] i.e. the adversary which for new independent instances of $(\mathbf{G}_{i+1}, B_{i+1}), \ldots, (\mathbf{G}_m, B_m)$ (with $B_{[i+1,m]} := B_{i+1} \oplus \cdots \oplus B_m$) simulates an interaction of $\mathbf{A}$ with $\mathbf{G}_1[\sigma_1] \| \cdots \| \mathbf{G}_{i-1}[\sigma_{i-1}] \| \mathbf{S} \| \mathbf{G}_{i+1} \| \cdots \| \mathbf{G}_m$, where $\mathbf{S}$ is the given system, and adds $b_{[1,i-1]} \oplus B_{[i+1,m]}$ to its output to obtain the actual output

A crucial observation is that the sampling processes SAMPLE and SAMPLE' compute the values $o_1, \ldots, o_\varphi$ by letting multiple, *independent*, instances of $\overline{\mathbf{A}}$ interact with the *same* instance of $\mathbf{G}_i$, and furthermore, we are interested in the probabilities π_1 and π_2 that a further, independent, instance of $\overline{\mathbf{A}}$ is successful in guessing the bit b when interacting (once again) with the same instance of $\mathbf{G}_i$. In the first random experiment, the instance of $\mathbf{G}_i$ is simulated by using the algorithm G_i. In the second case, this is done by choosing a random s and running $\mathbf{G}_i(s)$. However, in both cases, all probabilities (including the event $\mathcal{T}$) only depend on the input-output behavior of $\mathbf{G}_i$, and not on the way this instance is simulated, and hence we have $\pi_1 = \pi_2$, and $\mathcal{T}$ occurs with the same probability in both experiments.[24]

In the remainder of the proof we thus show that $\pi_2 \geq \frac{1+\epsilon}{2}$, which is much simpler than working with the original random experiment, as for each value s the system $\mathbf{G}_i(s)$ is stateless and we can conveniently upper-bound the error in the estimate of $\overline{\alpha_1}$ (using e.g. Hoeffding's inequality) due to the fact that the random variables $o_1, \ldots, o_\varphi$, conditioned on some fixed value s, are statistically independent. More formally, we denote the values $\overline{\alpha_1}$ and o_j associated with a particular choice of s as $\overline{\alpha_1}(s)$ and $o_j(s)$, respectively. We additionally define (for every $\eta \geq 0$) the set

$$\mathcal{G}_\eta := \big\{ s \in \mathcal{S} \,\big|\, |\alpha_1(s)| > \epsilon + \eta \big\}.$$

Note that $\mathcal{G}_\eta \subseteq \mathcal{G}_{\eta'}$ for all $\eta > \eta'$ and $\mathsf{P}[s \in \mathcal{G}_{\overline{\gamma}'}] > \overline{\gamma}'$ by the above assumption. Furthermore, with a slight abuse of notation, we denote by $\mathcal{G}_\eta$ the event $s \in \mathcal{G}_\eta$. In the following, we show that $\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \,|\, \mathcal{T}]$ is overwhelming.

Note that for a fixed s (because of the fact that $\mathbf{G}_i(s)$ is stateless) the random variables $o_j(s)$ (which here we exceptionally denote by lower-case letters) are independent binary variables, with

$$\mathsf{E}[o_j(s)] = p_1(s) = \mathsf{P}[\mathbf{A}(\mathbf{M}(\mathbf{G}_i(s))) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]} = 1]$$

for all $j = 1, \ldots, \varphi$. First, by Hoeffding's bound, and because of the factor two in the definition of α_1, we have for every fixed $s \in \mathcal{S}$,

$$\mathsf{P}\big[|\alpha_1(s) - \overline{\alpha_1}(s)| > \overline{\gamma}'/3\big] \leq \mathsf{P}\left[\left|\frac{1}{\varphi}\sum_{i=1}^{\varphi} o_i(s) - p_1(s)\right| > \overline{\gamma}'/6\right] \leq 2 \cdot e^{-\varphi(\overline{\gamma}'/6)^2}.$$

This in particular implies that for the randomly chosen s,

$$\mathsf{P}\big[\overline{\mathcal{G}_{\overline{\gamma}'/3}} \wedge \mathcal{T}\big] \leq \mathsf{P}\big[\mathcal{T} \,|\, \overline{\mathcal{G}_{\overline{\gamma}'/3}}\big] \leq \mathsf{P}\big[|\alpha_1(s) - \overline{\alpha_1}(s)| > \overline{\gamma}'/3 \,|\, s \notin \mathcal{G}_{\overline{\gamma}'/3}\big] \leq 2 \cdot e^{-\varphi(\overline{\gamma}'/6)^2}.$$

This yields

$$\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \,|\, \mathcal{T}] = \frac{\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}]}{\mathsf{P}\big[\overline{\mathcal{G}_{\overline{\gamma}'/3}} \wedge \mathcal{T}\big] + \mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}]} \geq \frac{\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}]}{\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}] + 2 \cdot e^{-\varphi(\overline{\gamma}'/6)^2}}. \tag{2}$$

As the function $x \mapsto \frac{x}{c+x}$ is non-decreasing over $[0,1]$ for every constant $c \in [0,1]$, it is now sufficient to find a non-trivial lower bound for $\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}]$: Since $\mathcal{G}_{\overline{\gamma}'} \subseteq \mathcal{G}_{\overline{\gamma}'/3}$, we have

$$\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}] \geq \mathsf{P}[\mathcal{G}_{\overline{\gamma}'} \wedge \mathcal{T}] = \mathsf{P}[\mathcal{G}_{\overline{\gamma}'}] \cdot \mathsf{P}[\mathcal{T} \,|\, \mathcal{G}_{\overline{\gamma}'}] \geq \overline{\gamma}' \cdot (1 - 2 \cdot e^{-\varphi(\overline{\gamma}'/6)^2}) \geq \overline{\gamma}' - 2 \cdot e^{-\varphi(\overline{\gamma}'/6)^2},$$

[24] Informally, this means that if we simulate a system $\mathbf{G} = \mathbf{G}(S)$ which is used only through black-box access, and we do this by means of an algorithm G, we *implicitly* sample a corresponding S during the interaction even though such S does not actually exist.

Adversary $\mathbf{A}'_i$: // given access to $\mathbf{G}_i$

for $j := 1$ to φ **do**

$o_j := \mathbf{A}(\mathbf{M}(\mathbf{G}_i)) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]}$;

$\overline{\alpha_1} := 2 \cdot \frac{1}{\varphi} \cdot \sum_{j=1}^{\varphi} o_j - 1$;

output 1 with probability $\min\{\max\{\frac{1}{2} + \frac{\overline{\alpha}_1}{2(\delta+\overline{\gamma}')}, 0\}, 1\}$.

Figure 2: Adversary $\mathbf{A}'_i$ in the proof of Lemma 8.

since $\mathsf{P}[\overline{\mathcal{T}} \,|\, \mathcal{G}_{\overline{\gamma}'}] \leq \mathsf{P}[|\alpha_1(S) - \overline{\alpha_1}(S)| > \overline{\gamma}'/3 \,|\, s \in \mathcal{G}_{\overline{\gamma}'}]$. (Note that the above inequality implies in particular $\mathsf{P}[\mathcal{T}] > 0$.) Plugging this into (2) yields $\mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \,|\, \mathcal{T}] \geq 1 - \frac{2}{\overline{\gamma}'} \cdot e^{-\varphi(\overline{\gamma}'/6)^2}$, and thus we conclude

$$\begin{aligned}
\pi_1 = \pi_2 &\geq \mathsf{P}[\mathcal{G}_{\overline{\gamma}'/3} \,|\, \mathcal{T}] \cdot \mathsf{P}[\mathbf{A}(\mathbf{M}(\mathbf{G}_i(S))) = b_{[1,i-1]} \oplus b \oplus B_{[i+1,m]} \,|\, S \in \mathcal{G}_{\overline{\gamma}'/3} \wedge \mathcal{T}] \\
&\geq \left(1 - \frac{2}{\overline{\gamma}'} \cdot e^{-\varphi(\overline{\gamma}'/6)^2}\right) \cdot \left(\frac{1 + \epsilon + \overline{\gamma}'/3}{2}\right) \\
&\geq \frac{1+\epsilon}{2} + \frac{\overline{\gamma}'}{6} - \frac{2}{\overline{\gamma}'} \cdot e^{-\varphi(\overline{\gamma}'/6)^2} > \frac{1+\epsilon}{2},
\end{aligned}$$

by the definition of φ.

Proof of Lemma 8. We construct the adversary $\mathbf{A}'_i$ for predicting the bit B_i given access to $\mathbf{G}_i$ as described in Figure 2: It estimates the value α_1 by repeatedly simulating the execution of $\mathbf{A}(\mathbf{M}(\cdot)) \oplus b_{[1,i-1]} \oplus B_{[i+1,m]}$ with the system $\mathbf{G}_i$, and finally outputs 1 with probability $\frac{1}{2} + \frac{\overline{\alpha_1}}{2(\delta+\overline{\gamma}')}$

As above, we assume that $(\mathbf{G}_i, B_i)$ is instantiated by first sampling $s \in \mathcal{S}$ according to P_S and then behaving as $(\mathbf{G}_i(s), B_i(s))$. This allows us to analyze the behavior of $\mathbf{A}'_i$ conditioned on each value $s \in \mathcal{S}$. In particular, we denote by $\overline{\alpha_1}(s)$ the value obtained when run on $\mathbf{G}(s)$.

We first consider the event $\mathcal{E}$ that for the chosen $S = s$ we have $\left|\overline{\alpha_1(s)} - \alpha_1(s)\right| > \overline{\gamma}'$. Note that the individual runs of the independent instances give independent outputs because $\mathbf{G}_i(s)$ is stateless. Hence, by Hoeffding's bound $\mathsf{P}[\mathcal{E}] \leq 2 \cdot e^{-\varphi(\overline{\gamma}/8)^2}$ (as similar argument was given above). Also as above, we let $\mathcal{G}_{\overline{\gamma}'}$ the set of values $s \in \mathcal{S}$ such that $|\alpha_1(s)| \leq \epsilon + \overline{\gamma}'$. (In particular, $\mathsf{P}[\overline{\mathcal{G}_{\overline{\gamma}}}] \leq \overline{\gamma}'$, by the assumption of the lemma.)

For all $s \in \mathcal{G}_{\overline{\gamma}'}$, we have

$$\mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(s)) = B_i(s)] \geq \mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(s)) = B_i(s) \wedge \overline{\mathcal{E}}] = \mathsf{P}[\overline{\mathcal{E}}] \cdot \mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(s)) = B_i(s) \,|\, \overline{\mathcal{E}}] \tag{3}$$

as well as

$$\mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(s)) = B_i(s) \,|\, \overline{\mathcal{E}}] \geq \frac{1}{2} + \frac{\alpha(s) - \overline{\gamma}'}{2(\epsilon + \overline{\gamma}')}. \tag{4}$$

Moreover,

$$\sum_{s \in \mathcal{G}_{\overline{\gamma}'}} \mathsf{P}_S(s) \cdot \alpha(s) = \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \alpha(s) - \sum_{s \notin \mathcal{G}_{\overline{\gamma}'}} \mathsf{P}_S(s) \cdot \alpha(s) \geq \mathsf{E}[\alpha(s)] - \mathsf{P}[\overline{\mathcal{G}_{\overline{\gamma}'}}] \geq \epsilon \cdot \delta + \overline{\gamma} - \overline{\gamma}', \tag{5}$$

which finally yields

$$\begin{aligned}
\mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(S)) = B_i(S) \wedge S \in \mathcal{G} \,|\, \overline{\mathcal{E}}] &= \sum_{s\in\mathcal{G}} \mathsf{P}_S(s) \cdot \mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i(s)) = B_i(s) \,|\, \overline{\mathcal{E}}] \\
&\overset{(4)}{\geq} \frac{1}{2} + \frac{\sum_{s\in\mathcal{G}} \mathsf{P}_S(s)\cdot\alpha(s) - \overline{\gamma}'}{2(\epsilon+\overline{\gamma}')} \\
&\overset{(5)}{\geq} \frac{1}{2} + \frac{\epsilon\cdot\delta + \overline{\gamma} - 2\overline{\gamma}'}{2(\epsilon+\overline{\gamma}')} \\
&= \frac{1}{2} + \frac{\delta\cdot(\epsilon+\overline{\gamma}') + \overline{\gamma} - \epsilon\cdot\overline{\gamma}' - 2\overline{\gamma}'}{2(\epsilon+\overline{\gamma}')} \\
&\geq \frac{1}{2} + \frac{\delta}{2} + \frac{\overline{\gamma}'}{2(\epsilon+\overline{\gamma}')} \geq \frac{1}{2} + \frac{\delta}{2} + \frac{\overline{\gamma}'}{4}.
\end{aligned}$$

Hence, using Equation 3 we conclude that

$$\mathsf{P}[\mathbf{A}'_i(\mathbf{G}_i) = B_i] \geq \frac{1+\delta}{2} + \frac{\overline{\gamma}'}{4} - 2e^{-\varphi(\overline{\gamma}/8)^2} = \frac{1+\delta}{2}$$

for the given value of φ.

