Security Amplification for the Cascade of Arbitrarily Weak PRPs: Tight Bounds via the Interactive Hardcore Lemma

Stefano Tessaro

Department of Computer Science and Engineering
University of California, San Diego
9500 Gilman Drive
La Jolla, CA
stessaro@cs.ucsd.edu

Abstract. A pseudorandom permutation (PRP) is a (keyed) permutation which, under a random key, can only be distinguished from a uniformly chosen permutation with negligible advantage over random guessing by a computationally bounded distinguisher. This paper considers the natural weakening of the PRP notion, called an ε-PRP, where we only require the advantage to be bounded by some non-necessarily negligible quantity ε: It is a natural and fundamental question – which we consider in this paper and refer to as *security amplification* – to efficiently construct a PRP from an ε-PRP for any $\varepsilon < 1$.

The simplest approach is the *cascade* (i.e, sequential composition) of weak PRPs, but determining its security-amplifying properties is a long-standing open problem: To date, partial results are limited to constant-length cascades [Luby and Rackoff (STOC '86), Myers '99], to the case $\varepsilon < \frac{1}{2}$, or to alternative, less efficient, approaches tweaking the cascade [Maurer and Tessaro (CRYPTO 2009)]. This paper closes this gap: we prove that the cascade of m ε-PRPs is an $((m-(m-1)\varepsilon)\varepsilon^m + \nu)$-PRP, where ν is a negligible function. This implies security amplification for all $\varepsilon < 1 - \frac{1}{\mathsf{poly}}$, and the result extends to *two-sided* PRPs, i.e., to the case where the inverse of the given permutation is also queried. Furthermore, we show that the bound is essentially tight.

Our approach relies on the first hardcore lemma for computational indistinguishability of *interactive* systems, which is our main technical and conceptual contribution, and is of independent interest: For any two systems $\mathbf{S}$ and $\mathbf{T}$, whose state does not depend on the interaction, and which no efficient adversary can distinguish with advantage better than ε, we show that there exist events $\mathcal{A}$ and $\mathcal{B}$ on the choices of the respective states, occurring each with probability at least $1-\varepsilon$, such that $\mathbf{S}$ and $\mathbf{T}$ are computationally indistinguishable conditioned on these events. As a corollary of the lemma, security amplification for the cascade follows from the fact, which we also prove, that for (essentially) all $\varepsilon < 1 - \frac{1}{N}$ the cascade of *two* independent (non-uniform) random permutations on an N-element set, whose distributions have both min-entropy at least $\log(N!) - \log\left((1-\varepsilon)^{-1}\right)$, is (information theoretically) indistinguishable from a *uniform* random permutation.

Keywords: Foundations, Security Amplification, Computational Indistinguishability, Pseudorandomness, Hardcore Lemmas, Block Cipher Cascading.

1 Introduction

1.1 Motivation: Security Amplification of PRPs

The security of most block-cipher based cryptographic schemes relies on the *unproven* assumption that the block cipher is a so-called *pseudorandom permutation* (PRP), i.e., a keyed family of permutations $E = \{E_k\}_{k\in\mathcal{K}}$ on a set $\mathcal{X}$ such that no computationally bounded adversary (usually called a *distinguisher*) is able to decide correctly whether it is given access to $E_K : \mathcal{X} \to \mathcal{X}$ under a random secret key $K \in \mathcal{K}$ or to a uniformly chosen permutation $\mathcal{X} \to \mathcal{X}$, except with a negligible advantage over random guessing. Continuous progress in cryptanalysis casts however some doubt as to whether block-cipher designs such as the *Advanced Encryption Standard* (AES) are indeed secure PRPs. It is therefore a prudent approach, as well as a central question in theoretical cryptography, to investigate *weaker* assumptions on a block cipher which are sufficient to efficiently solve a certain cryptographic task at hand.

A natural weakening of the PRP assumption, considered in this paper, is to only assume that the best *distinguishing advantage* of a computationally bounded adversary is bounded by a quantity $\varepsilon < 1$, where ε does not need to be a negligible function, but may also be a constant or even a quantity moderately converging to one as a function of the security parameter.[1] We consequently call a primitive satisfying this assumption an ε-PRP: Clearly, AES is much more likely to be a 0.99-PRP, rather than a fully-secure PRP.

This paper considers the natural question of *security amplification of PRPs*, i.e., we ask for constructions transforming an ε-PRP into a fully secure one. Ideally, such a construction should amplify security for arbitrary $\varepsilon < 1$ and call the weaker block cipher as few times as possible (that is, $\omega(\log n) \cdot (\log(1/\varepsilon))^{-1}$ times for security parameter n). This is in the same spirit of the huge body of literature on security amplification, initiated by Yao [46] in the context of one-way functions, and extended to a number of other cryptographic primitives, including (but not limited to) regular OWFs and OWPs [14, 17], two-party protocols [1, 39, 42, 45, 18, 16, 19, 4], key agreement and public-key encryption [10, 21, 23], and collision-resistant hash functions [3].

1.2 Our Result: Security Amplification of PRPs by Cascading

The most natural approach to strengthening a weak PRP E is to consider the *cascade* of length m which outputs

$$E'_{k_1,\ldots,k_m}(x) := (E_{k_1} \circ \cdots \circ E_{k_m})(x)$$

on input x and for keys $k_1, \ldots, k_m$ (which are chosen independently), where $\circ$ denotes the (sequential) composition of permutations.

Despite its apparent simplicity, determining the security amplifying properties of the cascade has been a long standing open problem. On the one hand, Luby and Rackoff [25] and Myers [35] showed that the cascade of c ε-PRPs is a $((2-\varepsilon)^{c-1}\varepsilon^c + \nu)$-PRP for any *constant* c, where ν is a negligible additive term, but this result is not sufficient to infer that a sufficiently-long cascade yields a fully-secure PRP for a non-negligible ε. On the other hand, Maurer and Tessaro [33] showed that the cascade of *arbitrary* (polynomial) length m is a $(2^{m-1}\varepsilon^m + \nu)$-PRP, but this bound only implies security amplification for $\varepsilon < \frac{1}{2}$ and is clearly not tight in view of the superior result for the constant-length case of [25, 35]. This leaves open the question of determining the exact behavior of the cascade of length m.

[1] Alternatives are restricting the type of interaction allowed by a distinguisher (e.g., by limiting it to non-adaptive or random queries [38, 6, 31, 32]) or only requiring unpredictability of the block cipher output, rather than pseudorandomness (cf. e.g. [8, 9]).

Our Result on Cascades. This paper closes this gap by providing an exact characterization of the security amplification properties of the cascade: We prove that the cascade of m ε-PRPs (with domain $\mathcal{X}$) is security amplifying for essentially any $\varepsilon < 1 - \frac{1}{|\mathcal{X}|}$.[2] In particular, we show that it is a $((m-(m-1)\varepsilon)\varepsilon^m + \nu)$-PRP, and the result extends to *two-sided* ε-PRPs, where the inverse can also be queried. We additionally prove our bound to be essentially tight. This result arises from the application of new *generic* techniques in the context of security amplification, which we illustrate in the next section at an informal level and which are of independent interest. It is our belief that they can potentially be applied to a number other questions within the wider scope of complexity-theoretic cryptography.

Further Related Work. Other less efficient constructions of fully secure PRPs from ε-PRPs exist: Maurer and Tessaro [34] showed that XORing two independent keys at both ends of the cascade yields an $(\varepsilon^m + \nu)$-PRP. Moreover, similar results [36, 7, 33] are known for strengthening the security of pseudorandom *functions* (PRF), where one drops the permutation requirement, but all these constructions fall short of implementing a permutation when instantiated with a keyed permutation, and an extra step (e.g., using [26]) is required to restore the permutation property. While these results may be considered sufficient for many purposes, we feel that the security-amplification result of this paper is important for at least two main reasons: First, we show that similar amplification properties are achieved with better efficiency by the most natural construction. Second, settling the open question of determining the security amplification properties of the plain cascade leads to the development of a new set of generic techniques which promise to be applicable in other problems.

We also point out that cascades have also been studied in other contexts. A first line of research [11, 28, 2, 12] has been devoted to studying generic attacks against the cascade of block ciphers. Furthermore, security-amplifying properties of cascades have been previously studied in the information-theoretic setting [44, 29, 30, 13] (i.e., with indistinguishability with respect to *unbounded* distinguishers), where, to the best of our knowledge, all obtained bounds are *not* tight. Finally, a line of work showed that the cascade of *non-adaptively* secure PRPs / PRFs is *not* an adaptively secure PRF / PRPs [37, 40], unless key agreement does not exist [41]. In particular, the statement has been shown to hold with respect to computationally unbounded distinguishers [29, 30, 13].

1.3 Our General Paradigm: The Interactive Hardcore Lemma and High-Entropy Permutations

A fundamental property of any two finite random variables X and Y (taking values from the same range) is that it is always possible to define events $\mathcal{A}$ and $\mathcal{B}$ on them (by means of conditional probability distributions $\mathsf{P}_{\mathcal{A}|X}$ and $\mathsf{P}_{\mathcal{B}|Y}$) such that:

(i) X and Y are equally distributed conditioned on the respective events, i.e., $\mathsf{P}_{X|\mathcal{A}} = \mathsf{P}_{Y|\mathcal{B}}$,
(ii) $\mathsf{P}[\mathcal{A}] = \mathsf{P}[\mathcal{B}] = 1 - d(X,Y)$, where $d(X,Y)$ is the so called *statistical distance*, which equals the best advantage of a computationally *unbounded* distinguisher in distinguishing X and Y.

A computational version of this statement is due to Maurer and Tessaro [34], and was used to prove security amplification results for PRGs. In this paper, we take this approach one step further by presenting a *computational* version of the above statement for discrete *interactive* systems.

CC-Stateless Systems. More specifically, we consider the general class of interactive systems called *convex-combination stateless* (or simply *cc-stateless*) [33], which includes a large number of cryptographic systems. These systems have the property that the answer of each query can be seen as only depending

[2] This restriction is necessary, as a permutation family with a fixed point (independent of the key value) is *at best* a $(1 - \frac{1}{|\mathcal{X}|})$-PRP, and the cascade obviously preserves such a fixed point. However, since *efficient* security amplification cannot be achieved for such high ε (m would be super-polynomial even assuming optimal security amplification), the restriction is irrelevant.

on the input of this query and on an *initial state*, but does not depend on previous queries and their answers. A simple example of a cc-stateless system is the system implementing a permutation E_K for a keyed family of permutations $\{E_k\}_{k\in\mathcal{K}}$ and a uniform random key $K \in \mathcal{K}$. A further example is a *uniform random permutation* (URP) $\mathbf{P} : \mathcal{X} \to \mathcal{X}$, a system choosing a permutation $P : \mathcal{X} \to \mathcal{X}$ uniformly at random, and answering each query x as $P(x)$. Moreover, a randomized encryption scheme where each encryption depends on the random key and some fresh randomness is also cc-stateless.

We stress that being cc-stateless is a property of the *input-output* behavior of a system, and not of its actual implementation: An implementation using such an initial state may be inefficient (e.g., due to the initial state being very large), but at the same time an efficient implementation of a cc-stateless system may indeed be fully stateful. For example, an efficient implementation of a URP does keep an interaction-dependent state and employs lazy sampling, returning for each new query a uniformly distributed value among those not returned yet.

The Hardcore Lemma Our main technical tool is the *Hardcore Lemma (HCL) for computational indistinguishability* (Theorem 2): Informally, it states that if all computationally bounded distinguishers only achieve advantage at most ε in distinguishing two cc-stateless systems $\mathbf{S}$ and $\mathbf{T}$, then there exist events $\mathcal{A}$ and $\mathcal{B}$, defined on the respective initial states of (the cc-stateless representations of) $\mathbf{S}$ and $\mathbf{T}$, such that the following holds:

(i) The systems $\mathbf{S}$ and $\mathbf{T}$ are computationally indistinguishable conditioned on the respective events $\mathcal{A}$ and $\mathcal{B}$.
(ii) Both events occur with probability at least $1 - \varepsilon$.

In addition, applications of the HCL typically require the ability to efficiently *simulate* $\mathbf{S}$ and $\mathbf{T}$ under the assumption that the associated events $\mathcal{A}$ and $\mathcal{B}$ occur (or do not occur), possibly with the help of some short (but not necessarily efficiently-samplable[3]) advice. In general, it is unclear whether this is possible given any two events satisfying (i) and (ii), even if both systems are efficiently implementable.

As an illustrative example, let $\mathbf{S} = E_K$ and $\mathbf{T} = \mathbf{P}$, where $K \in \mathcal{K}$ is uniformly distributed, $E = \{E_k\}_{k\in\mathcal{K}}$ an efficiently-computable family of permutations, and $\mathbf{P}$ is a URP, where all permutations are on the n-bit strings. If E is an ε-PRP, the HCL yields an event $\mathcal{A}$ defined on K and an event $\mathcal{B}$ defined on a uniformly chosen permutation table P, both occurring with probability at least $1 - \varepsilon$, such that $E_{K'}$ (for K' sampled from $\mathsf{P}_{K|\mathcal{A}}$) and a system $\mathbf{P}'$ (implementing a permutation table P' sampled from $\mathsf{P}_{P|\mathcal{B}}$) are computationally indistinguishable. While $E_{K'}$ is efficiently implementable given K', a representation of P' requires $2^{\Theta(n)}$ bits, and it is unclear how to define a short advice (i.e., with length $\mathsf{poly}(n)$) that can be used to efficiently simulate $\mathbf{P}'$. To overcome this issue, we will show that one can *always* find events $\mathcal{A}$ and $\mathcal{B}$ with the property that such advice exists as long as $\mathbf{S}$ and $\mathbf{T}$ are efficiently implementable. This will be the major challenge in proving the HCL for the interactive setting.

The core of our proof is a tight generalization (Theorem 1) of Impagliazzo's HCL [24] to the setting of guessing a random bit given access to some interactive system whose behavior is correlated with the bit value.

Cascade of Permutations with High Min-Entropy. We briefly illustrate how the HCL is used to prove our bounds for the cascade of ε-PRPs. The main observation is that P' as above has *min-entropy* at least

$$\mathsf{H}_\infty(P') = \log\left(\min_\pi \frac{1}{\mathsf{P}\left[P' = \pi\right]}\right) = \log\left(\min_\pi \frac{\mathsf{P}\left[\mathcal{B}\right]}{\mathsf{P}\left[P = \pi\right] \cdot \mathsf{P}\left[\mathcal{B} \,\middle|\, P = \pi\right]}\right) \geq \log\left(2^n!\right) - \log\left(\frac{1}{1-\varepsilon}\right),$$

i.e., at most $\log\left((1-\varepsilon)^{-1}\right)$ away from the maximal achievable min-entropy. Although this gap potentially makes $\mathbf{P}'$ easily distinguishable from a URP, we prove (Theorem 3) that the cascade of (at least)

[3] Our main result is in the non-uniform setting, thus efficient samplability is not a requirement.

two such permutations is indistinguishable from a URP for *computationally unbounded* distinguishers making at most an exponential number of queries and even when allowing queries to the inverse. (The proof uses techniques from the random systems framework [27], and is of independent interest.)

The main security amplification result (Theorem 4) roughly follows from the observation that by the above at least two (independent) permutations E_{K_i} and E_{K_j} (for $i \neq j$) in the cascade $E_{K_1} \circ \cdots \circ E_{K_m}$ (for independent keys $K_1, \ldots, K_m$) are computationally indistinguishable from $\mathbf{P}'$, except with probability $\varepsilon^m + m(1-\varepsilon)\varepsilon^{m-1}$, and in this case the cascade is computationally indistinguishable from a URP by Theorem 3. The final bound follows from a more fine-grained analysis.

Uniform vs. Non-Uniform Proofs. The results of this paper are formulated in a concrete, non-uniform computational model. This simplifies the presentation considerably and helps conveying the main ideas. Appendix F highlights the changes required in order to obtain uniform statements and proofs.

2 Preliminaries

Calligraphic letters $\mathcal{X}, \mathcal{Y}, \ldots$ denote sets and events, upper-case letters $X, Y, \ldots$ random variables (with expected values $\mathsf{E}[X], \mathsf{E}[Y], \ldots$), and lower-case letters $x, y, \ldots$ the values they take. Moreover, $\mathsf{P}[\mathcal{A}]$ is the probability of an event $\mathcal{A}$ (we denote as $\overline{\mathcal{A}}$ its complement) and 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,[4] and $x \xleftarrow{\$} \mathsf{P}_X$ is the action of sampling a value x with distribution P_X. (We use $x \xleftarrow{\$} \mathcal{S}$ to denote the special case where x is drawn uniformly from a finite set $\mathcal{S}$.) The *statistical distance* $d(X, Y)$ (or $d(\mathsf{P}_X, \mathsf{P}_Y)$) of X and Y (both with range $\mathcal{S}$) is defined as $d(X, Y) := \frac{1}{2} \sum_{x \in \mathcal{S}} |\mathsf{P}_X(x) - \mathsf{P}_Y(x)| = \sum_{x : \mathsf{P}_X(x) \geq \mathsf{P}_Y(x)} (\mathsf{P}_X(x) - \mathsf{P}_Y(x))$. Also, recall that a function is *negligible* if it vanishes faster than the inverse of any polynomial.

Computational Model. We consider *interactive* randomized stateful algorithms in some a-priori fixed RAM model of computation. Such an algorithm keeps a state (consisting, say, of the contents of the memory space it employs), and answers each query depending on the 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. We say that 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, of s, 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 $t_A(q) := t_A(q, 0)$. Furthermore, $s_A : \mathbb{N} \to \mathbb{N}$ is the *space complexity* of A, where $s_A(q)$ is the worst-case amount of memory used by A when answering any q queries.

Systems and Distinguishers. This paper considers abstract discrete interactive systems [27], denoted by bold-face letters $\mathbf{S}, \mathbf{T}, \ldots$, taking as inputs queries $X_1, X_2, \ldots$ and returning outputs $Y_1, Y_2, \ldots$. Such systems may be implemented by an interactive algorithm A (in which case we sometimes write A as a placeholder for the system it implements to explicit this fact), but may also arise from an arbitrary random process. The *input-output behavior* of the system $\mathbf{S}$ is fully described by the (infinite) family of conditional probability distributions $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}$ (for $i \geq 1$) of the i-th output Y_i given the first i queries $X^i = [X_1, \ldots, X_i]$, and the first $i-1$ outputs $Y^{i-1} = [Y_1, \ldots, Y_{i-1}]$. In general, every statement that involves a system $\mathbf{S}$ holds for any realization of the system $\mathbf{S}$, i.e., it only depends on its input-output behavior. In particular, we say that two systems $\mathbf{S}$ and $\mathbf{T}$ are *equivalent*, denoted $\mathbf{S} \equiv \mathbf{T}$, if they have the same input-output behavior, i.e., $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}} = \mathsf{p}^{\mathbf{T}}_{Y_i|X^iY^{i-1}}$ for all $i \geq 1$. Moreover, we say that an algorithm A *implements* the system $\mathbf{S}$ if $A \equiv \mathbf{S}$.