As for the complexity of the adversary, we note that the adversary $\mathbf{A}$' needs to hard-code a description of the states $\sigma_1, \ldots, \sigma_{i-1}$ (needing at most $l_{<i}$ bits, which are hence counted only once in the time-complexity) and runs φ copies of $\mathbf{A}$, making each time q_i queries to $\mathbf{G}_i$, and we additionally need to simulate $G_j[\sigma_j]$ for all $j = 1, \ldots, i-1$ (which takes time $t_{G_j}(q_j, l_j)$), and need to simulate new instances of $(\mathbf{G}_j, B_j)$ for $j = i+1, \ldots, m$ (which takes time $t_{(G_j,B_j)}(q_j)$).

3.4 Obtaining a Uniform Reduction

The results of this paper are presented in the concrete (non-asymptotic) setting. The result can be translated directly in the asymptotic setting by letting all quantities be functions of the considered security parameter, but our proof is inherently non-uniform. However, it can easily be extended to the uniform setting using standard techniques (cf. e.g. [11]). We only discuss the main modifications. Note that the asymptotic setting is mostly used in the case $(\mathbf{G}_1, B_1) = \cdots = (\mathbf{G}_m, B_m) = (\mathbf{G}, B)$, especially because the number of instances m grows with the security parameter, and we restrict ourselves to this case in the following discussion.

Given black box access to the adversary $\mathbf{A}$, the new adversary $\mathbf{S}^{\mathbf{A}}$ for predicting B given access to $\mathbf{G}$ proceeds as follows:

(1) Initially, it sets $b_{[1,0]} := 0$.

(2) For each $i = 1, \ldots, m-1$, given states $\sigma_1, \ldots, \sigma_{i-1}$, it repeatedly and independently runs process SAMPLE, without the final **output** statement. We consider two cases:

   (a) If the **if**-statement is satisfied in some run of SAMPLE (which would have output (σ, b)), then we set $\sigma_i := \sigma$ and $b_{[1,i]} := b \oplus b_{[1,i-1]}$, and move to the next index $i+1$.

   (b) If after a number $\mathrm{it}_{\max} := -\kappa/\log(1-\lambda'(\kappa))$ (where κ is the security parameter) of runs with the same index i of SAMPLE the **if**-statement was never satisfied, we run the adversary $\mathbf{A}'_i$ with the found states $\sigma_1, \ldots, \sigma_{i-1}$ and $b_{[1,i-1]}$ as above against the given system $\mathbf{G}$ to guess B.

(3) If the for loop is terminated (i.e. Case (b) was never met, and we end up with $i = m$), then states $\sigma_1, \ldots, \sigma_{m-1}$ and a bit $b_{[1,m-1]}$ have been found. We use them to run the adversary $\mathbf{A}(G[\sigma_1]\| \cdots \|G[\sigma_{m-1}]\|\cdot)$ on $\mathbf{G}$ and add $b_{[1,m-1]}$ to the resulting output.

The main idea is that for given i, in the case where $|\alpha_1(S)| > \epsilon + \overline{\gamma}'$ holds with probability at least $\overline{\gamma}'$, then a good state σ_i is found with overwhelming probability within the given number of iterations $\mathrm{it}_{\max}$, and we can move to the next i. If for some i the **if**-statement is never satisfied, then with overwhelming probability $|\alpha_1(S)| \leq \epsilon + \overline{\gamma}'$, and thus the attacker $\mathbf{A}'_i$ is successful.

4 A General Product Theorem for Neutralizing Constructions from the XOR-Lemma

4.1 Setting and Theorem Statement

Throughout this section, let $\mathbf{C}(\cdot)$ be a neutralizing construction for the real systems $\mathbf{F}_1, \ldots, \mathbf{F}_m$ and the corresponding ideal $\mathbf{I}_1, \ldots, \mathbf{I}_m$, of which all but $\mathbf{F}_m$ and $\mathbf{I}_m$ have to be cc-stateless. We prove a general product theorem upper bounding $\Delta_{t,q}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m))$ in terms of the individual advantages $\Delta_{t'_i,q'_i}(\mathbf{F}_i, \mathbf{I}_i)$ (for some related t'_i, q'_i). The theorem is a computational version of the information-theoretic product theorem from [26]: In particular, we inherit the same bounds, with an unavoidable additive term.

The theorem relies on the canonical implementation $\langle F_i, I_i\rangle_{B_i}$ of $\langle \mathbf{F}_i, \mathbf{I}_i\rangle_{B_i}$ which chooses a random bit $B_i \in \{0,1\}$ and answers each query using the implementations F_i and I_i (with respective complexities t_{F_i} and t_{I_i}) of $\mathbf{F}_i$ or of $\mathbf{I}_i$, respectively, depending on the value of B_i. (B_i is in particular part of the state.) It can be implemented with complexity $t_{\langle F_i,B_i\rangle_{B_i}}(q,s) = \max\{t_{F_i}(q,s), t_{I_i}(q,s)\} + \mathcal{O}(1)$. This also yields an implementation of $(\langle \mathbf{F}_i, \mathbf{I}_i\rangle_{B_i}, B_i)$ with the same complexity (by additionally outputting the bit B_i). Finally, we let l_i and $l_{<i}$ as above be defined with respect to $\langle F_i, I_i\rangle_{B_i}$, and let t_C be the time complexity of an efficient implementation of $\mathbf{C}(\cdot)$.

Theorem 9 (Product Theorem)**.** *Let $\mathbf{C}(\cdot)$ be as above, and let $q > 0$ be such that $\mathbf{C}(\cdot)$ makes q_i queries to its i-th subsystem when invoked q times. Then, for all $t, \gamma > 0$, if $\Delta_{t'_i,q'_i}(\mathbf{F}_i, \mathbf{I}_i) \leq \frac{1}{2}$ for all $i = 1, \ldots, m-1$,*

$$\Delta_{t,q}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)) \leq 2^{m-1} \cdot \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\mathbf{F}_i, \mathbf{I}_i) + 2\gamma,$$

where $t'_i := l_{<i} + \varphi \cdot \left[t + t_C(q) + \mathcal{O}\big(\sum_{j=1}^{i-1} t_{\langle F_j,I_j\rangle_{B_j}}(q_j, l_j) + \sum_{j=i+1}^{m} t_{\langle F_j,I_j\rangle_{B_j}}(q_j)\big)\right]$ and $q'_i := \varphi \cdot q_i$ for all $i = 1, \ldots, m-1$, whereas $t'_m := l_{<m} + t + t_C(q) + \mathcal{O}\left(\sum_{j=1}^{m-1} t_{\langle F_j,I_j\rangle_{B_j}}(q_j, l_j)\right)$ and $q'_m := q_m$.

The remainder of this section provides a simple proof sketch of Theorem 9. A full proof is given in the next section (Section 4.2), whereas Section 4.3 presents some applications of the product theorem.

Proof Sketch. We present a proof sketch of the above theorem for the case $m = 2$. For simplicity, let $\mathbf{F}_1 = \mathbf{F}$, $\mathbf{F}_2 = \mathbf{G}$, $\mathbf{I}_1 = \mathbf{I}$, and $\mathbf{I}_2 = \mathbf{J}$. The core of the proof is a generic argument (i.e. it holds for all distinguishers, regardless of their computing power) reducing the task of upper bounding

the distinguishing advantage for a neutralizing construction to the setting of the XOR-lemma.[25] It is easy to verify that (also cf. [26])

$$\begin{aligned}\Delta^{\mathbf{D}}(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})) &= 2\cdot\Delta^{\mathbf{D}}(\langle\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_B,\mathbf{C}(\mathbf{I},\mathbf{J}))\\ &= 2\cdot\left|\Gamma^{\mathbf{D}}(\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_B,\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_{B'},B')\right|,\end{aligned}$$

where B and B' are independent uniformly distributed random bits. Note that conditioned on $B'=0$, the system $\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_B,\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_{B'}$ behaves as $\mathbf{C}(\mathbf{F},\mathbf{G})$ with probability $\frac{1}{2}$, and as $\mathbf{C}(\mathbf{I},\mathbf{J})$ otherwise. On the other hand, conditioned on $B'=1$ it always behaves as $\mathbf{C}(\mathbf{I},\mathbf{J})$. In particular, this implies that (for independent uniform random bits B_1,B_2)

$$\big(\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_B,\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_{B'},B'\big)\equiv\big(\mathbf{C}(\langle\mathbf{F},\mathbf{I}\rangle_{B_1},\langle\mathbf{G},\mathbf{J}\rangle_{B_2}),B_1\oplus B_2\big),$$

because of the neutralizing property. We thus obtain

$$\Gamma^{\mathbf{D}}(\langle\langle\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_B,\mathbf{C}(\mathbf{I},\mathbf{J})\rangle_{B'},B')=\Gamma^{\mathbf{D}}(\mathbf{C}(\langle\mathbf{F},\mathbf{I}\rangle_{B_1},\langle\mathbf{G},\mathbf{J}\rangle_{B_2}),B_1\oplus B_2)$$

and we conclude the proof by "absorbing" the computation of $\mathbf{C}(\cdot)$ into $\mathbf{D}$, clearly without modifying the advantage. Using the XOR-lemma (Theorem 5) for $m=2$ we obtain

$$\begin{aligned}\Delta_{t,q}(\mathbf{C}(\mathbf{F},\mathbf{G}),\mathbf{C}(\mathbf{I},\mathbf{J})) &\le 2\cdot\Gamma_{t+t_C(q),q_1,q_2}(\langle\mathbf{F},\mathbf{I}\rangle_{B_1}\|\langle\mathbf{G},\mathbf{J}\rangle_{B_2},B_1\oplus B_2)\\ &\le 2\cdot\Gamma_{t'_1,q'_1}(\langle\mathbf{F},\mathbf{I}\rangle_{B_1},B_1)\cdot\Gamma_{t'_2,q'_2}(\langle\mathbf{G},\mathbf{J}\rangle_{B_2},B_2)+2\gamma.\end{aligned}$$

for appropriate t'_1,q'_1 and t'_2,q'_2. It is not hard to extend this argument to constructions $\mathbf{C}(\cdot)$ which are "splittable", i.e., one can write $\mathbf{C}(\mathbf{F}_1,\ldots,\mathbf{F}_m)$ as $\mathbf{C}(\mathbf{F}_1,\mathbf{C}(\mathbf{F}_2,\ldots,\mathbf{F}_m))$. However, extending the proof to *arbitrary* neutralizing constructions for $m>2$ requires some extra care, as we explain in the next section. In particular, the above argument can be extended to show that $\Delta^{\mathbf{D}}(\mathbf{C}(\mathbf{F}_1,\ldots,\mathbf{F}_m),\mathbf{C}(\mathbf{I}_1,\ldots,\mathbf{I}_m))=2^{m-1}\cdot\left|\Gamma^{\mathbf{D}}(\langle\mathbf{F}_1,\mathbf{I}_1\rangle_{B_1}\|\ldots\|\langle\mathbf{F}_m,\mathbf{I}_m\rangle_{B_m},B_1\oplus\cdots\oplus B_m)\right|$, but this is not sufficient due to the extra term γ appearing in the upper bound given by the XOR-lemma.

4.2 Proof of the Product Theorem

The key technique used in the proof of Theorem 9 is the extension of the isolation lemma (Lemma 6) to the setting of computational indistinguishability amplification. While this follows the same lines as the proof sketch given above, we rely on the stronger statement given by the isolation lemma (Lemma 6) in order to prove the result for the most general setting where no structural requirements are made on the construction $\mathbf{C}(\cdot)$.

Isolation Lemma. In order to simplify notation, we denote as $G_1,G_2,\ldots$ the implementations $\langle F_1,I_1\rangle_{B_1},\langle F_2,I_2\rangle_{B_2},\ldots$ of the systems $\langle\mathbf{F}_1,\mathbf{I}_1\rangle_{B_1},\langle\mathbf{F}_2,\mathbf{I}_2\rangle_{B_2},\ldots$ for uniform random bits $B_1,B_2,\ldots$. It is also occasionally convenient to write $\mathbf{C}(\mathbf{F}_1\|\ldots\|\mathbf{F}_m)$ instead of $\mathbf{C}(\mathbf{F}_1,\ldots,\mathbf{F}_m)$, i.e., we see the construction as accessing one system (namely, the parallel composition of all subsystems), rather than the individual subsystems. Furthermore, we define the shorthands $\mathbf{F}_{[i,j]}:=\mathbf{F}_i\|\ldots\|\mathbf{F}_j$ and $\mathbf{I}_{[i,j]}:=\mathbf{I}_i\|\ldots\|\mathbf{I}_j$.

[25] A similar argument was implicitly used in the information-theoretic product theorem of [26].