[4] 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.

A *distinguisher* $\mathbf{D}$ is a special type of system which interacts with another system $\mathbf{S}$ by means of q queries and outputs a decision bit $\mathbf{D}(\mathbf{S})$ depending on their outputs: Its *advantage* in distinguishing systems $\mathbf{S}$ and $\mathbf{T}$ is

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

Moreover, $\Delta_q(\mathbf{S},\mathbf{T})$ is the best distinguishing advantage $\Delta^{\mathbf{D}}(\mathbf{S},\mathbf{T})$ over *all* q-query $\mathbf{D}$, whereas $\Delta_{t,q}(\mathbf{S},\mathbf{T})$ is used when the maximization is restricted to distinguishers implemented by an algorithm with time complexity t.

Stateless Systems. A system $\mathbf{S}$ is called *stateless* if the i-th answer Y_i only depends on the i-th query X_i, i.e., there exists $\mathsf{p}^{\mathbf{S}}_{Y|X}$ such that $\mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}}(y_i,x^i,y^{i-1}) = \mathsf{p}^{\mathbf{S}}_{Y|X}(y_i,x_i)$ for all $i\geq 1$, $x^i=[x_1,\ldots,x_i]$, and $y^i=[y_1,\ldots,y_i]$. Furthermore, $\mathbf{S}$ is *convex-combination-stateless* (or simply *cc-stateless*) [33] if there exists a system $\mathbf{T}(\cdot)$ accessing a random variable S (called the *initial state*) such that $\mathbf{S}\equiv\mathbf{T}(S)$ and $\mathbf{T}(s)$ is stateless for all values s taken by S. To save on notation, we usually write $\mathbf{S}(\cdot)$ instead of $\mathbf{T}(\cdot)$, but we stress that $\mathbf{S}(\cdot)$ and $\mathbf{S}$ are different objects, despite their notational similarity. We refer to $\mathbf{S}(S)$ as the *cc-stateless representation* of $\mathbf{S}$.

It is crucial to remark that being cc-stateless is a property of the *input-output behavior* of a system: Its (efficient) implementation may well be stateful, and its cc-stateless representation may be completely inefficient (e.g., because the description of the initial state is even too large to be processed by an efficient algorithm).

Random Functions and Permutations. A system $\mathbf{F}$ taking inputs from a set $\mathcal{X}$ and returning outputs in $\mathcal{Y}$ is a *random function* $\mathcal{X}\to\mathcal{Y}$ if for any two equal queries $X_i=X_j$ we have $Y_i=Y_j$ for the respective answers. Furthermore, if $\mathcal{X}=\mathcal{Y}$, it is called a *random permutation* if $X_i\neq X_j$ also implies $Y_i\neq Y_j$. Typical (cc-stateless) examples are *uniform random function* (URF) $\mathbf{R}:\mathcal{X}\to\mathcal{Y}$, which answers according to a uniformly chosen function $\mathcal{X}\to\mathcal{Y}$, a *uniform random permutation* (URP) $\mathbf{P}:\mathcal{X}\to\mathcal{X}$, implementing a uniformly chosen permutation $\mathcal{X}\to\mathcal{X}$, or E_K for a permutation family $\{E_k\}_{k\in\mathcal{K}}$ and a random $K\in\mathcal{K}$.

The initial state of a cc-stateless random function $\mathbf{F}$ can always be seen without loss of generality as a (randomly chosen) *function table* F according to which $\mathbf{F}$ answers its queries, and usually write $\mathbf{F}(x)$ instead of $F(x)$. In particular, the *inverse* $\mathbf{Q}^{-1}$ of a cc-stateless permutation $\mathbf{Q}$ is well-defined, and $\langle\mathbf{Q}\rangle$ is the *two-sided* random permutation which allows both forward queries $(x,+)$ returning $\mathbf{Q}(x)$ as well as *backward queries* $(y,-)$ returning $\mathbf{Q}^{-1}(y)$. The *cascade* $\mathbf{Q}'\rhd\mathbf{Q}''$ of two random permutations is the system which on input x returns $\mathbf{Q}''(\mathbf{Q}'(x))$, i.e., it implements the composition of the associated permutation tables. (This extends naturally to longer cascades.) Note in particular that for any cascade we have $\mathbf{Q}_1\rhd\cdots\rhd\mathbf{Q}_m\equiv\mathbf{P}$ whenever there exists i such that $\mathbf{Q}_i\equiv\mathbf{P}$ for a URP $\mathbf{P}$. Moreover, we let $\langle\mathbf{Q}'\rangle\rhd\langle\mathbf{Q}''\rangle := \langle\mathbf{Q}'\rhd\mathbf{Q}''\rangle$.

An efficiently implementable family of permutations $E=\{E_k\}_{k\in\mathcal{K}}$ with domain $\mathcal{X}$ and indexed by keys $k\in\mathcal{K}$ is an *ε-pseudorandom permutation* (ε-PRP) if $\Delta_{t,q}(E_K,\mathbf{P})\leq\varepsilon$ for all polynomially bounded t and q, a uniform $K\in\mathcal{K}$, and a URP $\mathbf{P}$. It is a *two-sided* ε-PRP if $\langle E_K\rangle$ is efficiently implementable and $\Delta_{t,q}(\langle E_K\rangle,\langle\mathbf{P}\rangle)\leq\varepsilon$ for all polynomially bounded t and q.

3 Hardcore Lemmas for Interactive Systems

3.1 System-Bit Pairs, Measures, and State Samplers

We consider the general setting of *system-bit pairs* [33] $(\mathbf{S},B)$ consisting of a bit B (with an associated probability distribution P_B), and a system $\mathbf{S}=\mathbf{S}(B)$ whose behavior depends on the outcome of the bit B. A system-bit pair $(\mathbf{S},B)$ is to be interpreted as a system which parallely composes $\mathbf{S}$ and a correlated bit B (which is *initially* chosen, before any interaction with $\mathbf{S}$ has taken place). The notion

of a cc-stateless system-bit pair $(\mathbf{S}(S), B(S))$ is obtained naturally. Also, an implementation $A_{(\mathbf{S},B)}$ of a system-bit pair $(\mathbf{S}, B)$ is without loss of generality an algorithm which outputs the bit B and then simulates the system $\mathbf{S}(B)$.

We associate with every system-bit pair $(\mathbf{S}, B)$ a game where an *adversary* $\mathbf{A}$ interacts with $\mathbf{S}(B)$ and outputs a binary guess $\mathbf{A}(\mathbf{S}(B)) \in \{0,1\}$ for B: Its *guessing advantage* is defined as the quantity

$$\mathrm{Guess}^{\mathbf{A}}(B \,|\, \mathbf{S}) := 2 \cdot \mathsf{P}[\mathbf{A}(\mathbf{S}(B)) = B] - 1 \in [-1, 1].$$

If $\mathrm{Guess}^{\mathbf{A}}(B \,|\, \mathbf{S}) = 1$, then $\mathbf{A}$ always guesses B correctly, whereas $\mathrm{Guess}^{\mathbf{A}}(B \,|\, \mathbf{S}) = -1$ means that $\mathbf{A}$ is always wrong (though flipping $\mathbf{A}$'s output bit yields an adversary which is always correct.) The shorthand $\mathrm{Guess}_{t,q}(B \,|\, \mathbf{S})$ denotes the best guessing advantage taken over all adversaries with time complexity t and issuing at most q queries to $\mathbf{S}$.

Example 1. An example is the (cc-stateless) system-bit pair $(\mathbf{R}, B)$ for a URF $\mathbf{R} : \mathcal{X} \to \{0,1\}$ and $B := \oplus_{x \in \mathcal{X}} \mathbf{R}(x)$ is the parity of its function table. It is easy to see that $\mathrm{Guess}_q(B \,|\, \mathbf{R}) = 0$ for all $q < |\mathcal{X}|$.

Example 2. If $(\mathbf{F}, B)$ is such that B is uniform, and $\mathbf{F}$ behaves as a system $\mathbf{S}$ if $B = 0$, and as another system $\mathbf{T}$ if $B = 1$, then $\mathrm{Guess}^{\mathbf{D}}(B \,|\, \mathbf{F}) = \Delta^{\mathbf{D}}(\mathbf{S}, \mathbf{T})$ for all $\mathbf{D}$ by a standard argument. Note that if both $\mathbf{S}$ and $\mathbf{T}$ are cc-stateless, then $(\mathbf{F}, B)$ is also cc-stateless.

Measures. A *measure* $\mathcal{M}$ for a cc-stateless system $\mathbf{S} \equiv \mathbf{S}(S)$, where $S \in \mathcal{S}$ is the initial state, is a mapping $\mathcal{M} : \mathcal{S} \to [0,1]$. Its *density* is defined as $\mu(\mathcal{M}) := \mathsf{E}[\mathcal{M}(S)] = \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathcal{M}(s)$. The measure $\mathcal{M}$ is naturally associated with a probability distribution $\mathsf{P}_{\mathcal{M}}$ on $\mathcal{S}$ such that $\mathsf{P}_{\mathcal{M}}(s) := \mathsf{P}_S(s) \cdot \mathcal{M}(s) \cdot \mu(\mathcal{M})^{-1}$ for all $s \in \mathcal{S}$. Also, we define the complement of a measure $\mathcal{M}$ as the measure $\overline{\mathcal{M}}$ such that $\overline{\mathcal{M}}(s) := 1 - \mathcal{M}(s)$ for all $s \in \mathcal{S}$. We repeatedly abuse notation writing $S \xleftarrow{\$} \mathcal{M}$ instead of $S \xleftarrow{\$} \mathsf{P}_{\mathcal{M}}$.

Traditionally, measures are seen as "fuzzy" subsets of $\mathcal{S}$. Alternatively, it is convenient to think of $\mathcal{M}$ in terms of a conditional probability distribution $\mathsf{P}_{\mathcal{A}|S}$ with $\mathsf{P}_{\mathcal{A}|S}(s) := \mathcal{M}(s)$ which adjoins the event $\mathcal{A}$ on the choice of S: In particular, $\mu(\mathcal{M}) = \mathsf{P}[\mathcal{A}]$, $\mathsf{P}_{\mathcal{M}} = \mathsf{P}_{S|\mathcal{A}}$, and $\mathsf{P}_{\overline{\mathcal{M}}} = \mathsf{P}_{S|\overline{\mathcal{A}}}$. In the following, we stick to measures for stating and proving our hardcore lemmas, while an event-based view will be useful when exercising these results.

State Samplers. Ideally, the hardcore lemma for a cc-stateless system-bit pair $(\mathbf{S}, B) \equiv (\mathbf{S}(S), B(S))$ (for initial state $S \in \mathcal{S}$) states that if $\mathrm{Guess}_{t,q}(B \,|\, \mathbf{S}) \leq \varepsilon$, then there exists a measure $\mathcal{M}$ on $\mathcal{S}$ such that (i) $\mu(\mathcal{M}) \geq 1 - \varepsilon$ and (ii) $\mathrm{Guess}_{t',q'}(B(S') \,|\, \mathbf{S}(S')) \approx 0$ for $S' \xleftarrow{\$} \mathcal{M}$ and t', q' as close as possible to t, q. Whenever $\mathbf{S}(S)$ is a random variable, this is equivalent to (a tight) version of Impagliazzo's Hardcore Lemma [22]. However, applications of the hardcore lemma (as the one we give later in this paper) require the ability, possibly given some short advice, to efficiently simulate $(\mathbf{S}(S'), B(S'))$ for $S' \xleftarrow{\$} \mathcal{M}$ or $(\mathbf{S}(S''), B(S''))$ for $S'' \xleftarrow{\$} \overline{\mathcal{M}}$.[5] While in the context of random variables the advice is generally a sample of S' itself, this approach fails in the setting of interactive systems: Recall that the representation $(\mathbf{S}(S), B(S))$ is possibly only a thought experiment, and the description of S' may be of exponential size, or no efficient algorithm implementing $(\mathbf{S}, B)$ from S' exists, even if the system-bit pair itself is efficiently implementable.

To formalize the concept of an advice distribution, we introduce the notion of a state sampler for a cc-stateless system (such as e.g. a system-bit pair).

Definition 1 (State Samplers). *Let $\mathbf{S} \equiv \mathbf{S}(S)$ be a cc-stateless system with implementation $A_{\mathbf{S}}$ and $S \in \mathcal{S}$, let $\zeta_1, \zeta_2 \in [0,1]$, and let $\mathcal{M} : \mathcal{S} \to [0,1]$ be a measure for $\mathbf{S}$. A (ζ_1, ζ_2)-(state) sampler O for $\mathcal{M}$ and $A_{\mathbf{S}}$ with length ℓ is a random process O such that:*

[5] Formally, one actually needs to prove that $\mathrm{Guess}_{t',q'}(B(S') \,|\, \mathbf{S}(S')) \approx 0$ holds even given access to the advice: While this is implicit in the non-uniform setting (every adversary with advice can be turned in an equally good one without advice), the proof is more challenging in the uniform setting, cf. e.g. [21] and Appendix F.

(i) O *always returns a pair* (σ, z) *with* σ *being a valid state for* $A_{\mathbf{S}}$ *with* $|\sigma| \leq \ell$ *and* $z \in [0,1]$*;*

(ii) For $(\Sigma, Z) \xleftarrow{\$} \mathsf{O}$*, we have*[6]

$$(A_{\mathbf{S}}[\Sigma], Z) \equiv (\mathbf{S}(S), Z'(S)),$$

where $Z'(S) \in [0,1]$ *is a random variable (which only depends on* S*) which differs from* $\mathcal{M}(S)$ *by at most* ζ_1*, except with probability* ζ_2*, for any value taken by* S*.*

Example 3. For all implementations $A_{\mathbf{S}}$ of $\mathbf{S}$, the all-one measure (i.e., $\mathsf{P}_{\mathcal{M}} = \mathsf{P}_S$) admits an error-less sampler O which returns the initial (void) state for $A_{\mathbf{S}}$ and $z = 1$. We will see further examples of state samplers in the proof of Theorem 1 below.

Note that O is *not* required to be efficiently implementable. State samplers allow for efficient simulation of $\mathbf{S}(S')$ for $S' \xleftarrow{\$} \mathcal{M}$: Given the output (Σ, Z) sampled from a (ζ_1, ζ_2)-sampler O, we flip a coin B with $\mathsf{P}_B(1) = Z$: Consider the distribution $\mathsf{P}_{\Sigma|B=1}$ of Σ conditioned on the outcome $B = 1$. If $\zeta_1 = \zeta_2 = 0$, it is not hard to verify that $A_{\mathbf{S}}[\Sigma'] \equiv \mathbf{S}(S')$ for $\Sigma' \xleftarrow{\$} \mathsf{P}_{\Sigma|B=1}$ and $S' \xleftarrow{\$} \mathcal{M}$. This is because, by definition, we have $(A_{\mathbf{S}}[\Sigma], Z, B) \equiv (\mathbf{S}(S), \mathcal{M}(S), B')$, where B' is a bit which is 1 with probability $\mathcal{M}(S)$, and thus in particular $A_{\mathbf{S}}[\Sigma'] \equiv \mathbf{S}(S')$ where $S' \xleftarrow{\$} \mathsf{P}_{S|B'=1}$. However, since $\mathsf{P}_{B'|S}(1, s) := \mathcal{M}(s)$ and $\mathsf{P}_{B'}(1) := \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathcal{M}(s) = \mu(\mathcal{M})$,

$$\mathsf{P}_{S|B'}(s, 1) = \mathcal{M}(s) \cdot \mathsf{P}_S(s) \cdot \mu(\mathcal{M})^{-1} = \mathsf{P}_{\mathcal{M}}(s).$$

Of course, one can similarly simulate for $S'' \xleftarrow{\$} \mathsf{P}_{\overline{\mathcal{M}}}$, as we obtain a corresponding sampler by just replacing z by $1 - z$ in the output (σ, z). This approach can be extended to non-zero errors ζ_1 and ζ_2.

3.2 The Hardcore Lemma for System-Bit Pairs

In the following, for understood parameters γ, ε, ζ_1, and ζ_2, we define

$$\varphi_{\mathsf{hc}} := \frac{6400}{\gamma^2(1-\varepsilon)^4} \cdot \ln\left(\frac{160}{\gamma(1-\varepsilon)^3}\right) \text{ and } \psi_{\mathsf{hc}} := \frac{200}{\gamma^2(1-\varepsilon)^4\zeta_1^2} \cdot \ln\left(\frac{2}{\zeta_2}\right).$$

We now state the HCL for cc-stateless system-bit pairs. Even though we apply the result only in a more restricted setting, we prove a more general statement for arbitrary cc-stateless system-bit pairs.

Theorem 1 (HCL for System-Bit Pairs). *Let* $(\mathbf{S}, B) \equiv (\mathbf{S}(S), B(S))$ *be a cc-stateless system-bit pair admitting an implementation* $A_{(\mathbf{S},B)}$ *with space complexity* $s_{A_{(\mathbf{S},B)}}$*. Furthermore, for some integers* $t, q > 0$ *and some* $\varepsilon \in [0,1)$*,*

$$\mathrm{Guess}_{t,q}(B \,|\, \mathbf{S}) \leq \varepsilon.$$

Then, for all $0 < \zeta_1, \zeta_2 < 1$ *and all* $0 < \gamma \leq \frac{1}{2}$*, there exists a measure* $\mathcal{M}$ *for* $(\mathbf{S}, B)$ *with* $\mu(\mathcal{M}) \geq 1 - \varepsilon$ *such that the following two properties are satisfied:*

(i) For $S' \xleftarrow{\$} \mathcal{M}$*,* $t' := t/\varphi_{\mathsf{hc}}$*, and* $q' := q/\varphi_{\mathsf{hc}}$*,*

$$\mathrm{Guess}_{t',q'}(B(S') \,|\, \mathbf{S}(S')) \leq \gamma.$$

(ii) There exists a (ζ_1, ζ_2)*-sampler for* $\mathcal{M}$ *and* $A_{(\mathbf{S},B)}$ *with length* $s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}} \cdot q')$*. Moreover, if* $(\mathbf{S}(s), B(s))$ *is deterministic for all* s*, then there also exists a* $(0,0)$*-sampler for* $\mathcal{M}$ *and* $A_{(\mathbf{S},B)}$ *with length* $s_{A_{(\mathbf{S},B)}}((7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1) \cdot q')$*.*

In the remainder of this section, we outline the proof intuition, while the full proof is postponed to Appendix C.1. We refer the interested reader to Appendix B for an example application in the most general setting where only a sampler with non-zero error is given by the theorem.