Lemma 10. *Let $i \in \{1, \ldots, m-1\}$, let $\overline{\gamma} > 0$, and let $\mathbf{D}$ be a distinguisher with complexity t and making q queries to $\mathbf{C}(\cdot)$, resulting in $q_1, \ldots, q_m$ queries to the respective subsystems. Moreover, let $\sigma_1, \ldots, \sigma_{i-1}$ be valid states for the systems $\langle \mathbf{F}_1, \mathbf{I}_1 \rangle_{B_1}, \ldots, \langle \mathbf{F}_{i-1}, \mathbf{I}_{i-1} \rangle_{B_{i-1}}$, respectively, with $|\sigma_j| \leq l_j$ (for $j = 1, \ldots, i-1$), and let $b_{[1,i-1]} \in \{0,1\}$ be a binary value with the property that*

$$\Gamma^{\mathbf{D}}(\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_{i-1}[\sigma_{i-1}] \| \langle \mathbf{F}_{[i,m]}, \mathbf{I}_{[i,m]} \rangle_B), b_{[1,i-1]} \oplus B) > 2 \cdot \epsilon \cdot \delta + 2\overline{\gamma}.$$

Then, at least one of the following two statements is true:

(i) There exists a valid state σ_i for G_i with $|\sigma_i| \leq l_i$ and a bit $b_{[1,i]} \in \{0,1\}$ such that

$$\Gamma^{\mathbf{D}}(\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_i[\sigma_i] \| \langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_B), b_{[1,i]} \oplus B) > \epsilon;$$

(ii) There exists a distinguisher $\mathbf{D}'_i$ with running time t'_i making q'_i queries such that $\Delta^{\mathbf{D}'_i}(\mathbf{F}_i, \mathbf{I}_i) \geq \Gamma^{\mathbf{D}'_i}(\langle \mathbf{F}_i, \mathbf{I}_i \rangle_{B_i}, B_i) > \delta$, and

$$t'_i = l_{<i} + \varphi \cdot \left[t + \mathcal{O}\left(\sum_{j=1}^{i-1} t_{\langle F_j, I_j \rangle_{B_j}}(q_j, l_j) + \sum_{j=i+1}^{m} t_{\langle F_j, I_j \rangle_{B_j}}(q_j) \right) \right]$$

and $q'_i = \varphi' \cdot q_i$.

Proof. Using the shorthand $G = G_1[\sigma_1] \| \cdots \| G_{i-1}[\sigma_{i-1}]$, a straightforward calculation yields (for an additional uniform random bit B' independent of B)

$$\Gamma^{\mathbf{D}}(\mathbf{C}(G \| \langle \mathbf{F}_{[i,m]}, \mathbf{I}_{[i,m]} \rangle_B), b_{[1,i-1]} \oplus B) = 2 \cdot \Gamma^{\mathbf{D}}(\mathbf{C}(G \| \langle \langle \mathbf{F}_{[i,m]}, \mathbf{I}_{[i,m]} \rangle_{B'}, \mathbf{I}_{[i,m]} \rangle_B), b_{[1,i-1]} \oplus B). \quad (6)$$

Furthermore, we observe that the fact that $\mathbf{C}(\cdot)$ is neutralizing for $\mathbf{F}_1, \ldots, \mathbf{F}_m$ and $\mathbf{I}_1, \ldots, \mathbf{I}_m$ implies that $\mathbf{C}(\cdot)$ is neutralizing for $\langle \mathbf{F}_1, \mathbf{I}_1 \rangle_{B_1}, \ldots, \langle \mathbf{F}_{i-1}, \mathbf{I}_{i-1} \rangle_{B_{i-1}}, \mathbf{F}_i, \ldots, \mathbf{F}_m$ and $\mathbf{I}_1, \ldots, \mathbf{I}_m$. In turn, because $\langle \mathbf{F}_1, \mathbf{I}_1 \rangle_{B_1}, \ldots, \langle \mathbf{F}_{i-1}, \mathbf{I}_{i-1} \rangle_{B_{i-1}}$ are cc-stateless systems, the construction $\mathbf{C}(\cdot)$ is neutralizing for $G_1[\sigma_1], \ldots, G_{i-1}[\sigma_{i-1}], \mathbf{F}_i, \ldots, \mathbf{F}_m$ and $\mathbf{I}_1, \ldots, \mathbf{I}_m$ as well. This in particular implies that

$$\mathbf{C}(G \| \mathbf{F}_i \| \mathbf{I}_{[i+1,m]}) \equiv \mathbf{C}(G \| \mathbf{I}_i \| \mathbf{F}_{[i+1,m]}) \equiv \mathbf{C}(G \| \mathbf{I}_{[i,m]}) \equiv \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m),$$

and this yields the equivalence

$$\begin{aligned}(\mathbf{C}(G \| \langle \langle \mathbf{F}_{[i,m]}, \mathbf{I}_{[i,m]} \rangle_{B'}, \mathbf{I}_{[i,m]} \rangle_B), b_{[1,i-1]} \oplus B) \equiv \\ \equiv (\mathbf{C}(G \| \langle \mathbf{F}_i, \mathbf{I}_i \rangle_{B_i} \| \langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_{B''}), b_{[1,i-1]} \oplus B_i \oplus B'') \quad (7)\end{aligned}$$

for independent uniform random bits B_i and B'', since both system-bit pairs behave as $\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)$ whenever the bit is $1 - b_{[1,i-1]}$, whereas otherwise they behave as $\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)$ or $\mathbf{C}(G \| \mathbf{F}_{[i,m]})$ with probability $1/2$ each. Therefore, combining (6) and (7), the assumption of the lemma can equivalently be reformulated as

$$2 \cdot \Gamma^{\mathbf{D}}(\mathbf{C}(G \| \langle \mathbf{F}_i, \mathbf{I}_i \rangle_{B_i} \| \langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_{B''}), b_{[1,i-1]} \oplus B_i \oplus B'') > 2\epsilon\delta + 2\overline{\gamma},$$

which is equivalent to

$$\Gamma^{\mathbf{DC}}(G \| \langle \mathbf{F}_i, \mathbf{I}_i \rangle_{B_i} \| \langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_{B'}, B_i \oplus B') > \epsilon\delta + \overline{\gamma},$$

where $\mathbf{DC}$ is the distinguisher which, given a system $\mathbf{S}$, simulates the interaction between $\mathbf{D}$ and $\mathbf{C}(\mathbf{S})$, outputting $\mathbf{D}$'s decision bit. We can now apply the isolation lemma (Lemma 6) to obtain the desired statement, and note that the running time to implement $(\langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_{B'}, B')$ when accessed by $\mathbf{D}$ is roughly $\sum_{j=i+1}^{m} t_{\langle F_j, I_j \rangle_{B_j}}(q_j)$ by construction of $\langle F_j, I_j \rangle_{B_j}$. $\square$

From the Isolation Lemma to the Product Theorem. For given $t, q, \lambda > 0$ (with $\overline{\gamma} := \gamma/(m-1)$, $q_1, \ldots, q_m$ such that q queries result in $q_1, \ldots, q_m$ queries of $\mathbf{C}(\cdot)$ to the respective subsystems, we let $t'_1, \ldots, t'_m$ and $q'_1, \ldots, q'_m$ be as in the statement of Theorem 9, and define $\delta_i := \Delta_{t'_i, q'_i}(\mathbf{F}_i, \mathbf{I}_i)$. Furthermore, recall that $\delta_i \leq \frac{1}{2}$ for all $i = 1, \ldots, m-1$.

From now on, let $\mathbf{D}$ be the distinguisher with running time t making q queries to $\mathbf{C}(\cdot)$ such that

$$\Delta^{\mathbf{D}}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)) = \Gamma^{\mathbf{D}}(\mathbf{C}(\langle \mathbf{F}_{[1,m]}, \mathbf{I}_{[1,m]} \rangle_B), B) \geq 2^{m-1} \cdot \delta_1 \cdots \delta_m + 2\gamma$$

for a uniform random bit B. Define $\gamma_i := (m-1-i) \cdot \overline{\gamma}$ for $i = 0, \ldots, m-1$ and consider the statements $\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, b)$ for a bit $b \in \{0,1\}$ and valid states $\sigma_1, \ldots, \sigma_i$ of $G_1, \ldots, G_i$ respectively, with $|\sigma_j| \leq l_j$ for $j = 1, \ldots, i$ which holds if and only if

$$\begin{aligned}\Gamma^{\mathbf{D}}(\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_i[\sigma_i] \| \langle \mathbf{F}_{[i+1,m]}, \mathbf{I}_{[i+1,m]} \rangle_{B'}), b \oplus B') &\geq 2^{m-1-i} \cdot \delta_{i+1} \cdots \delta_m + \gamma_i \\ &\geq 2 \cdot \delta_{i+1} \cdot (2^{m-1-(i+1)} \cdot \delta_{i+2} \cdots \delta_m + \gamma_{i+1}) + \overline{\gamma},\end{aligned}$$

where we have used the facts that $\gamma_i := \gamma_{i-1} + \overline{\gamma}$ and $2 \cdot \delta_{i+1} \leq 1$. In particular, by our assumption $\mathsf{STAT}_0(0)$ holds. Furthermore, given $\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_{i-1}, b)$ holds for some i we apply the isolation Lemma (Lemma 10). Note that condition (ii) cannot hold, as otherwise there exists a distinguisher $\mathbf{D}'_i$ such that $\Delta^{\mathbf{D}'_i}(\mathbf{F}_i, \mathbf{I}_i) > \delta_i$, but this contradicts the assumed indistinguishability of $\mathbf{F}_i$ and $\mathbf{I}_i$, as we have defined δ_i with respect to the maximal running time of a constructed distinguisher $\mathbf{D}'_i$. Therefore, condition (i) must hold: In other words, we have shown that for all $i = 0, \ldots, m-2$

$$\mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, b) \Longrightarrow \exists \sigma_{i+1}, b' : \mathsf{STAT}_i(\sigma_1, \ldots, \sigma_i, \sigma_{i+1}, b \oplus b'),$$

where all the states $\sigma_1, \sigma_2, \ldots$ are valid for their respective implementations and are such that $|\sigma_i| \leq l_i$. Iterating the argument we obtain that there exist $\sigma_1, \ldots, \sigma_{m-1}$ and a bit b such that

$$\Gamma^{\mathbf{D}}(\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_{m-1}] \| \langle \mathbf{F}_m, \mathbf{I}_m \rangle_{B_m}), b \oplus B_m) > \delta_m + \widetilde{\gamma}_{m-1} = \delta_m.$$

However, we can now consider the distinguisher $\mathbf{D}' := \mathbf{D}(\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_i] \| \cdot)) \oplus b$ which given access to $\mathbf{S} \in \{\mathbf{F}_m, \mathbf{I}_m\}$ simulates $\mathbf{D}$ interacting with $\mathbf{C}(G_1[\sigma_1] \| \cdots \| G_{m-1}[\sigma_i] \| \mathbf{S})$, obtaining output b'. Finally, $\mathbf{D}'$ outputs $b \oplus b'$. The distinguisher $\mathbf{D}'$ obtains advantage larger than $> \delta_m$, contradicting the assumed indistinguishability of $\mathbf{F}_m$ and $\mathbf{I}_m$.

4.3 Applications of the Product Theorem

Sums of PRFs. Let $\mathbf{F}_1, \ldots, \mathbf{F}_m : \mathcal{X} \to \mathcal{Y}$ be cc-stateless random functions (in fact, $\mathbf{F}_m$ can possibly be stateful), and let $\star$ be a quasi-group operation on $\mathcal{Y}$. The operator $\star$ is neutralizing, as discussed in Section 2.2, for $\mathbf{F}_1, \ldots, \mathbf{F}_m$ and ideal systems $\mathbf{I}_1 = \cdots = \mathbf{I}_m = \mathbf{R}$, where $\mathbf{R} : \mathcal{X} \to \mathcal{Y}$ is a URF. In order to simplify the time complexity statements, we assume that there exist efficient algorithms implementing $\mathbf{F}_i(\cdot)$ such that $\mathbf{F}_i(s, x)$ is computed in time t_{F_i} given s and x (this holds in the interesting case where we apply the result to PRFs) and elements of $\mathcal{Y}$ can be encoded using $\ell \approx \log |\mathcal{Y}|$ bits. Note that the canonical implementation of $\mathbf{R}$ keeps a linearly-growing state of size $s = \mathcal{O}(q \cdot \ell)$ after q queries, and answers each query in time $\mathcal{O}(\log(s))$. Therefore, with $t_{\langle F_i, R \rangle_{B_i}}(q, s) = \mathcal{O}(q \cdot \max\{t_{F_i}, \log(s + q\ell)\})$ and $l_{<i} = \mathcal{O}((i-1)\varphi q \ell)$, we apply Theorem 9 to obtain the following result (we tacitly assume that all advantages are bounded by $\frac{1}{2}$):

Corollary 11. *For all $t, q, \gamma > 0$,*

$$\Delta_{t,q}(\mathbf{F}_1 \star \cdots \star \mathbf{F}_m, \mathbf{R}) \leq 2^{m-1} \cdot \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\mathbf{F}_i, \mathbf{R}) + 2\gamma.$$

We remark that the analogous result for PRGs follows as a special case, since a PRG can be seen as a one-input PRF.

A weaker version of this result was shown by Dodis et al. [6] for the special case $\star = \oplus$: Their bounds depend in particular on the number of queries, and is hence far away from the tight information-theoretic bound which is matched (up to the term γ) by our result.