[6] That is, we consider the parallel composition of a system (either $A_{\mathbf{S}}[\Sigma]$ or $\mathbf{S}(S)$) and a correlated $[0,1]$-valued random variable.

PROOF OUTLINE. The proof is by contradiction: We assume that for all measures $\mathcal{M}$ with $\mu(\mathcal{M}) \geq 1-\varepsilon$ admitting a (ζ_1, ζ_2)-sampler as in (ii), there exists an adversary $\mathbf{A}$ with time complexity t' and query complexity q' such that $\text{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma$ for $S' \xleftarrow{\$} \mathcal{M}$. The core of the proof consists of proving that, under this assumption, there exists a sufficiently small family of adversaries $\mathcal{A}$ (more specifically, $|\mathcal{A}| = 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$) such that either (A) $\alpha(S) > \gamma$ holds with probability higher than $1 - \frac{1-\varepsilon}{4}$ over the choice of S, where $\alpha(s) := \mathsf{E}\left[\text{Guess}^{\mathbf{A}'}(B(s) \,|\, \mathbf{S}(s))\right]$ for all s, where $\mathbf{A}' \xleftarrow{\$} \mathcal{A}$, or (B) $\mathsf{E}\left[\alpha(S')\right] > \Theta\left((1-\varepsilon)^2\gamma\right)$ *for all* measures $\mathcal{M}$ with density $1-\varepsilon$ and $S' \xleftarrow{\$} \mathcal{M}$.

In Case (A), a simple majority-voting based strategy yields a good adversary breaking the assumed hardness of $(\mathbf{S}, B)$, whereas in Case (B) such an adversary can be built from $\mathcal{A}$ using techniques similar to the case of random variables [24, 21]. Both adversaries heavily rely on the cc-stateless property of $(\mathbf{S}, B)$.

The existence of an appropriate family $\mathcal{A}$ is shown by means of an iterative process, similar to the ones used by Impagliazzo [24] and by Holenstein [21]: We associate with each family $\mathcal{A}$ and $\tau \in \mathbb{N}$ a measure $\mathcal{M}_{\mathcal{A},\tau}$ such that elements for which $\mathcal{A}$ is worst, i.e., $N_{\mathcal{A}}(s) \leq \tau$, are given high weight (i.e. $\mathcal{M}_{\mathcal{A},\tau}(s) = 1$), whereas elements for which $\mathcal{A}$ performs well, i.e., $N_{\mathcal{A}}(s) \geq \tau + \frac{1}{\gamma(1-\varepsilon)}$, are not chosen ($\mathcal{M}_{\mathcal{A},\tau}(s) = 0$). An intermediate measure value is assigned to states not falling into one of these two categories. In particular, $\mathcal{M}_{\emptyset,0}$ is the all-one measure (i.e., $\mathsf{P}_{\mathcal{M}}$ equals the state distribution P_S), which has density $1 \geq 1-\varepsilon$. A crucial property, which we show, is that $\mathcal{M}_{\mathcal{A},\tau}$ *always* admits an (ζ_1, ζ_2)-state sampler for all $\mathcal{A}$ and τ. We then consider the following iterative process: It starts with $\mathcal{A} := \emptyset$ and then, at each round, it possibly increases τ to ensure that $\mu(\mathcal{M}_{\mathcal{A},\tau}) \geq 1-\varepsilon$ and then uses the assumption of the HCL being wrong to find an adversary achieving advantage larger than γ for $\mathcal{M}_{\mathcal{A},\tau}$, and adds it $\mathcal{A}$. We prove that within $7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ iterations $\mathcal{A}$ satisfies (A) or (B).

Remark 1. A natural question is whether the HCL can be extended to arbitrary system-bit pairs, where the measure is defined on the randomness of the implementation of the system-bit pair, regardless of the system having a cc-stateless representation. Yet, techniques similar to the ones used in counter-examples to soundness amplification for interactive arguments via parallel repetition [1, 42] yield (non cc-stateless) efficiently implementable system-bit pairs for which, given multiple independent instances of the system-bit pair, the probability of guessing all of the bits given access to all of the associated systems in parallel does not decrease with the number of instances. If such a a general HCL were true, then it is not hard to prove that the guessing probability *would* decrease exponentially in the number of instances. We postpone a discussion to a later version of this paper.

3.3 The Hardcore Lemma for Computational Indistinguishability

This section presents the hardcore lemma for computational indistinguishability of *interactive* systems. In particular, this result generalizes the statement for random variables previously shown in [34].

Theorem 2 (HCL for Computational Indistinguishability). *Let $\mathbf{S} \equiv \mathbf{S}(S)$ and $\mathbf{T} \equiv \mathbf{T}(T)$ be cc-stateless systems, with respective implementations $A_{\mathbf{S}}$ (with space complexity $s_{A_{\mathbf{S}}}$) and $A_{\mathbf{T}}$ (with space complexity $s_{A_{\mathbf{T}}}$). Furthermore, for some integers $t, q > 0$ and some $\varepsilon \in [0, 1)$,*

$$\Delta_{t,q}(\mathbf{S}, \mathbf{T}) \leq \varepsilon.$$

Then, for all $0 < \zeta_1, \zeta_2 < 1$ and all $0 < \gamma \leq \frac{1}{2}$, there exist measures $\mathcal{M}_{\mathbf{S}}$ and $\mathcal{M}_{\mathbf{T}}$ such that $\mu(\mathcal{M}_{\mathbf{S}}) \geq 1-\varepsilon$ and $\mu(\mathcal{M}_{\mathbf{T}}) \geq 1-\varepsilon$ and the following properties hold:

(i) For $S' \xleftarrow{\$} \mathcal{M}_{\mathbf{S}}$, $T' \xleftarrow{\$} \mathcal{M}_{\mathbf{T}}$, $t' := t/\varphi_{\mathsf{hc}}$, and $q' := q/\varphi_{\mathsf{hc}}$, we have

$$\Delta_{t',q'}(\mathbf{S}(S'), \mathbf{T}(T')) \leq 2\gamma;$$

(ii) There exist a (ζ_1, ζ_2)-sampler $\mathsf{O}_{\mathbf{S}}$ for $\mathcal{M}_{\mathbf{S}}$ and $A_{\mathbf{S}}$ with length $s_{A_{\mathbf{S}}}(\psi_{\mathsf{hc}} \cdot q')$ and a (ζ_1, ζ_2)-sampler $\mathsf{O}_{\mathbf{T}}$ for $\mathcal{M}_{\mathbf{T}}$ and $A_{\mathbf{T}}$ with length $s_{A_{\mathbf{T}}}(\psi_{\mathsf{hc}} \cdot q')$. Furthermore, if both $\mathbf{S}$ and $\mathbf{T}$ are random functions, then both samplers can be made error-less with lengths $s_{A_{\mathbf{S}}}(\psi \cdot q')$ and $s_{A_{\mathbf{T}}}(\psi \cdot q')$, where $\psi := 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$.

We postpone the full proof to Appendix D, which relies on Theorem 1, and only present the main ideas in the following. Furthermore, we state a uniform version of the theorem in Appendix F.

Proof Sketch. We define $(\mathbf{F}, B) \equiv (\mathbf{F}(X, B), B)$ to be the cc-stateless system-bit pair with a uniform random bit B and where $\mathbf{F}$ behaves as $\mathbf{S}$ if $B = 0$ and as $\mathbf{T}$ if $B = 1$. In particular, the initial state (X, B) of $(\mathbf{F}, B)$ is sampled by first letting $B \xleftarrow{\$} \{0,1\}$, and then choosing $X \xleftarrow{\$} \mathsf{P}_S$ if $B = 0$ and $X \xleftarrow{\$} \mathsf{P}_T$ otherwise, and

$$(\mathbf{F}(x,b), B(x,b)) = \begin{cases} (\mathbf{S}(x), 0) & \text{if } b = 0, \\ (\mathbf{T}(x), 1) & \text{if } b = 1. \end{cases}$$

By a standard argument $\Delta_{t,q}(\mathbf{S}, \mathbf{T}) = \mathrm{Guess}_{t,q}(B \,|\, \mathbf{F}) \leq \varepsilon$ holds (also cf. Example 2), and Theorem 1 thus implies that there exists a measure $\mathcal{M}$ for $(\mathbf{F}, B)$ such that $\mu(\mathcal{M}) \geq 1-\varepsilon$, and $\mathrm{Guess}_{t',q'}(B' \,|\, \mathbf{F}(X')) \leq \gamma$, where $(X', B') \xleftarrow{\$} \mathcal{M}$, $t' = t/\varphi_{\mathsf{hc}}$, and $q' = q/\varphi_{\mathsf{hc}}$. Define $\mathcal{M}_{\mathbf{S}}(s) := \mathcal{M}(s, 0)$ and $\mathcal{M}_{\mathbf{T}}(t) := \mathcal{M}(t, 1)$, and note that

$$\mathsf{P}_{X'B'}(s,0) = \frac{1}{2\mu(\mathcal{M})} \cdot \mathsf{P}_S(s) \cdot \mathcal{M}_{\mathbf{S}}(s) \quad \text{and} \quad \mathsf{P}_{X'B'}(t,1) = \frac{1}{2\mu(\mathcal{M})} \cdot \mathsf{P}_T(t) \cdot \mathcal{M}_{\mathbf{T}}(t). \tag{1}$$