Cascade of PRPs. Let $\mathbf{P} : \{0,1\}^n \to \{0,1\}^n$ be a URP and let $\mathbf{Q}_1, \ldots, \mathbf{Q}_m : \{0,1\}^n \to \{0,1\}^n$ be cc-stateless random permutations. Recall that the $\triangleright$ operator is neutralizing for $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ (all with ideal system $\mathbf{P}$), as well as for $\langle \mathbf{Q}_1 \rangle, \ldots, \langle \mathbf{Q}_m \rangle$ (all with ideal system $\langle \mathbf{P} \rangle$). As above, we assume that both $\mathbf{Q}_i(s, x)$ and $\mathbf{Q}_i^{-1}(s, y)$ are computable in time t_{Q_i}. Furthermore, simulating the URP $\mathbf{P}$ (as well as the two-sided URP $\langle \mathbf{P} \rangle$) requires the same complexity as implementing a URF. Therefore, with $t_{\langle Q_i, P \rangle_{B_i}}(q, s) = t_{\langle \langle Q_i \rangle, \langle P \rangle \rangle_{B_i}}(q, s) = \mathcal{O}(q \cdot \max\{t_{Q_i}, \log(s + qn)\})$ and $l_{<i} = \mathcal{O}((i-1)\varphi qn)$, Theorem 9 yields the following corollary:

Corollary 12. *For all $t, q, \gamma > 0$,*

$$\Delta_{t,q}(\mathbf{Q}_1 \triangleright \cdots \triangleright \mathbf{Q}_m, \mathbf{P}) \leq 2^{m-1} \cdot \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\mathbf{Q}_i, \mathbf{P}) + 2\gamma,$$

and

$$\Delta_{t,q}(\langle \mathbf{Q}_1 \rangle \triangleright \cdots \triangleright \langle \mathbf{Q}_m \rangle, \langle \mathbf{P} \rangle) \leq 2^{m-1} \cdot \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\langle \mathbf{Q}_i \rangle, \langle \mathbf{P} \rangle) + 2\gamma.$$

Furthermore, we note that $\mathbf{Q}_1$ is allowed to be stateful in the one-sided case, as Theorem 9 allows one system to be stateful: In fact, $\triangleright$ is not necessarily neutralizing whenever at least two permutations are stateful.

We remark that this is the first result considering *two-sided* PRPs, and even in the one-sided setting only the case $m = 2$ was considered by Luby and Rackoff [23], and subsequently extended to any constant m by Myers [28]. (Also, a slightly weaker result for the case $m = \mathcal{O}(\log \log n)$ is given in [28].) However, all of these results fall short of achieving security amplification, as they are not able to transform a δ-PRP for a non-negligible δ into a fully secure PRP.

Asymptotic Interpretation. In the asymptotic setting, Corollary 11 can be interpreted as follows. If $\mathbf{F}(\cdot)$ is a δ-PRF (for some $\delta < \frac{1}{2}$), it follows that $\mathbf{F}(S_1) \star \cdots \star \mathbf{F}(S_m)$, for independent keys $S_1, \ldots, S_m$, is a $2^{m-1} \cdot \delta^m$-PRF: For t, q polynomially bounded in n, we have $\Delta_{t,q}(\mathbf{F}(S_1) \star \cdots \star \mathbf{F}(S_m), \mathbf{R}) \leq 2^{m-1} \cdot \delta^m + \nu(n) + 1/p(n)$ for all polynomials p (and some negligible function ν), as both t'_i and q'_i are polynomial as well. Moreover, Corollary 12 implies a similar statement for the cascade of δ-(two-sided)-PRPs.

5 A Strong Product Theorem for Randomized Neutralizing Constructions

5.1 A Product Theorem from Self-Independence

Since Theorem 9 holds for arbitrary neutralizing constructions, one cannot avoid the factor 2^{m-1} in the bound. This section shows that a subclass of neutralizing constructions satisfying a simple information-theoretic property yield a *strong* product theorem, i.e., the obtained upper bound is roughly the product of the individual advantages.

Self-Independence. The notion of self-independence of an ideal system $\mathbf{I}$ under a construction $\mathbf{C}(\cdot)$ captures the property that a computationally unbounded distinguisher cannot tell apart the scenario where the *same* instance of $\mathbf{I}$ is accessed through independent instances of $\mathbf{C}(\cdot)$ from the setting where each instance of $\mathbf{C}(\cdot)$ accesses an independent instance of $\mathbf{I}$.

Definition 3. The system $\mathbf{I}$ is *η-self-independent under* $\mathbf{C}(\cdot)$ for a function $\eta : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$, if for all $q, \lambda > 0$, the best (information-theoretic) distinguishing advantage when allowing q queries to each subsystem satisfies

$$\Delta_{q,\ldots,q}(\mathbf{C}_1(\mathbf{I}) \| \ldots \| \mathbf{C}_\lambda(\mathbf{I}), \mathbf{C}_1(\mathbf{I}_1) \| \ldots \| \mathbf{C}_\lambda(\mathbf{I}_\lambda)) \leq \eta(q, \lambda),$$

where $\mathbf{C}_1(\cdot), \ldots, \mathbf{C}_\lambda(\cdot)$ and $\mathbf{I}_1, \ldots, \mathbf{I}_\lambda$ are independent copies of $\mathbf{C}(\cdot)$ and $\mathbf{I}$, respectively.

As an example, consider the construction $\mathbf{C}(\cdot)$ which generates a (secret) random n-bit offset Z, and given access to a random function $\mathbf{F} : \{0,1\}^n \to \{0,1\}^n$, $\mathbf{C}(\mathbf{F})$ returns $\mathbf{F}(x \oplus Z)$ upon each query x. It is not hard to show, e.g. using the tools from [24], that a URF $\mathbf{R} : \{0,1\}^n \to \{0,1\}^n$ is η-self-independent under $\mathbf{C}(\cdot)$ for $\eta(q, \lambda) \leq \frac{q^2\lambda^2}{2} \cdot 2^{-n}$, i.e., the probability that for some distinct $i \neq j$ the instances $\mathbf{C}_i(\cdot)$ and $\mathbf{C}_j(\cdot)$ invoke $\mathbf{R}$ with the same input.

Restricted Attacks on Cryptographic Functions. Indistinguishability-based security definitions can also be weakened by restricting the distinguisher's access to the given system. For instance, the standard PRF notion considering an (adaptive) *chosen-input attack* can be weakened to non-adaptive chosen-input attacks or even *(known) random-input* attacks. (Keyed functions which are secure under the latter notion are usually called *weak PRFs* [30] in the literature.[26]) This is conveniently modeled by letting the distinguisher access either of $\mathbf{E}(\mathbf{F})$ and $\mathbf{E}(\mathbf{G})$, where the construction $\mathbf{E}(\cdot)$ enforces a particular type of access, and $\mathbf{F}$ and $\mathbf{G}$ are the systems to be distinguished. For a chosen-input attack, $\mathbf{E}$ would just give full access to the underlying system (i.e. $\mathbf{E}(\cdot)$ is the *identity*), and the following are two additional examples:

- *Random-input attacks* against an $(\mathcal{X}, \mathcal{Y})$-system are modeled by $\mathbf{K}(\cdot)$ that, upon each invocation (with some dummy input), generates a fresh uniformly-chosen element $r \in \mathcal{X}$, makes a query with input r to the given subsystem, obtaining $y \in \mathcal{Y}$, and returns (r, y).
- For a quasi-group operation $*$ on $\mathcal{X}$ (usually $\oplus$), a *random-offset attack* is modeled by a construction $\mathbf{Z}(\cdot)$ which initially generates a random offset $Z \in \mathcal{X}$, and upon each invocation with input $x \in \mathcal{X}$, makes a query to the given subsystem with input $x \star Z$, and outputs the returned value y. (To our knowledge, this notion was not previously considered in the literature.)

[26] The name is slightly misleading within the context of this paper, as it can been used [29] to describe an ϵ-PRF for a non-negligible $\epsilon < 1$.

A feature of the product theorem of this section is that it is easily applicable also to the restricted-access case.

THE STRONG PRODUCT THEOREM. In the following, let $\mathbf{C}(\cdot)$ be a neutralizing construction for systems $\mathbf{F}_1, \ldots, \mathbf{F}_m$ and ideal system $\mathbf{I}_1, \ldots, \mathbf{I}_m$, all of which (with the possible exception of $\mathbf{F}_m$ and $\mathbf{I}_m$) are cc-stateless. Furthermore, we assume that $\mathbf{F}_i(\cdot)$ is efficiently implementable for all $i = 1, \ldots, m-1$,[27] and the corresponding (short) random variable S_i is drawn from the set $\mathcal{S}_i$. Also, we let $\mathbf{E}(\cdot)$ be a construction restricting access to $\mathbf{F}_i$ and $\mathbf{I}_i$. Finally, for $i = 1, \ldots, m$, and for $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$ we define

$$\mathbf{C}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot) := \mathbf{C}(\mathbf{F}_1(s_1), \ldots, \mathbf{F}_{i-1}(s_{i-1}), \cdot, \mathbf{F}_{i+1}, \ldots, \mathbf{F}_m)$$

and consider the following two properties:

(i) For all $i = 1, \ldots, m-1$ (the property is not necessary for $i = m$) and all $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$, the ideal system $\mathbf{I}_i$ is η-self-independent under the construction $\mathbf{C}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$ for some small function η.

(ii) For all $i = 1, \ldots, m$ and $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$, there exists a construction $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$ with the property that for independent instances $\mathbf{T}_1(\cdot), \ldots, \mathbf{T}_\lambda(\cdot)$ and $\mathbf{C}_1(\cdot), \ldots, \mathbf{C}_\lambda(\cdot)$ of $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$ and $\mathbf{C}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$, respectively, and all compatible systems $\mathbf{S}$,

$$\mathbf{T}_1(\mathbf{E}(\mathbf{S}))\| \cdots \|\mathbf{T}_\lambda(\mathbf{E}(\mathbf{S})) \equiv \mathbf{C}_1(\mathbf{S})\| \cdots \|\mathbf{C}_\lambda(\mathbf{S}).$$

We define t_{T_i} as the maximal complexity (taken over all $s_1, \ldots, s_{i-1}$) for implementing the construction $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$.

In the following, we define $\lambda := \left(\frac{4m}{\gamma}\right)^2 \cdot \ln\left(\frac{4m}{\gamma}\right)$, for understood m and γ.

Theorem 13 (Strong Product Theorem)**.** *Let $q > 0$, $\mathbf{C}(\cdot)$, and $\mathbf{E}(\cdot)$ be as above satisfying conditions* (i) *and* (ii)*, and assume that upon q queries, $\mathbf{C}(\cdot)$ makes at most q_i queries to the i-th subsystem. Then, for all $t, \gamma > 0$,*

$$\Delta_{t,q}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)) \leq \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\mathbf{E}(\mathbf{F}_i), \mathbf{E}(\mathbf{I}_i)) + \sum_{i=1}^{m-1} \eta(q_i, \lambda) + \gamma,$$

where $t'_i := \lambda \cdot (t + \mathcal{O}(t_{T_i}(q_i)))$ and $q'_i := \lambda \cdot q_i$ for all $i = 1, \ldots, m-1$, whereas $t'_m := t + \mathcal{O}(t_{T_m}(q))$ and $q'_m := q_m$.

We first give a proof sketch for the case $m = 2$ (and $\mathbf{E}(\cdot)$ being the identity) to illustrate the main ideas used in the proof. The full proof of Theorem 13 is deferred to Section 5.2. It abstracts and generalizes the proof technique used by Myers [29] (which was in turn based on Levin's proof of the XOR-lemma [22, 11]).

[27] While the same techniques as in the proof of Theorem 5 could be used to address general cc-stateless systems where $\mathbf{F}(\cdot)$ is not necessarily efficient, this will not be necessary for our applications.

PROOF IDEA FOR $m = 2$. For convenience we set $\mathbf{F}_1 = \mathbf{F} = \mathbf{F}(S), \mathbf{F}_2 = \mathbf{G}$, $\mathbf{I}_1 = \mathbf{I}$, $\mathbf{I}_2 = \mathbf{J}$, and we omit $\mathbf{E}(\cdot)$, as it is the identity, We assume that S is drawn (according to some distribution) from a set $\mathcal{S}$. Let $\mathbf{D}$ be a distinguisher contradicting Theorem 13, i.e., such that

$$\mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}, \mathbf{G})) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{I}, \mathbf{J})) = 1] > \delta \cdot \epsilon + \eta + \gamma.$$

We show that there exist distinguishers $\mathbf{D}'$ and $\mathbf{D}''$ with complexities close to that of $\mathbf{D}$ such that $\Delta^{\mathbf{D}'}(\mathbf{F}, \mathbf{I}) > \delta$ or $\Delta^{\mathbf{D}''}(\mathbf{G}, \mathbf{J}) > \epsilon$ holds.

For all $s \in \mathcal{S}$, we define

$$\alpha(s) := \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}(s), \mathbf{G}) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{I}, \mathbf{J})) = 1].$$

We note that, on the one hand, if there existed some s with $\alpha(s) > \epsilon$, we would be able to distinguish $\mathbf{G}$ from $\mathbf{J}$ by means of a distinguisher $\mathbf{D}'' := \mathbf{D}(\mathbf{C}(\mathbf{F}(s), \cdot))$, since using the fact that $\mathbf{C}(\cdot)$ is neutralizing we have

$$\Delta^{\mathbf{D}''}(\mathbf{G}, \mathbf{J}) \geq \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}(s), \mathbf{G})) = 1] - \mathsf{P}[\mathbf{D}(\underbrace{\mathbf{C}(\mathbf{F}(s), \mathbf{J})}_{\equiv \mathbf{C}(\mathbf{I}, \mathbf{J})}) = 1] = \alpha(s) > \epsilon.$$

On other hand, if $\alpha(s) \leq \epsilon$, a simple averaging argument implies the existence of a set $\mathcal{G} \subseteq \mathcal{S}$ such that $\mathsf{P}[S \in \mathcal{G}] > \delta + \eta$ and $\alpha(s) > \gamma$ for all $s \in \mathcal{G}$.