If B' were uniformly distributed (i.e., $\sum_s \mathsf{P}_{X'B'}(s,0) = \sum_t \mathsf{P}_{X'B'}(t,1) = \frac{1}{2}$), we then would have $\mu(\mathcal{M}_{\mathbf{S}}) = \mu(\mathcal{M}_{\mathbf{T}}) = \mu(\mathcal{M}) \geq 1 - \varepsilon$ by (1), and (X', B') could be sampled by choosing B' uniformly, and letting $X' = S' \xleftarrow{\$} \mathcal{M}_{\mathbf{S}}$ if $B' = 0$, and $X' = T' \xleftarrow{\$} \mathcal{M}_{\mathbf{T}}$ if $B' = 1$. This would also yield

$$\Delta_{t',q'}(\mathbf{S}(S'), \mathbf{T}(T')) = \mathrm{Guess}_{t',q'}(B' \,|\, \mathbf{F}(X')) \leq \gamma,$$

concluding the proof. The main challenge in the full proof (cf. Appendix D) is dealing with the fact that B' is generally only $\Theta(\gamma)$-close to uniform.

Remark 2. Theorem 2 can be seen as a computational analogue of Lemma 5 in [30], which shows a similar property for information-theoretic indistinguishability (i.e., with respect to computationally unbounded distinguishers). Theorem 2 can of course also be used in the IT setting, and it is somewhat stronger in that it yields events defined on the initial state of the system, instead of interaction-dependent sequences of events as in [30]. However, Lemma 5 in [30] holds for *arbitrary* systems and presents a tight reduction with $q' = q$ and no additive term γ, which we do not know how to achieve in the computational setting.

Connection to Computational Entropy. Let $\mathbf{Q}$ be a cc-stateless random permutation on $\mathcal{X}$ (with $N := |\mathcal{X}|$) with function table Q and such that $\Delta_{t,q}(\mathbf{Q}, \mathbf{P}) \leq \varepsilon$ for a URP $\mathbf{P}$. Theorem 2 yields events $\mathcal{A}$ on Q and $\mathcal{B}$ on a uniform permutation table P such that $\mathsf{P}[\mathcal{A}] \geq 1-\varepsilon$, $\mathsf{P}[\mathcal{B}] \geq 1-\varepsilon$, and $\Delta_{t',q'}(\mathbf{Q}', \mathbf{P}') \leq \gamma$, where $\mathbf{Q}'$ and $\mathbf{P}'$ are cc-stateless random functions with function tables $Q' \xleftarrow{\$} \mathsf{P}_{Q|\mathcal{A}}$ and $P' \xleftarrow{\$} \mathsf{P}_{P|\mathcal{B}}$, respectively. In particular, $\mathsf{P}_{P'}(\pi) = \frac{\mathsf{P}_P(\pi) \cdot \mathsf{P}_{\mathcal{B}|P}(\pi)}{\mathsf{P}[\mathcal{B}]} \leq \frac{1}{(1-\varepsilon)\cdot(N!)}$ for all permutations π, and the *min-entropy* $\mathsf{H}_\infty(P') := -\log \max_\pi \mathsf{P}_{P'}(\pi)$ is at least $\log(N!) - \log\left((1-\varepsilon)^{-1}\right)$. Informally, this can be interpreted as the fact that Q has *"computational"* min-entropy at most $\log\left((1-\varepsilon)^{-1}\right)$ away from the maximal achievable entropy $\log(N!)$ with probability $1-\varepsilon$.[7] Clearly, the statement also extends to the two-sided case as well as to other types of systems.

[7] We stress, however, that the distribution P' depends on t, q, as well as on γ.

Remark 3. Another useful fact is that P' has statistical distance ε from P. This can be shown using the fact that the distribution of P' is a convex combination of flat distributions over subsets of size at least $(1-\varepsilon)\cdot(N!)$: As each such distribution is ε-away from uniform, the bound follows from the convexity of the statistical distance. Therefore $\Delta_{t,q}(\mathbf{P}',\mathbf{P}) \le \Delta_{t,q}(\langle\mathbf{P}'\rangle,\langle\mathbf{P}\rangle) \le d(P',P) \le \varepsilon$ for *any* t,q.

4 Cascade of Weak Permutations

4.1 Cascade of Permutations with Large Entropy

Let $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be two independent cc-stateless random permutations on the set $\mathcal{X}$ (with $N := |\mathcal{X}|$) with the property that the *min*-entropies of their respective function tables Q_1 and Q_2 satisfy $\mathsf{H}_\infty(Q_1) \ge \log(N!) - \log\left((1-\varepsilon)^{-1}\right)$ and $\mathsf{H}_\infty(Q_2) \ge \log(N!) - \log\left((1-\varepsilon)^{-1}\right)$ for some $\varepsilon \in \left[0, 1-\frac{1}{N}\right)$. We prove that the cascade $\mathbf{Q}_1 \rhd \mathbf{Q}_2$ is indistinguishable from a URP $\mathbf{P}$ for computationally *unbounded* distinguishers, both in the one- and in the two-sided cases.

Theorem 3 (Cascade of Large-Min-Entropy Permutations). *For all $q, \Lambda \ge 1$,*

$$\Delta_q(\mathbf{Q}_1 \rhd \mathbf{Q}_2, \mathbf{P}) \le \Delta_q(\langle\mathbf{Q}_1 \rhd \mathbf{Q}_2\rangle, \langle\mathbf{P}\rangle) \le \tfrac{4q\Lambda}{N} + \tfrac{2\Lambda(q+\Lambda)}{(1-\varepsilon)N} + 2\left(\frac{q\log\left((1-\varepsilon)^{-1}\right)}{\Lambda}\right)^{\frac{1}{2}}.$$

The same bound applies to any cascade $\mathbf{Q}'_1 \rhd \cdots \rhd \mathbf{Q}'_m$ of m independent cc-stateless random permutations such that $\mathbf{Q}'_i \equiv \mathbf{Q}_1$ and $\mathbf{Q}'_j \equiv \mathbf{Q}_2$ for some $i < j$, as such a cascade can be seen as the cascade of two permutations $\overline{\mathbf{Q}}_1 := \mathbf{Q}'_1 \rhd \cdots \rhd \mathbf{Q}'_i$ and $\overline{\mathbf{Q}}_2 := \mathbf{Q}'_{i+1} \rhd \cdots \rhd \mathbf{Q}'_m$ with the same min-entropy guarantees on their function tables. The theorem allows free choice of Λ: For our purposes, it suffices to set $\Lambda := (\log N)^\zeta$ (for a slowly growing $\zeta = \omega(1)$ in the security parameter $\log N$) to achieve indistinguishability for $q = \mathsf{poly}(\log N)$ queries and any $\varepsilon \le 1 - \frac{(\log N)^{3\zeta}}{N}$.

The core of the proof (given in Appendix E) is a lemma stating that $\langle\mathbf{Q}_i\rangle$ (for $i = 1, 2$) is indistinguishable from a random permutation $\left\langle (\mathbf{Q}_i)_{\overline{\mathbf{D}}_i} \right\rangle$ which is initialized by letting a carefully-chosen distinguisher $\overline{\mathbf{D}}_i$ (making Λ queries) interact with $\langle\mathbf{Q}_i\rangle$, and then answering queries according to a randomly chosen permutation consistent with $\overline{\mathbf{D}}_i$'s interaction. (This extends a previous result by Unruh [43] to random permutations.) We employ tools from the random systems framework [27] (including a new lemma) to prove that the cascade of two independent such permutations is indistinguishable from a URP.

4.2 Security Amplification of Weak PRPs

Let $\mathbf{Q}$ be a cc-stateless random permutation with domain $\mathcal{X}$ (for $N := |\mathcal{X}| = 2^n$, where n is the security parameter) such that $\langle\mathbf{Q}\rangle$ is implemented by the algorithm $A_{\langle\mathbf{Q}\rangle}$ with time complexity $t_{A_{\langle\mathbf{Q}\rangle}}$ and space complexity $s_{A_{\langle\mathbf{Q}\rangle}}$. We also consider the canonical (efficient) implementation of a two-sided URP $\langle\mathbf{P}\rangle$ that maintains a table consisting of all input-output pairs (x_i, y_i) of previous queries as its state, and, upon a new query $(x, +)$ or $(y, -)$, it chooses uniformly at random a y' (or x') not appearing as the second (first) element in a previous input-output pair, and adds (x, y') (or (x', y)) to the table. (If a corresponding pair is in the table, it answers accordingly.) Thus each query is answered in time $\mathcal{O}(\log(s))$, where s is the size of the table, and $s = \mathcal{O}(q \cdot n)$ after q queries.

The following is the main security amplification result of this paper.