Let now $\pi := \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{I}, \mathbf{J}) = 1]$: We consider a distinguisher $\mathbf{D}'$ which given access to $\mathbf{S} \in \{\mathbf{F}, \mathbf{I}\}$ computes the outputs $o_1, \ldots, o_\lambda \in \{0, 1\}$ obtained by letting λ independent instances of $\mathbf{D}(\mathbf{C}(\cdot, \mathbf{G}))$ sequentially interact with the given system $\mathbf{S}$. Finally, it outputs 1 if and only if the average $\overline{o}$ of these bits is larger than $\pi + \gamma/2$, and 0 otherwise.

On the one hand, if $\mathbf{S} = \mathbf{F}(s)$ for some fixed $s \in \mathcal{S}$, the system $\mathbf{S}$ is stateless and hence the variables $o_1, \ldots, o_\lambda$ are independent with $\mathsf{E}[o_j] = \mathsf{P}[\mathbf{D}(\mathbf{D}(\mathbf{F}(s), \mathbf{G})) = 1] > \pi + \gamma$ (since $\alpha(s) > \gamma$): By Hoeffding's inequality (cf. Appendix A) the average $\overline{o}$ of $o_1, \ldots, o_\lambda \in \{0, 1\}$ is larger than $\pi + \gamma/2$, except with negligible probability, and thus 1 is output with overwhelming probability by $\mathbf{D}'$. On the other hand, the variables $o_1, \ldots, o_\lambda$ are in general *not* independent if $\mathbf{S} = \mathbf{I}$. However, we can exploit self-independence under $\mathbf{C}(\cdot, \mathbf{G})$ to show that the probability that $\overline{o} < \pi + \gamma/2$ is at most an additive term η larger than the probability that the same happens for the average $\overline{o}'$ of the outputs $o'_1, \ldots, o'_\lambda$ of λ *independent* instances of $\mathbf{D}(\mathbf{C}(\mathbf{I}, \mathbf{G}))$. Due to independence, this last probability is again negligible by Hoeffding's inequality, and thus $\mathbf{D}'$ outputs 1 with probability bounded by η plus a negligible term. Approximating overwhelming and negligible by 1 and 0, respectively, the final advantage of $\mathbf{D}'$ is hence at least $\mathsf{P}[S \in \mathcal{G}] \cdot 1 - \eta \geq \delta + \eta - \eta = \delta$.

The full proof is obtained by a careful analysis of the involved quantities and extending this argument (in an inductive way) to m systems.

5.2 Proof of the Strong Product Theorem

This section presents a proof of Theorem 13. As in the cases of Theorems 5 and 9, the proof is based on the isolation technique. At the end of this section we briefly mention the main changes to be done in order to obtain a uniform reduction.

THE ISOLATION LEMMA. In the following, we let $\mathbf{I} := \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_m)$ and define $\lambda = \lambda(m, \gamma) := \left(\frac{4}{\gamma}\right)^2 \cdot \ln\left(\frac{4}{\gamma}\right)$. Also, in the following, let λ be defined as above with respect to some understood γ, whereas with respect to $\overline{\gamma}$ we define $\lambda := \left(\frac{4}{\overline{\gamma}}\right)^2 \cdot \ln\left(\frac{4}{\overline{\gamma}}\right)$.

Lemma 14. *Let $i \in \{1, \ldots, m-1\}$ and $\overline{\gamma} > 0$. Assume that $\mathbf{D}$ makes q queries to $\mathbf{C}(\cdot)$ resulting in q_j queries to each of the m sub-systems $\mathbf{F}_j$, and that for values $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$ we have*

$$\mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}_1(s_1), \ldots, \mathbf{F}_{i-1}(s_{i-1}), \mathbf{F}_i, \cdots, \mathbf{F}_m)) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1] > \epsilon \cdot \delta + \zeta + \overline{\gamma},$$

then at least one of the two following statements holds:

(i) There exists a value $s_i \in \mathcal{S}_i$ such that

$$\mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}_1(s_1), \ldots, \mathbf{F}_i(s_i), \mathbf{F}_{i+1}, \cdots, \mathbf{F}_m)) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1] > \epsilon + \zeta - \eta(q_i, \lambda).$$

(ii) There exists a distinguisher $\mathbf{D}'_i$ such that $\Delta^{\mathbf{D}'_i}(\mathbf{E}(\mathbf{F}_i), \mathbf{E}(\mathbf{I}_i)) > \delta$ which makes $q'_i := \lambda \cdot q_i$ queries and has time complexity $t'_i := \lambda \cdot (t + t_{T_i}(q))$.

From the Isolation Lemma to Theorem 13. Similarly to the cases of Theorems 5 and 9, we apply the isolation lemma iteratively to obtain Theorem 13. Fix $t, q, \gamma > 0$ (and let $\overline{\gamma} := \gamma/(m-1)$) such that upon q queries $\mathbf{C}(\cdot)$ makes $q_1, \ldots, q_m$ queries to its subsystems. For all $i = 1, \ldots, m$, let t'_i, q'_i as in the statement of Theorem 13, and let $\delta_i := \Delta_{t'_i, q'_i}(\mathbf{F}_i, B_i)$.

Assume towards a contradiction that there exists a distinguisher $\mathbf{D}$ with running time t making q queries to $\mathbf{C}$ such that

$$\Delta^{\mathbf{D}}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m), \mathbf{I}) \geq \mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_m)) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1] \geq \delta_1 \cdots \delta_m + \sum_{i=1}^{m-1} \eta(q_i, \lambda) + \gamma.$$

and define $\gamma_i := (m - 1 - i) \cdot \overline{\gamma}$ for $i = 0, \ldots, m-1$ and consider the statements $\mathsf{STAT}_i(s_1, \ldots, s_i)$ which holds if and only if

$$\begin{aligned}\mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}_1(s_1), \ldots, \mathbf{F}_i(s_i), \mathbf{F}_{i+1}, \ldots, \mathbf{F}_m)) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1] &\geq \delta_{i+1} \cdots \delta_m + \sum_{j=i+1}^{m-1} \eta(q_j, \lambda) + \gamma_i \\ &\geq \delta_{i+1} \cdot (\delta_{i+2} \cdots \delta_m + \gamma_{i+1}) + \sum_{j=i+1}^{m-1} \eta(q_j, \lambda) + \overline{\gamma}.\end{aligned}$$

In particular, by our assumption STAT_0 holds. By the Isolation Lemma (Lemma 14) and the fact that $\Delta_{t'_{i+1}, q'_{i+!}}(\mathbf{F}_{i+1}, \mathbf{I}_{i+1}) \leq \delta_{i+1}$, given $\mathsf{STAT}_i(s_1, \ldots, s_i)$ holds for some $s_1, \ldots, s_i$, then condition (ii) cannot hold, and hence condition (i) *must* hold, i.e., there must exists s_{i+1} such that $\mathsf{STAT}_{i+1}(s_1, \ldots, s_{i+1})$ holds. In other words, we have shown that for all $i = 0, \ldots, m-2$

$$\mathsf{STAT}_i(s_1, \ldots, s_i) \Longrightarrow \exists s_{i+1} : \mathsf{STAT}_{i+1}(s_1, \ldots, s_{i+1}).$$

Iterating the argument we obtain that there exist $s_1, \ldots, s_{m-1}$ such that $\mathsf{STAT}_{m-1}(s_1, \ldots, s_{m-1})$ holds, i.e.

$$\mathsf{P}[\mathbf{D}(\mathbf{C}(\mathbf{F}_1(s_1), \ldots, \mathbf{F}_{m-1}(s_{m-1}), \mathbf{F}_m) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1] \geq \delta_m.$$

This, however, gives rise to a distinguisher $\mathbf{D}'_m := \mathbf{D}(\mathbf{T}_{m-1, s_1, \ldots, s_{m-1}}(\cdot))$ with advantage δ_m and the given complexity t'_m, contradicting the indistinguishability assumption on $\mathbf{F}_m$ and $\mathbf{I}_m$.

Distinguisher $\mathbf{D}'_i$: // for $\mathbf{S} \in \{\mathbf{E}(\mathbf{F}_i), \mathbf{E}(\mathbf{I}_i)\}$

for $j := 1, \ldots, \lambda$ **do**
  $o_j := \mathbf{D}(\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\mathbf{S}))$;
$\overline{o} := \frac{1}{\lambda} \cdot \sum_{j=1}^{\lambda} o_j$;
if $\overline{o} - \pi > \overline{\gamma}/4$ **then**
  output 1
else
  output 0;

Distinguisher $\mathbf{D}''_i$: // for $\mathbf{S} \in \{\mathbf{F}_i, \mathbf{I}_i\}$

for $j := 1, \ldots, \lambda$ **do**
  $o_j := \mathbf{D}(\mathbf{C}^{(i)}(\mathbf{S}))$;
$\overline{o} := \frac{1}{\lambda} \cdot \sum_{j=1}^{\lambda} o_j$;
if $\overline{o} - \pi > \overline{\gamma}/4$ **then**
  output 1
else
  output 0;

Figure 3: Distinguishers $\mathbf{D}'_i$ and $\mathbf{D}''_i$ in the proof of Lemma 14. In both cases, π equals $\mathsf{P}[\mathbf{D}(\mathbf{I}) = 1]$.

Proof of the Isolation Lemma (Lemma 14). We fix $i \in \{1, \ldots, m\}$ as in the statement of the lemma. It is convenient to denote $\eta := \eta(q_i, \lambda)$, as well as $\mathbf{C}^{(i)}(\cdot) := \mathbf{C}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$. Note that since $\mathbf{C}(\cdot)$ is neutralizing, we have $\mathbf{C}^{(i)}(\mathbf{I}_i) \equiv \mathbf{I}$. Furthermore, we also define the function $\alpha : \mathcal{S}_i \to [-1, 1]$ such that

$$\alpha(s) := \mathsf{P}[\mathbf{D}(\mathbf{C}^{(i)}(\mathbf{F}(s))) = 1] - \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1].$$

Clearly, by the assumption of the lemma, we have $\mathsf{E}[\alpha(S_i)] > \epsilon \cdot \delta + \zeta + \overline{\gamma}$. Furthermore, if there existed an s_i such that $\alpha(s_i) > \epsilon + \zeta - \eta$, the first statement of the lemma would be directly true. Therefore, we assume in the following that $\alpha(s) \leq \epsilon + \zeta - \eta$ for all $s \in \mathcal{S}_i$. Furthermore, we define the set

$$\mathcal{G} := \{s \in \mathcal{S}_i \,|\, \alpha(s) > \overline{\gamma}/2\}.$$

The following claim gives a lower bound on the probability that S_i is in the set $\mathcal{G}$.

Claim 1. $\mathsf{P}[S_i \in \mathcal{G}] > \delta + \eta + \overline{\gamma}/2$.

Proof. Assume, towards a contradiction, that $\mathsf{P}[S_i \in \mathcal{G}] \leq \delta + \overline{\gamma}/2 + \eta$. Then,

$$\begin{aligned}\mathsf{E}[\alpha(S_i)] &= \mathsf{P}[S_i \in \mathcal{G}] \cdot \mathsf{E}[\alpha(S_i)|S_i \in \mathcal{G}] + \mathsf{P}[S_i \notin \mathcal{G}] \cdot \mathsf{E}[\alpha(S_i)|S_i \notin \mathcal{G}] \\ &\leq (\delta + \eta + \overline{\gamma}/2) \cdot (\epsilon + \zeta - \eta) + \overline{\gamma}/2 \leq \delta\epsilon + \zeta - \eta + \eta + \overline{\gamma}/2 + \overline{\gamma}/2 = \delta\epsilon + \zeta + \overline{\gamma}.\end{aligned}$$

which is in contradiction with the assumption on $\mathbf{D}$. □

We construct the distinguisher $\mathbf{D}'_i$ for distinguishing $\mathbf{E}(\mathbf{F}_i)$ from $\mathbf{E}(\mathbf{I}_i)$ as specified on the left-hand-side of Figure 3: With $\pi := \mathsf{P}[\mathbf{D}(\mathbf{I}) = 1]$, $\mathbf{D}'_i$ simulates λ independent instances of $\mathbf{D}(\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot))$ sequentially interacting with the given system $\mathbf{S}$, and computes the average $\overline{o}$ of the λ bits output by these instances, and finally outputs 1 if $\overline{o} > \pi + \overline{\gamma}/2$, and 0 otherwise. Note that by the property of $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}$ we can equivalently consider a distinguisher $\mathbf{D}''_i$ for $\mathbf{F}_i$ and $\mathbf{I}_i$ (depicted on the right-hand-side of Figure 3) which computes the bits $o_1, \ldots, o_\lambda$ by letting independent instances of $\mathbf{D}(\mathbf{C}^{(i)}(\cdot))$ interact with the given system. Clearly, $\Delta^{\mathbf{D}'_i}(\mathbf{E}(\mathbf{F}_i), \mathbf{E}(\mathbf{I}_i)) = \Delta^{\mathbf{D}''_i}(\mathbf{F}_i, \mathbf{I}_i)$, and we thus focus on the latter advantage.

The remainder of the proof consists of the following two lemmas.

Lemma 15. $\mathsf{P}[\mathbf{D}''_i(\mathbf{F}_i(s)) = 1] > 1 - e^{-\lambda(\overline{\gamma}/4)^2}$ *for all* $s \in \mathcal{G}$.

Lemma 16. $\mathsf{P}[\mathbf{D}''_i(\mathbf{I}_i) = 1] \leq e^{-\lambda(\overline{\gamma}/4)^2} + \eta(q_i, \lambda)$.