Theorem 4. *Let $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ be independent instances of $\mathbf{Q}$ and let $\mathbf{P}$ be a URP, and assume that for some t, q we have $\Delta_{t,q}(\langle\mathbf{Q}\rangle, \langle\mathbf{P}\rangle) \le \varepsilon$. For all $\gamma > 1$ and $\Lambda > 0$,*

$$\Delta_{t'',q''}(\langle\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m\rangle, \langle\mathbf{P}\rangle) \le (m-(m-1)\varepsilon)\cdot\varepsilon^m + \tfrac{4q''\Lambda}{N} + \tfrac{2\Lambda(q''+\Lambda)}{(1-\varepsilon)N} + 2\left(\frac{q''\log\left((1-\varepsilon)^{-1}\right)}{\Lambda}\right)^{\frac{1}{2}} + (2m+2)\gamma$$

where $t'' = t/\varphi_{\mathsf{hc}} - (m-1)\max\left\{t_{A_{\langle\mathbf{Q}\rangle}}(q'', s_{A_{\langle\mathbf{Q}\rangle}}(q''\cdot\psi)), \mathcal{O}\left(q''\log(q''\cdot(\psi+1)n)\right)\right\}$ *and* $q'' = q/\varphi_{\mathsf{hc}}$ *for* $\psi := 7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$ *and* φ_{hc} *as in Theorem 2.*

Essentially the same result can be proven for the single-sided case. The proof of Theorem 4 follows the intuition that, with very high probability, at least two permutations in the cascade are computational indistinguishable from random permutations with high-entropy, allowing application of Theorem 3. Extra work is required to prove a non-trivial bound for the case where *at most* one permutation is guaranteed to have high-entropy. We postpone a proof of the tightness of these bounds to Section 4.3.

Proof. Theorem 2 implies that we can define (two-sided) random permutations $\langle\mathbf{Q}'\rangle, \langle\mathbf{Q}''\rangle$, and $\langle\mathbf{P}'\rangle$ such that the following three properties hold for some $p \leq \varepsilon$: (i) The function table of $\langle\mathbf{P}'\rangle$ has min-entropy at least $\log(N!) - \log\left((1-\varepsilon)^{-1}\right)$, (ii) $\langle\mathbf{Q}\rangle$ behaves as $\langle\mathbf{Q}'\rangle$ with probability $1-p$ and as $\langle\mathbf{Q}''\rangle$ with probability p, and (iii) $\Delta_{t',q''}(\langle\mathbf{Q}'\rangle, \langle\mathbf{P}'\rangle) \leq 2\gamma$ for $t' := t/\varphi_{\mathsf{hc}}$. Furthermore, $\langle\mathbf{Q}'\rangle$ and $\langle\mathbf{Q}''\rangle$ can both be perfectly implemented using $A_{\langle\mathbf{Q}\rangle}$ initialized with some appropriately distributed state of length at most $s_{A_{\langle\mathbf{Q}\rangle}}(q''\cdot\psi)$ given as advice. Similarly, $\langle\mathbf{P}'\rangle$ can be simulated by running the above canonical algorithm initialized with an appropriate state of length $\mathcal{O}(q''\cdot\psi\cdot n)$. (See the discussion in Section 3.1.)

Additionally, for $\mathcal{I} \subseteq \{1,\ldots,m\}$, let $\mathcal{A}_I$ be the event that $\langle\mathbf{Q}_i\rangle$ behaves as $\langle\mathbf{Q}'\rangle$ for all $i \in \mathcal{I}$ whereas $\langle\mathbf{Q}_i\rangle$ behaves as $\langle\mathbf{Q}''\rangle$ for all $i \notin \mathcal{I}$. Likewise, for independent instances $\langle\mathbf{Q}'_i\rangle$ and $\langle\mathbf{Q}''_i\rangle$ (for $i = 1,\ldots,m$) of $\langle\mathbf{Q}'\rangle$ and $\langle\mathbf{Q}''\rangle$, respectively, let $\mathbf{Q}_{\mathcal{I}} := \mathbf{S}_1 \rhd \cdots \rhd \mathbf{S}_m$, where $\mathbf{S}_i := \mathbf{Q}'_i$ for all $i \in \mathcal{I}$ and $\mathbf{S}_i := \mathbf{Q}''_i$ for all $i \notin \mathcal{I}$

We now fix some distinguisher $\mathbf{D}$ with time complexity t'' and making q'' queries, and we first observe that

$$\delta^{\mathbf{D}}(\langle\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m\rangle, \langle\mathbf{P}\rangle) = \sum_{\mathcal{I}\subseteq\{1,\ldots,m\}} q_{\mathcal{I}}\cdot\delta^{\mathbf{D}}(\langle\mathbf{Q}_{\mathcal{I}}\rangle, \langle\mathbf{P}\rangle), \tag{2}$$

where $\delta^{\mathbf{D}}(\mathbf{F},\mathbf{G}) := \mathsf{P}\left[\mathbf{D}(\mathbf{F})=1\right] - \mathsf{P}\left[\mathbf{D}(\mathbf{G})=1\right]$ and $q_{\mathcal{I}} := \mathsf{P}\left[\mathcal{A}_{\mathcal{I}}\right] = (1-p)^{|\mathcal{I}|}\cdot p^{m-|\mathcal{I}|}$.

We first upper bound the summands corresponding to sets $\mathcal{I}$ with at most one element. To this end, for all $i = 1,\ldots,m$, we define the distinguisher $\mathbf{D}_i$ which, given access to a two-sided random permutation $\langle\mathbf{S}\rangle$, outputs $\mathbf{D}(\langle\mathbf{Q}''_1 \rhd \cdots \rhd \mathbf{Q}''_{i-1} \rhd \mathbf{S} \rhd \mathbf{Q}''_{i+1} \rhd \cdots \rhd \mathbf{Q}''_m\rangle)$: Note that $\mathbf{D}_i$ can be implemented with time complexity $t'' + (m-1)t_{A_{\langle\mathbf{Q}\rangle}}(q', s_{A_{\langle\mathbf{Q}\rangle}}(\psi\cdot q')) \leq t'$ given the appropriate advice.

We have $\delta'_i := \delta^{\mathbf{D}_i}(\langle\mathbf{Q}'\rangle, \langle\mathbf{P}\rangle) = \delta^{\mathbf{D}_i}(\langle\mathbf{Q}'\rangle, \langle\mathbf{P}'\rangle) + \delta^{\mathbf{D}_i}(\langle\mathbf{P}'\rangle, \langle\mathbf{P}\rangle) \leq 2\gamma + \varepsilon$, where the bound on the first term follows from the hardcore lemma (for every fixed value of the advice), whereas the bound on the second term follows from the fact that $\delta^{\mathbf{D}_i}(\langle\mathbf{P}'\rangle, \langle\mathbf{P}\rangle) \leq \varepsilon$ (cf. Remark 3). Additionally, $\delta^{\mathbf{D}_i}(\langle\mathbf{Q}\rangle, \langle\mathbf{P}\rangle) = (1-p)\cdot\delta'_i + p\cdot\delta''_i \leq \varepsilon$ with $\delta''_i := \delta^{\mathbf{D}_i}(\langle\mathbf{Q}''\rangle, \langle\mathbf{P}\rangle)$ by the indistinguishability assumption on $\langle\mathbf{Q}\rangle$ and the fact that $t' < t$. Using the fact that $\left\langle\mathbf{Q}''_1 \rhd \cdots \rhd \mathbf{Q}''_{i-1} \rhd \mathbf{P} \rhd \mathbf{Q}''_{i+1} \rhd \cdots \rhd \mathbf{Q}''_m\right\rangle \equiv \langle\mathbf{P}\rangle$, we obtain $\delta^{\mathbf{D}}(\langle\mathbf{Q}_\emptyset\rangle, \langle\mathbf{P}\rangle) = \delta''_i$ and $\delta^{\mathbf{D}}(\langle\mathbf{Q}_{\{i\}}\rangle, \langle\mathbf{P}\rangle) = \delta'_i$ for all $i \in \{1,\ldots,m\}$, and thus

$$\begin{aligned}\sum_{|\mathcal{I}|\leq 1} q_{\mathcal{I}}\cdot\delta^{\mathbf{D}}(\langle\mathbf{Q}_{\mathcal{I}}\rangle), \langle\mathbf{P}\rangle) &= \sum_{i=1}^{m}\frac{1}{m}\cdot p^m\cdot\delta''_i + p^{m-1}(1-p)\cdot\delta'_i \\ &\leq \max_{i\in\{1,\ldots,m\}}\left\{p^m\cdot\delta''_i + m\cdot p^{m-1}\cdot(1-p)\cdot\delta'_i\right\}.\end{aligned}$$

However, for all $i \in \{1,\ldots,m\}$, we combine all of the above observations to obtain

$$\begin{aligned}p^m\delta''_i + mp^{m-1}(1-p)\delta'_i &= p^{m-1}(p\delta''_i + (1-p)\delta'_i) + (m-1)p^{m-1}(1-p)\delta'_i \\ &\leq p^{m-1}\varepsilon + (m-1)p^{m-1}(1-p)\varepsilon + 2\gamma \\ &\leq \varepsilon^m + (m-1)\varepsilon^m(1-\varepsilon) + 2\gamma = \varepsilon^m(m-(m-1)\varepsilon) + 2\gamma,\end{aligned}$$

where we also have used $p \leq \varepsilon$ and the fact that $p^m + (m-1)p^{m-1}(1-p)$ grows monotonically for $p \in [0,1]$.