Before we turn to the proofs of the two lemmas, we note that they suffice to obtain the isolation lemma, since by Claim 1

$$\begin{aligned}\Delta^{\mathbf{D}''_i}(\mathbf{F}_i, \mathbf{I}_i) &\geq \mathsf{P}[S_i \in \mathcal{G}] \cdot \mathsf{P}[\mathbf{D}''_i(\mathbf{F}_i(S_i)) = 1 | S_i \in \mathcal{G}] - \mathsf{P}[\mathbf{D}''_i(\mathbf{I}_i) = 1] \\ &> (\delta + \overline{\gamma}/2 + \eta)(1 - e^{-\lambda(\overline{\gamma}/4)^2}) - e^{-\lambda(\overline{\gamma}/4)^2} - \eta \\ &\geq \delta + \overline{\gamma}/2 + \eta - e^{-\lambda(\overline{\gamma}/4)^2} - e^{-\lambda(\overline{\gamma}/4)^2} - \eta = \delta + \overline{\gamma}/2 - 2e^{-\lambda(\overline{\gamma}/4)^2} = \delta\end{aligned}$$

for the chosen value of λ.

Proof of Lemma 15. Since the system $\mathbf{F}_i(s)$ is stateless, note that for a fixed $s \in \mathcal{G}$ the random variables $o_1, \ldots, o_\lambda$ are independent binary variables with

$$p(s) := \mathsf{P}[o_j = 1] = \mathsf{P}[\mathbf{D}(\mathbf{C}^{(i)}(\mathbf{F}_i(s))) = 1],$$

and thus $\mathsf{E}[\overline{o}] = p(s)$. We know that $p(s) - \pi > \overline{\gamma}/2$, since $s \in \mathcal{G}$, and thus $\overline{o} - \pi \leq \overline{\gamma}/4$ implies that $\overline{o} < p(s) - \overline{\gamma}/4$. Therefore, by Hoeffding's bound, we have

$$\mathsf{P}[\mathbf{D}''_i(\mathbf{F}_i(s)) = 0] = \mathsf{P}[\overline{o} - \pi \leq \overline{\gamma}/4] \leq \mathsf{P}[\overline{o} < p(s) - \overline{\gamma}/4] < e^{-\lambda(\overline{\gamma}/4)^2}. \qquad \square$$

Proof of Lemma 16. Consider a distinguisher $\overline{\mathbf{D}}''_i$ which given access to the parallel composition $\mathbf{S}_1 \| \ldots \| \mathbf{S}_\lambda$ of λ systems $\mathbf{S}_1, \ldots, \mathbf{S}_\lambda$, computes $o_j := \mathbf{D}(\mathbf{S}_j)$ for all $j = 1, \ldots, \lambda$, and then outputs its decision bit as $\mathbf{D}''_i$. Clearly, for independent instances $\mathbf{C}^{(i)}_1(\cdot), \ldots, \mathbf{C}^{(i)}_\lambda(\cdot)$ of $\mathbf{C}^{(i)}(\cdot)$ we have

$$\mathsf{P}[\mathbf{D}''_i(\mathbf{I}_i) = 1] = \mathsf{P}[\overline{\mathbf{D}}''_i(\mathbf{C}^{(i)}_1(\mathbf{I}_i) \| \ldots \| \mathbf{C}^{(i)}_\lambda(\mathbf{I}_i)) = 1],$$

i.e., where every system in the parallel composition accesses the same instance $\mathbf{I}_i$. Note that because of the η-self-independence of $\mathbf{I}_i$ under $\mathbf{C}^{(i)}$ we have, for independent instances $\mathbf{I}_{i,1}, \ldots, \mathbf{I}_{i,\lambda}$ of $\mathbf{I}_i$

$$\Delta^{\overline{\mathbf{D}}''_i}(\mathbf{C}^{(i)}_1(\mathbf{I}_i) \| \ldots \| \mathbf{C}^{(i)}_\lambda(\mathbf{I}_i), \mathbf{C}^{(i)}_1(\mathbf{I}_{i,1}) \| \ldots \| \mathbf{C}^{(i)}_\lambda(\mathbf{I}_{i,\lambda})) \leq \eta(q_i, \lambda),$$

from which we directly infer

$$\mathsf{P}[\mathbf{D}''_i(\mathbf{I}_i) = 1] \leq \eta(q_i, \lambda) + \mathsf{P}[\overline{\mathbf{D}}''_i(\mathbf{C}^{(i)}_1(\mathbf{I}_{i,1}) \| \ldots \| \mathbf{C}^{(i)}_\lambda(\mathbf{I}_{i,\lambda})) = 1]. \tag{8}$$

We hence upper bound the probability on the right-hand side. Note that, in this case, the variables $o_1, \ldots, o_\lambda$ are assigned *independent* equally-distributed values, with $\mathsf{P}[o_j = 1] = \pi$ all $j = 1, \ldots, \lambda$ (by the neutralizing property of $\mathbf{C}(\cdot)$), and thus $\mathsf{E}[\overline{o}] = \pi$ holds as well. Then, by Hoeffding's inequality,

$$\mathsf{P}[\overline{\mathbf{D}}''_i(\mathbf{C}^{(i)}_1(\mathbf{I}_{i,1}) \| \ldots \| \mathbf{C}^{(i)}_\lambda(\mathbf{I}_{i,\lambda})) = 1] = \mathsf{P}[\overline{o} - \pi > \overline{\gamma}/4] < e^{-\lambda \cdot (\overline{\gamma}/4)^2},$$

which combined with (8) yields the desired upper bound. $\square$

Obtaining a Uniform Reduction. We can adapt the above proof to obtain a uniform reduction, similarly to Section 3.4. Obviously, the main difference is that we cannot fix values $s_1, s_2, \ldots$ as in the above proof, but we have to be able to sample such values. An additional requirement is thus that there exists a uniform algorithm which (on input i and $s_1, \ldots, s_{i-1}$) implements $\mathbf{T}^{(i)}_{s_1, \ldots, s_{i-1}}(\cdot)$.

For increasing i, we attempt to sample a good s_i for which $\alpha(s_i) > \epsilon + \sum_{j=i+1}^{m-1} \eta(q_j, \lambda) + \overline{\gamma}'$ (for some extra term $\overline{\gamma}' < \overline{\gamma}$) by repeated sampling and testing, and if we fail, we have the guarantee that $\alpha(s_i) \leq \epsilon + \sum_{j=i+1}^{m-1} \eta(q_j, \lambda) + \overline{\gamma}'$ with good probability over the choice of S_i, and we run the distinguisher $\mathbf{D}'_i$. The only difference is that the probability π must be estimated as well. Also, if states $s_1, \ldots, s_{m-1}$ are found, then we conclude as in the non-uniform case.

The computation has to be modified to take into account the extra error term $\overline{\gamma}'$

5.3 Applications of the Strong Product Theorem

This section presents a number of new results which follow as simple applications of Theorem 13. Let $\mathbf{Q}_1, \ldots, \mathbf{Q}_m : \{0,1\}^n \to \{0,1\}^n$ be cc-stateless random permutations, and let $\mathbf{F}_1, \ldots, \mathbf{F}_m : \{0,1\}^n \to \{0,1\}^\ell$ be cc-stateless random functions. In particular, we let $\mathbf{Q}_i \equiv \mathbf{Q}_i(S_i)$ for some random variable S_i over a set $\mathcal{S}_i$, for all $i = 1, \ldots, m$, and analogously we have $\mathbf{F}_i \equiv \mathbf{F}_i(S_i)$. Furthermore, let $\mathbf{P} : \{0,1\}^n \to \{0,1\}^n$ and $\mathbf{R} : \{0,1\}^n \to \{0,1\}^\ell$ be a URP and URF, respectively. Assume that $\mathbf{Q}_i(s,x)$ (and $\mathbf{Q}_i^{-1}(s,y)$) and $\mathbf{F}_i(s,x)$ can be computed in time t_{Q_i} and t_{F_i}, respectively, for all s, x, and y.

5.3.1 Randomized Cascade of PRPs

The perhaps most surprising application is a *strong* product theorem for (two-sided) PRPs. We modify the (two-sided) cascade $\langle \mathbf{Q}_1 \rangle \rhd \cdots \rhd \langle \mathbf{Q}_m \rangle$ by choosing two independent random offsets that are added to the inputs and the outputs, i.e., we consider $\langle \oplus_{Z_1} \rangle \rhd \langle \mathbf{Q}_1 \rangle \rhd \cdots \rhd \langle \mathbf{Q}_m \rangle \rhd \langle \oplus_{Z_2} \rangle$ for two independent uniform n-bit strings Z_1, Z_2, where for some $z \in \{0,1\}^n$ the system $\langle \oplus_z \rangle$ is the bi-directional mapping which answers a forward query $(x,+)$ with $x \oplus z$ and a backward query $(y,-)$ with $y \oplus z$. The computational overhead is minimal compared to the regular cascade, and requires only additional storage for two n-bit strings (which are to be seen as part of the secret key).[28] Clearly the neutralizing property of the original cascade is preserved.

The central observation is that the two additional random offsets ensure property (i) and allow for applying Theorem 13. More formally, for some $i \in \{1, \ldots, m\}$ and $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$ we define (for a system $\mathbf{S}$ implementing a two-sided permutation from n bits to n bits) the construction $\mathbf{N}(\cdot)$ as choosing two offsets Z_1, Z_2 independently and uniformly at random, as well as $S_{i+1}, \ldots, S_m$, and such that

$$\mathbf{N}(\mathbf{S}) := \langle \oplus_{Z_1} \rangle \rhd \langle \mathbf{Q}_1(s_1) \rangle \rhd \cdots \rhd \langle \mathbf{Q}_{i-1}(s_{i-1}) \rangle \rhd \mathbf{S} \rhd \langle \mathbf{Q}_{i+1}(S_{i+1}) \rangle \rhd \cdots \rhd \langle \mathbf{Q}_m(S_m) \rangle \rhd \langle \oplus_{Z_2} \rangle.$$

The following lemma states the self-independence of a (bi-directional) URP $\langle \mathbf{P} \rangle$ under $\mathbf{N}(\cdot)$.

Lemma 17. *Let $\mathbf{N}_1(\cdot), \ldots, \mathbf{N}_\lambda(\cdot)$ be independent instances of $\mathbf{N}$, and let $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda, \mathbf{P} : \{0,1\}^n \to \{0,1\}^n$ be independent URPs. Then*

$$\Delta_{q,\ldots,q}(\mathbf{N}_1(\langle \mathbf{P} \rangle) \| \cdots \| \mathbf{N}_\lambda(\langle \mathbf{P} \rangle), \mathbf{N}_1(\langle \mathbf{P}_1 \rangle) \| \cdots \| \mathbf{N}_\lambda(\langle \mathbf{P}_\lambda \rangle)) \leq \lambda^2 \cdot q^2 \cdot 2^{-n}.$$

In other words, $\langle \mathbf{P} \rangle$ is η-self-independent under $\mathbf{N}(\cdot)$ for $\eta(q,\lambda) := \lambda^2 q^2 2^{-n}$.

Proof. For notational simplicity, we define

$$\mathbf{H}_0 := \mathbf{N}_1(\langle \mathbf{P} \rangle) \| \cdots \| \mathbf{N}_\lambda(\langle \mathbf{P} \rangle) \text{ and } \mathbf{H}_1 := \mathbf{N}_1(\langle \mathbf{P}_1 \rangle) \| \cdots \| \mathbf{N}_\lambda(\langle \mathbf{P}_\lambda \rangle).$$

For both systems and all $t = 1, \ldots, \lambda$, let $\mathcal{X}_i^{(t)}$ contain those n-bit strings x for which when processing the first i queries $\mathbf{N}_t(\cdot)$ has issued a forward query $(x,+)$, or alternatively a backward query $(y,-)$ of $\mathbf{N}_t(\cdot)$ was answered with the value x by the given permutation. Analogously, let $\mathcal{Y}_i^{(t)}$ contain those n-bit strings y for which when processing the first i queries $\mathbf{N}_t(\cdot)$ has issued a backward query $(y,-)$, or alternatively a forward query $(x,+)$ of $\mathbf{N}_t(\cdot)$ was answered by y.

[28] The idea of adding random offsets at both ends of a permutation was already used by Even and Mansour [9], though in a completely different context.

We define on both systems $\mathbf{H}_0$ and $\mathbf{H}_1$ the respective MES $\mathcal{A} = A_0, A_1, \ldots$ and $\mathcal{B} = B_0, B_1, \ldots$ where A_i (B_i) fails after query i if for some $t \neq t'$ we have $\mathcal{X}_i^{(t)} \cap \mathcal{X}_i^{(t')} \neq \emptyset$ or $\mathcal{Y}_i^{(t)} \cap \mathcal{Y}_i^{(t')} \neq \emptyset$. Then, we can easily show that $\mathbf{H}_0|\mathcal{A} \equiv \mathbf{H}_1|\mathcal{B}$. Furthermore, we have for all i

$$\mathsf{p}^{\mathbf{H}_0}_{A_i|X^iY^{i-1}A_{i-1}} \geq \mathsf{p}^{\mathbf{H}_1}_{B_i|X_iY^{i-1}B_{i-1}}$$

because for system $\mathbf{H}_0$ $\mathcal{X}_i^{(t)} \cap \mathcal{X}_i^{(t')} = \emptyset$ holds for all $t \neq t'$ if and only if $\mathcal{Y}_i^{(t)} \cap \mathcal{Y}_i^{(t')} = \emptyset$ for all $t \neq t'$. Therefore, by Lemma 1 (ii) we obtain

$$\Delta^{\mathbf{D}}(\mathbf{H}_0, \mathbf{H}_1) \leq \nu^{\mathbf{D}}(\mathbf{H}_1, \overline{B_{\lambda q}})$$

for any distinguisher $\mathbf{D}$ issuing at most q queries to each subsystem (and thus making totally at most $\lambda \cdot q$ queries).

We thus focus on upper bounding the probability $\nu^{\mathbf{D}}(\mathbf{H}_1, \overline{B_{\lambda q}})$ in the following. More specifically, let $Z_1^{(t)}, Z_2^{(t)}, S_{i+1}^{(t)}, \ldots, S_m^{(t)}$ be the values taken by $Z_1, Z_2, S_{i+1}, \ldots, S_m$ in $\mathbf{N}_t(\cdot)$ for $t = 1, \ldots, \lambda$. Note that for all values $z_1^{(t)}, z_2^{(t)}$ and $s_{i+1}^{(t)}, \ldots, s_m^{(t)}$ taken by these variables, the system $\mathbf{H}_1$ behaves as the parallel composition of λ independent (bi-directional) random permutations. We can thus apply Lemma 2: We assume that we are given some possible interaction with input-output pairs $(u_1^{(t)}, v_1^{(t)}), \ldots, (u_q^{(t)}, v_q^{(t)})$ for $\mathbf{N}_t(\langle \mathbf{P}_t \rangle)$ for $t = 1, \ldots, \lambda$, i.e., the i'th query to $\mathbf{N}_t(\langle \mathbf{P}_t \rangle)$ either is a forward query $u_i^{(t)}$ with output $v_i^{(t)}$ or is a backward query $v_i^{(t)}$ with output $u_i^{(t)}$. Moreover, we define for all $i = 1, \ldots, q$ and $t = 1, \ldots, \lambda$

$$\begin{aligned} P_i^{(t)} &:= \big(\mathbf{Q}_1(s_1) \rhd \cdots \rhd \mathbf{Q}_{i-1}(s_{i-1})\big)(u_i^{(t)} \oplus Z_1^{(t)}) \\ Q_i^{(t)} &:= \big(\mathbf{Q}_m^{-1}(S_m^{(t)}) \rhd \cdots \rhd \mathbf{Q}_{i+1}^{-1}(S_{i+1}^{(t)})\big)(v_i^{(t)} \oplus Z_2^{(t)}), \end{aligned}$$

i.e., these are the corresponding input-output pairs for the underlying permutations $\langle \mathbf{P}_1 \rangle, \ldots, \langle \mathbf{P}_\lambda \rangle$. Note that all of these random variables are (individually) uniformly distributed and furthermore, given $t \neq t'$, i, j we have $\mathsf{P}[P_i^{(t)} = P_j^{(t')}] = 2^{-n}$ and $\mathsf{P}[Q_i^{(t)} = Q_j^{(t')}] = 2^{-n}$. Therefore, $\overline{B_{\lambda q}}$ implies that there exist distinct t, t' and (not necessarily distinct) i, j such that $P_i^{(t)} = P_j^{(t')}$ or $Q_i^{(t)} = Q_j^{(t')}$. By the union bound we conclude that

$$\nu^{\mathbf{D}}(\mathbf{H}_1, \overline{B_{\lambda q}}) \leq 2 \cdot \binom{\lambda}{2} \cdot q^2 \cdot 2^{-n} \leq \lambda^2 \cdot q^2 \cdot 2^{-n}. \qquad \square$$

Therefore, we can combine Theorem 13 (with $\mathbf{E}(\cdot)$ being the identity) and Lemma 17 to obtain the following result.

Corollary 18. *For all $t, q, \gamma > 0$, and independent uniform n-bit strings Z_1, Z_2,*

$$\Delta_{t,q}(\langle \oplus_{Z_1} \rangle \rhd \langle \mathbf{Q}_1 \rangle \rhd \cdots \rhd \langle \mathbf{Q}_m \rangle \rhd \langle \oplus_{Z_2} \rangle, \langle \mathbf{P} \rangle) \leq \prod_{i=1}^{m} \Delta_{t'_i, q'_i}(\langle \mathbf{Q}_i \rangle, \langle \mathbf{P} \rangle) + \tfrac{mq^2\lambda^2}{2^n} + \gamma,$$

where $t'_i := \lambda \cdot \big(t + \mathcal{O}(q \cdot \sum_{j \neq i} t_{Q_j})\big)$ and $q'_i := \lambda \cdot q$ for all $i = 1, \ldots, m-1$, whereas $t'_m := t + \mathcal{O}(q \cdot \sum_{j=1}^{m-1} t_{Q_j})$ and $q'_m := q$.

The result can be used to obtain a δ^m-two-sided PRP from any δ-two-sided PRP. (Note that the η-dependent term is negligible for polynomial t, q and any γ which is the inverse of a polynomial.) It can be shown that the second random offset Z_2 is superfluous in the one-sided case.

5.3.2 Sum of Random-Input PRFs

The construction $\mathbf{K}(\mathbf{F}_1 \oplus \cdots \oplus \mathbf{F}_m)$ (i.e. the XOR of the functions accessed in a random-input attack) is clearly neutralizing (the ideal system being $\mathbf{K}(\mathbf{R})$). However, the fact that $\mathbf{F}_1 \oplus \cdots \oplus \mathbf{F}_m$ and $\mathbf{R}$ are invoked on random inputs only allows for proving a much stronger result using Theorem 13, since it satisfies property (i) needed by the theorem.

More precisely, let us fix $i \in \{1, \ldots, m\}$ and $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$, and define for a random function $\mathbf{S} : \{0,1\}^n \to \{0,1\}^\ell$ the construction $\mathbf{W}(\cdot)$ as

$$\mathbf{W}(\mathbf{S}) := \mathbf{K}(\mathbf{F}_1(s_1) \oplus \cdots \oplus \mathbf{F}_{i-1}(s_{i-1}) \oplus \mathbf{S} \oplus \mathbf{F}_{i+1} \oplus \cdots \oplus \mathbf{F}_m).$$

Then, we can show the following.

Lemma 19. *Let $\mathbf{R}, \mathbf{R}_1, \ldots, \mathbf{R}_\lambda : \{0,1\}^n \to \{0,1\}^\ell$ be independent URFs, and let $\mathbf{W}_1(\cdot), \ldots, \mathbf{W}_\lambda(\cdot)$ be independent instances of $\mathbf{W}(\cdot)$, then*

$$\Delta_{q,\ldots,q}(\mathbf{W}_1(\mathbf{R}) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}), \mathbf{W}_1(\mathbf{R}_1) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}_\lambda)) \leq \frac{q^2\lambda^2}{2} 2^{-n}.$$

In other words, the URF $\mathbf{R}$ is η-self-independent under $\mathbf{W}$ for $\eta(q,\lambda) \leq \frac{q^2\lambda^2}{2} \cdot 2^{-n}$.

Proof. Query i to the sub-system t is associated with a random input $r_{i,t}$ at which the XOR is evaluated. Consider the MES $\mathcal{A} := A_0, A_1, \ldots$ such that A_i holds as long as within the first i queries there exists no two random inputs $r_{j,t} = r_{k,t'}$ for distinct $t \neq t'$: As long as $\mathcal{A}$ holds, the random function $\mathbf{R}$ is never evaluated at the same input by two constructions $\mathbf{W}_t(\cdot)$ and $\mathbf{W}_{t'}(\cdot)$, and thus

$$(\mathbf{W}_1(\mathbf{R}) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R})) | \mathcal{A} \equiv \mathbf{W}_1(\mathbf{R}_1) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}_\lambda),$$

from which we infer by Lemma 1 (i) that for all distinguishers $\mathbf{D}$ issuing at most q queries to each of the λ subsystems we have

$$\begin{aligned}\Delta^{\mathbf{D}}(\mathbf{W}_1(\mathbf{R}) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}), \mathbf{W}_1(\mathbf{R}_1) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}_\lambda)) &\leq \nu^{\mathbf{D}}(\mathbf{W}_1(\mathbf{R}) \| \cdots \| \mathbf{W}_\lambda(\mathbf{R}), \overline{A_{\lambda \cdot q}}) \\ &\leq \binom{\lambda}{2} q^2 2^{-n}.\end{aligned}$$

□

Moreover, for all i and keys $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$, the appropriate construction $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$ generates random keys $S_{i+1}, \ldots, S_m$ and whenever invoked, it issues a query to $\mathbf{K}(\mathbf{S})$, obtaining (r, y), and outputs the pair

$$\Big(r, \bigoplus_{j=1}^{i-1} \mathbf{F}_j(s_j, r) \oplus y \oplus \bigoplus_{j=i+1}^{m} \mathbf{F}_j(S_j, r)\Big).$$

It is easy to see that these constructions satisfy property (ii), since $\mathbf{K}(\cdot)$ evaluates the given function at a fresh random input upon each invocation. Therefore, Theorem 13 yields the following result.

Corollary 20. *For all $t, q, \gamma > 0$,*

$$\Delta_{t,q}(\mathbf{K}(\mathbf{F}_1 \oplus \cdots \oplus \mathbf{F}_m), \mathbf{K}(\mathbf{R})) \leq \prod_{i=1}^{m} \Delta_{t'_i,q'_i}(\mathbf{K}(\mathbf{F}_i), \mathbf{K}(\mathbf{R})) + \tfrac{(m-1)q^2\lambda^2}{2^{n+1}} + \gamma,$$

where $t'_i := \lambda\big(t + \mathcal{O}(q \cdot \sum_{j \neq i} t_{F_j})\big)$ and $q'_i := \lambda \cdot q$ for all $i = 1, \ldots, m-1$, whereas $t'_m := t + \mathcal{O}(q \cdot \sum_{j=1}^{m-1} t_{F_j})$ and $q'_m := q$.

The result holds for any other quasi-group operation. It is remarkable that XOR satisfies much stronger indistinguishability amplification properties under random-input attacks than under chosen-input attacks. This is particularly interesting, as a wide number of applications, such as secure symmetric message encryption, can efficiently be based on this weaker PRF notion (cf. [5, 27]).

5.3.3 Randomized XOR of PRFs

The first product theorem for PRFs, due to Myers [29], considered the neutralizing composition $\mathbf{Z}_1(\mathbf{F}_1) \oplus \cdots \oplus \mathbf{Z}_m(\mathbf{F}_m)$ for independent instances of $\mathbf{Z}(\cdot)$ (which we assume to be instantiated with $\oplus$, but any other quasi-group operation would work). We show that this result is implied by Theorem 13. In fact, the same result holds for the construction $\mathbf{Z}(\mathbf{F}_1 \oplus \cdots \oplus \mathbf{F}_m)$ using the *same* offset for all invocations.

More precisely, fix $i \in \{1, \ldots, m\}$ and $s_1 \in \mathcal{S}_1, \ldots, s_{i-1} \in \mathcal{S}_{i-1}$, and define the constructions $\mathbf{M}(\cdot)$ and $\mathbf{M}'(\cdot)$ such that for any random function $\mathbf{S} : \{0,1\}^n \to \{0,1\}^\ell$,

$$\mathbf{M}(\mathbf{S}) := \mathbf{Z}\big(\mathbf{F}_1(s_1) \oplus \cdots \oplus \mathbf{F}_{i-1}(s_{i-1}) \oplus \mathbf{S} \oplus \mathbf{F}_{i+1} \oplus \cdots \oplus \mathbf{F}_m\big).$$

and

$$\mathbf{M}'(\mathbf{S}) := \mathbf{Z}_1(\mathbf{F}_1(s_1)) \oplus \cdots \oplus \mathbf{Z}_{i-1}(\mathbf{F}_{i-1}(s_{i-1})) \oplus \mathbf{Z}_i(\mathbf{S}) \oplus \mathbf{Z}_{i+1}(\mathbf{F}_{i+1}) \oplus \cdots \oplus \mathbf{Z}_m(\mathbf{F}_m)\big),$$

Lemma 21. *Let $\mathbf{R}, \mathbf{R}_1, \ldots, \mathbf{R}_\lambda : \{0,1\}^n \to \{0,1\}^\ell$ be independent URFs, and let $\mathbf{M}_1(\cdot), \ldots, \mathbf{M}_\lambda(\cdot)$ be independent instances of $\mathbf{M}(\cdot)$, then*

$$\Delta_{q,\ldots,q}(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}), \mathbf{M}_1(\mathbf{R}_1) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}_\lambda)) \leq \frac{q^2 \cdot \lambda^2}{4} \cdot 2^{-n}$$

and

$$\Delta_{q,\ldots,q}(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}), \mathbf{M}_1(\mathbf{R}_1) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}_\lambda)) \leq \frac{q^2 \cdot \lambda^2}{4} \cdot 2^{-n}.$$

In other words, the URF $\mathbf{R}$ is η-self-independent under $\mathbf{M}(\cdot)$ and $\mathbf{M}'(\cdot)$ for $\eta(q, \lambda) \leq \frac{q^2\lambda^2}{2} \cdot 2^{-n}$.

Proof. We just prove the statement for $\mathbf{M}(\cdot)$, as the proof for $\mathbf{M}'(\cdot)$ is fully analogous. Let $Z_1, \ldots, Z_\lambda$ be the random offsets chosen by $\mathbf{Z}(\cdot)$ in $\mathbf{M}_1, \ldots, \mathbf{M}_\lambda$, respectively, and consider the MES $\mathcal{A} := A_0, A_1, \ldots$ such that A_i holds as long as within the first i queries there exists no two queries X_j, X_k, $j \neq k$ to distinct sub-systems $t, t' \in \{1, \ldots, \lambda\}$ such that $X_i \oplus Z_t = X_j \oplus Z_{t'}$. Clearly, as long as this does not hold, every query (regardless of the systems) returns an independent random output, i.e.,

$$(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R})) | \mathcal{A} \equiv \mathbf{M}_1(\mathbf{R}_1) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}_\lambda),$$

from which we infer by Lemma 1 (i) that for all distinguishers $\mathbf{D}$ issuing at most q queries to each of the λ subsystems we have

$$\Delta^{\mathbf{D}}(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}), \mathbf{M}_1(\mathbf{R}_1) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}_\lambda)) \leq \nu^{\mathbf{D}}(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}), \overline{A_{\lambda \cdot q}}).$$

Clearly, $\mathcal{A}$ is independent of the output values, and by Lemma 3 we can restrict ourselves to non-adaptive strategies for making $\mathcal{A}$ fail. For each valid sequence of $\lambda \cdot q$ queries and for each pair of

systems $t \neq t'$ there are at most q^2 pairs of possibly colliding queries $X_i \oplus Z_t$, $X_j \oplus Z_{t'}$, and the probability that they collide is exactly 2^{-n}. By the union bound we have

$$\nu^{\mathbf{D}}(\mathbf{M}_1(\mathbf{R}) \| \cdots \| \mathbf{M}_\lambda(\mathbf{R}), \overline{A_{\lambda \cdot q}}) \leq \binom{\lambda}{2} \cdot q^2 \cdot 2^{-n}. \qquad \square$$

However, a major advantage of Myers' original construction (which was unobserved so far) is that independent instances of the construction can be simulated even when only given access to $\mathbf{Z}(\mathbf{S})$ (with $\mathbf{S} \in \{\mathbf{F}_i, \mathbf{R}\}$). The relevance of this observation is due to the fact that the best advantage under $\mathbf{Z}(\cdot)$ can be significantly smaller than under direct access: Consider e.g. a good PRF which is modified to have the additional property of outputting the zero string when evaluated at some fixed known input, regardless of the key.

In order to apply Theorem 13, the corresponding construction $\mathbf{T}^{(i)}_{s_1,\ldots,s_{i-1}}(\cdot)$ chooses independent instances $\mathbf{F}_{i+1}, \ldots, \mathbf{F}_m$, $\mathbf{Z}_1(\cdot), \ldots, \mathbf{Z}_{i-1}(\cdot), \mathbf{Z}_{i+1}(\cdot), \ldots, \mathbf{Z}_m(\cdot)$, and a random n-bit string Z, and on input x queries $x \oplus Z$ to $\mathbf{Z}(S)$, obtaining $y \in \{0,1\}^\ell$, and outputs

$$y \oplus \bigoplus_{j=1}^{i-1} \mathbf{Z}_j(\mathbf{F}_j(s_j))(x) \oplus \bigoplus_{j=i+1}^{m} \mathbf{Z}_j(\mathbf{F}_j)(x),$$

where $\mathbf{Z}_j(\mathbf{F}_j)(x)$ is the result of invoking the system $\mathbf{Z}_j(\mathbf{F}_j)$ on input x.

Once again, condition (ii) is easily verified by the fact that access through $\mathbf{Z}(\cdot)$ can be re-randomized by simply adding a fresh random offset to all inputs. Thus, Theorem 13 yields the following strengthened version of the main result of [29].

Corollary 22. *For all $t, q, \gamma > 0$, and for independent instances $\mathbf{Z}_1(\cdot), \ldots, \mathbf{Z}_m(\cdot)$ of $\mathbf{Z}(\cdot)$,*

$$\Delta_{t,q}(\mathbf{Z}_1(\mathbf{F}_1) \oplus \cdots \oplus \mathbf{Z}_m(\mathbf{F}_m), \mathbf{R}) \leq \prod_{i=1}^{m} \Delta_{t'_i, q'_i}(\mathbf{Z}(\mathbf{F}_i), \mathbf{Z}(\mathbf{R})) + \tfrac{(m-1)q^2\lambda^2}{2^{n+1}} + \gamma,$$

where $t'_i := \lambda\big(t + \mathcal{O}(q \cdot \sum_{j \neq i} t_{F_j})\big)$ and $q'_i := \lambda \cdot q$ for all $i = 1, \ldots, m-1$, whereas $t'_m := t + \mathcal{O}(q \cdot \sum_{j=1}^{m-1} t_{F_j})$ and $q'_m := q$.

Acknowledgments

We thank Peter Gaži for helpful feedback. This research was partially supported by the Swiss National Science Foundation (SNF), project no. 200020-113700/1.

References

[1] M. Bellare, R. Impagliazzo, and M. Naor, "Does parallel repetition lower the error in computationally sound protocols?," in *FOCS '97: Proceedings of the 38th IEEE Annual Symposium on Foundations of Computer Science*, pp. 374–383, 1997.

[2] M. Bellare and P. Rogaway, "The security of triple encryption and a framework for code-based game-playing proofs," in *Advances in Cryptology — EUROCRYPT 2006*, vol. 4004 of *Lecture Notes in Computer Science*, pp. 409–426, 2006.

[3] R. Canetti, S. Halevi, and M. Steiner, "Hardness amplification of weakly verifiable puzzles," in *Theory of Cryptography — TCC 2005*, vol. 3378 of *Lecture Notes in Computer Science*, pp. 17–33, 2005.

[4] R. Canetti, R. L. Rivest, M. Sudan, L. Trevisan, S. P. Vadhan, and H. Wee, "Amplifying collision resistance: A complexity-theoretic treatment," in *Advances in Cryptology — CRYPTO 2007*, vol. 4622 of *Lecture Notes in Computer Science*, pp. 264–283, 2007.

[5] I. B. Damgård and J. B. Nielsen, "Expanding pseudorandom functions; or: From known-plaintext security to chosen-plaintext security," in *Advances in Cryptology — CRYPTO 2002*, vol. 2442 of *Lecture Notes in Computer Science*, pp. 449–464, 2002.

[6] Y. Dodis, R. Impagliazzo, R. Jaiswal, and V. Kabanets, "Security amplification for interactive cryptographic primitives," in *Theory of Cryptography — TCC 2009*, vol. 5444 of *Lecture Notes in Computer Science*, pp. 128–145, 2009.

[7] C. Dwork, M. Naor, and O. Reingold, "Immunizing encryption schemes from decryption errors," in *Advances in Cryptology — EUROCRYPT 2004*, vol. 3027 of *Lecture Notes in Computer Science*, pp. 342–360, 2004.

[8] S. Even and O. Goldreich, "On the power of cascade ciphers," *ACM Trans. Comput. Syst.*, vol. 3, no. 2, pp. 108–116, 1985.

[9] S. Even and Y. Mansour, "A construction of a cipher from a single pseudorandom permutation," *Journal of Cryptology*, vol. 10, no. 3, pp. 151–162, 1997.

[10] O. Goldreich, R. Impagliazzo, L. A. Levin, R. Venkatesan, and D. Zuckerman, "Security preserving amplification of hardness," in *FOCS '90: Proceedings of the 31st IEEE Annual Symposium on Foundations of Computer Science*, pp. 318–326, 1990.

[11] O. Goldreich, N. Nisan, and A. Wigderson, "On Yao's XOR-lemma," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 2, no. 50, 1995.

[12] I. Haitner, D. Harnik, and O. Reingold, "On the power of the randomized iterate," in *Advances in Cryptology — CRYPTO 2006*, vol. 4117 of *Lecture Notes in Computer Science*, pp. 22–40, 2006.

[13] S. Halevi and T. Rabin, "Degradation and amplification of computational hardness," in *Theory of Cryptography — TCC 2008*, vol. 4948 of *Lecture Notes in Computer Science*, pp. 626–643, 2008.

[14] J. Håstad, R. Impagliazzo, L. A. Levin, and M. Luby, "A pseudorandom generator from any one-way function," *SIAM Journal on Computing*, vol. 28, no. 4, pp. 1364–1396, 1999.

[15] A. Herzberg, "On tolerant cryptographic constructions," in *CT-RSA 2005*, vol. 3376 of *Lecture Notes in Computer Science*, pp. 172–190, 2005.

[16] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.

[17] T. Holenstein, "Key agreement from weak bit agreement," in *STOC '05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pp. 664–673, 2005.

[18] T. Holenstein and R. Renner, "One-way secret-key agreement and applications to circuit polarization and immunization of public-key encryption," in *Advances in Cryptology — CRYPTO 2005*, vol. 3621 of *Lecture Notes in Computer Science*, pp. 478–493, 2005.

[19] N. Hopper, D. Molnar, and D. Wagner, "From weak to strong watermarking," in *Theory of Cryptography — TCC 2007*, vol. 4392 of *Lecture Notes in Computer Science*, pp. 362–382, 2007.

[20] R. Impagliazzo, "Hard-core distributions for somewhat hard problems," in *FOCS '95: Proceedings of the 36th IEEE Annual Symposium on Foundations of Computer Science*, pp. 538–545, 1995.

[21] R. Impagliazzo, R. Jaiswal, and V. Kabanets, "Chernoff-type direct product theorems," in *Advances in Cryptology — CRYPTO 2007*, vol. 4622 of *Lecture Notes in Computer Science*, pp. 500–516, 2007.

[22] L. A. Levin, "One way functions and pseudorandom generators," *Combinatorica*, vol. 7, no. 4, pp. 357–363, 1987.

[23] M. Luby and C. Rackoff, "Pseudo-random permutation generators and cryptographic composition," in *STOC '86: Proceedings of the 18th Annual ACM Symposium on Theory of Computing*, pp. 356–363, 1986.

[24] U. Maurer, "Indistinguishability of random systems," in *Advances in Cryptology — EUROCRYPT 2002*, vol. 2332 of *Lecture Notes in Computer Science*, pp. 110–132, 2002.

[25] U. Maurer and J. L. Massey, "Cascade ciphers: The importance of being first," *Advances in Cryptology — CRYPTO '93*, vol. 6, no. 1, pp. 55–61, 1993.

[26] U. Maurer, K. Pietrzak, and R. Renner, "Indistinguishability amplification," in *Advances in Cryptology — CRYPTO 2007*, vol. 4622 of *Lecture Notes in Computer Science*, pp. 130–149, Aug. 2007.

[27] U. Maurer and J. Sjödin, "A fast and key-efficient reduction of chosen-ciphertext to known-plaintext security," in *Advances in Cryptology — EUROCRYPT 2007*, vol. 4515 of *Lecture Notes in Computer Science*, pp. 498–516, 2007.

[28] S. Myers, "On the development of block-ciphers and pseudo-random function generators using the composition and XOR operators." Master's thesis, University of Toronto, 1999.

[29] S. Myers, "Efficient amplification of the security of weak pseudo-random function generators," *Journal of Cryptology*, vol. 16, pp. 1–24, 2003.

[30] M. Naor and O. Reingold, "Synthesizers and their application to the parallel construction of pseudo-random functions," *Journal of Computer and System Sciences*, vol. 58, no. 2, pp. 336–375, 1999.

[31] R. Pass and M. Venkitasubramaniam, "An efficient parallel repetition theorem for Arthur-Merlin games," in *STOC '07: Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 420–429, 2007.

[32] K. Pietrzak and D. Wikström, "Parallel repetition of computationally sound protocols revisited," in *Theory of Cryptography — TCC 2007*, vol. 4392 of *Lecture Notes in Computer Science*, pp. 86–102, 2007.

[33] R. Shaltiel and E. Viola, "Hardness amplification proofs require majority," in *STOC '08: Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pp. 589–598, 2008.

[34] S. Vaudenay, "Provable security for block ciphers by decorrelation," in *STACS '98*, vol. 1373 of *Lecture Notes in Computer Science*, pp. 249–275, 1998.

[35] S. Vaudenay, "Adaptive-attack norm for decorrelation and super-pseudorandomness," in *Selected Areas in Cryptography — SAC '99*, vol. 1758 of *Lecture Notes in Computer Science*, pp. 49–61, 1999.

[36] J. Wullschleger, "Oblivious-transfer amplification," in *Advances in Cryptology — EUROCRYPT 2007*, vol. 4515 of *Lecture Notes in Computer Science*, pp. 555–572, 2007.

[37] A. C. Yao, "Theory and applications of trapdoor functions," in *FOCS '82: Proceedings of the 23rd IEEE Annual Symposium on Foundations of Computer Science*, pp. 80–91, 1982.

A Hoeffding's Inequality

The following well-known result from probability theory [16] is repeatedly used throughout this paper.

Lemma 23 (Hoeffding's Inequalities)**.** *Let $X_1, \ldots, X_\varphi$ be independent random variables with range $[0,1]$, and let $\overline{X} := \frac{1}{t}\sum_{i=1}^{\varphi} X_i$. Then, for all $\epsilon > 0$ we have*

$$\mathsf{P}[\overline{X} \geq E[\overline{X}] + \epsilon] \leq e^{-\varphi\epsilon^2} \text{ and } \mathsf{P}[\overline{X} \leq E[\overline{X}] - \epsilon] \leq e^{-\varphi\epsilon^2}.$$

In particular,

$$\mathsf{P}\left[\left|\overline{X} - \mathsf{E}[\overline{X}]\right| \geq \epsilon\right] \leq 2 \cdot e^{-\varphi\epsilon^2}.$$