To bound the remaining summands of Equation (2), we use a standard hybrid argument. Let $\langle \mathbf{P}'_1 \rangle, \ldots, \langle \mathbf{P}'_m \rangle$ be independent instances of $\langle \mathbf{P}' \rangle$. For given $\mathcal{I}$ (with $|\mathcal{I}| \geq 2$) and $i \in \{0, \ldots, m\}$, we define $\mathbf{Q}_{\mathcal{I},i}$ as the cascade $\mathbf{S}_1 \rhd \cdots \rhd \mathbf{S}_m$ where $\mathbf{S}_j = \mathbf{Q}''_j$ if $j \notin \mathcal{I}$, $\mathbf{S}_j = \mathbf{P}'_j$ if $j \in \mathcal{I}$ and $j \leq i$ and $\mathbf{S}_j := \mathbf{Q}'_j$ if $j \in \mathcal{I}$ and $j > i$. Then,

$$\delta^{\mathbf{D}}(\langle \mathbf{Q}_{\mathcal{I}} \rangle, \langle \mathbf{P} \rangle) = \sum_{i=1}^{m} \delta^{\mathbf{D}}(\langle \mathbf{Q}_{\mathcal{I},i-1} \rangle, \langle \mathbf{Q}_{\mathcal{I},i} \rangle) + \delta^{\mathbf{D}}(\langle \mathbf{Q}_{\mathcal{I},m} \rangle, \langle \mathbf{P} \rangle).$$

Let $\mathbf{D}'_i$ be the distinguisher which outputs $\mathbf{D}(\langle \mathbf{S}_1 \rhd \cdots \rhd \mathbf{S}_{i-1} \rhd \mathbf{S} \rhd \mathbf{S}_{i+1} \rhd \cdots \rhd \mathbf{S}_m \rangle)$ given access a system $\mathbf{S} \in \{\langle \mathbf{Q}' \rangle, \langle \mathbf{P}' \rangle\}$, with $\mathbf{S}_j$ defined as above, but simulated using the corresponding implementations of $\langle \mathbf{Q} \rangle$ and $\langle \mathbf{P} \rangle$, as well as appropriate advice: In particular, its time complexity is $t'' + (i-1) \cdot \mathcal{O}(q'' \log(q'' \cdot (\psi + 1)n)) + (m-i+1) \cdot t_{A_{\langle \mathbf{A} \rangle}}(q', s_{A_{\langle \mathbf{Q} \rangle}}(\psi \cdot q')) \leq t'$ and thus $\delta^{\mathbf{D}}(\langle \mathbf{Q}_{\mathcal{I},i-1} \rangle, \langle \mathbf{Q}_{\mathcal{I},i} \rangle) = \delta^{\mathbf{D}_i}(\langle \mathbf{Q}' \rangle, \langle \mathbf{P}' \rangle) \leq 2\gamma$ by Theorem 2 and the fact that this holds for all values of the advice. Finally, as $|\mathcal{I}| \geq 2$ and for all $j \in \mathcal{I}$ the function table of $\left\langle \mathbf{P}'_j \right\rangle$ in the cascade $\langle \mathbf{Q}_{\mathcal{I},m} \rangle$ has min-entropy at least $\log(N!) - \log\left((1-\varepsilon)^{-1}\right)$, we can bound $\delta^{\mathbf{D}}(\langle \mathbf{Q}_{\mathcal{I},m} \rangle, \langle \mathbf{P} \rangle)$ using Theorem 3. We conclude the proof by noticing that $\delta_{t',q'}(\langle \mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m \rangle, \langle \mathbf{P} \rangle) = \Delta_{t',q'}(\langle \mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m \rangle, \langle \mathbf{P} \rangle)$. □

This in particular yields the following corollary, by applying the above argument to all $\gamma = 1/p$ (for some polynomial p in n) and to all polynomially bounded t, q, and by choosing an appropriate $\Lambda := n^{\omega(1)}$:

Corollary 1. *If $E = \{E_k\}_{k \in \mathcal{K}}$ is a (two-sided) ε-PRP for $\varepsilon \leq 1 - \frac{1}{n^{\mathcal{O}(1)}}$ (for security parameter n), then for any $m = \mathsf{poly}(n)$ the cascade $\{E_{k_1} \circ \cdots \circ E_{k_m}\}_{k_1, \ldots, k_m \in \mathcal{K}}$ is a (two-sided) $(\varepsilon^m(m-(m-1)\varepsilon)+\nu)$-PRP for some negligible function ν, where $\circ$ denotes permutation composition.*

4.3 Tightness

Let $\varepsilon < 1 - 2^{-n}$ be such that $\log\left((1-\varepsilon)^{-1}\right) \in \{1, \ldots, n\}$. Let $\mathbf{Q} : \{0,1\}^n \to \{0,1\}^n$ be the cc-stateless random permutation which initially chooses $B \in \{0,1\}$ with $\mathsf{P}_B(0) = \varepsilon$. If $B = 0$, then $\mathbf{Q}$ behaves as the identity permutation id, whereas if $B = 1$ it behaves as a uniformly chosen permutation Q' with the constraint that the first $\log\left((1-\varepsilon)^{-1}\right)$ bits of $Q'(0^n)$ are all equal to 0. Clearly, it is possible to give an efficient *stateful* algorithm implementing $\mathbf{Q}$ (or $\langle \mathbf{Q} \rangle$) by using lazy sampling.[8]

In the following, let $\mathbf{P} : \{0,1\}^n \to \{0,1\}^n$ be a URP, and let $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ be *independent* instances of $\mathbf{Q}$. We prove the following two statements:

(i) For all distinguishers $\mathbf{D}$ we have $\Delta^{\mathbf{D}}(\langle \mathbf{Q} \rangle, \langle \mathbf{P} \rangle) \leq \varepsilon$, regardless of their computing power.
(ii) There exists a constant-time distinguisher $\mathbf{D}^*$ making *one* single forward query such that

$$\Delta^{\mathbf{D}^*}(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m, \mathbf{P}) = \Delta^{\mathbf{D}^*}(\langle \mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m \rangle, \langle \mathbf{P} \rangle) \geq (m - (m-1)\varepsilon)\varepsilon^m - \frac{1}{2^n}.$$

These two facts imply that the bound of Theorem 4 cannot be substantially improved, even if allowing a huge security loss (i.e., $t'' << t$ and $q'' << q$). This in particular extends to arbitrary m a previous tightness result given by Myers [35] for the special case $m = 2$.

[8] Also, from any PRP $E = \{E_k\}_{k \in \{0,1\}^n}$ with n-bit string domain, a permutation family $E' = \{E'_{k'}\}_{k' \in \{0,1\}^{\log(1/\varepsilon)+n}}$ which is computationally indistinguishable from $\mathbf{Q}$ under a uniform $(\log(1/\varepsilon)+n)$-bit random key can be defined as follows: For all $k' \in \{0,1\}^{\log(1/\varepsilon)}$ and $k \in \{0,1\}^n$, let $E'_{k'\|k}(x) := x$ if $k' = 0^{\log(1/\varepsilon)}$, and $E'_{k'\|k} := E_k(x) \oplus E_k(0^n)|_{\log((1-\varepsilon)^{-1})}$ otherwise, where $x|_r$ i sets the last $n-r$ bits of x to be 0 (and leaves the first r unchanged) and $\|$ denotes string concatenation.

$\mathbf{Q}$ IS A TWO-SIDED ε-PRP. In the following, let Q and P be random variables representing the distributions of the permutation tables of $\mathbf{Q}$ and $\mathbf{P}$, respectively. There are $(1-\varepsilon)(2^n!)$ permutations π for which the last $\log\left((1-\varepsilon)^{-1}\right)$ bits of $\pi(0^n)$ all equal to 0, and the identity id is one such permutation. Hence,

$$\mathsf{P}_Q(\mathrm{id}) = \varepsilon + (1-\varepsilon)\cdot\frac{1}{(1-\varepsilon)(2^n!)} = \varepsilon + \frac{1}{2^n!} \geq \frac{1}{2^n!} = \mathsf{P}_P(\mathrm{id}).$$

For all $\pi \neq \mathrm{id}$, we have $\mathsf{P}_Q(\pi) \leq (1-\varepsilon)\cdot\frac{1}{(1-\varepsilon)(2^n!)} = \frac{1}{2^n!} = \mathsf{P}_P(\pi)$. This yields $\Delta^{\mathbf{D}}(\langle\mathbf{P}\rangle, \langle\mathbf{Q}\rangle) \leq d(P,Q) = \mathsf{P}_Q(\mathrm{id}) - \mathsf{P}_P(\mathrm{id}) = \varepsilon$ for *all* distinguishers $\mathbf{D}$.

LOWER BOUND FOR DISTINGUISHING THE CASCADE. We define $\mathbf{D}^*$ as the distinguisher querying 0^n and outputting 1 if and only if the first $\log\left((1-\varepsilon)^{-1}\right)$ bits of the resulting output are all 0, and outputting 0 otherwise. It is easy to verify that $\mathsf{P}[\mathbf{D}^*(\mathbf{P})=1] = 2^{-\log\left((1-\varepsilon)^{-1}\right)} = 1-\varepsilon$, as the output of $\mathbf{P}$ on input 0^n is a uniformly distributed n-bit string.

Denote as B_i the bit B associated with the i-th instance $\mathbf{Q}_i$, and let $\mathcal{A}_{\mathcal{I}}$ for $\mathcal{I} \subseteq \{1,\ldots,m\}$ be the event that $B_i = 1$ for all $i \in \mathcal{I}$ and $B_i = 0$ for all $i \notin \mathcal{I}$. Furthermore, let $\mathcal{E}$ be the event that $\mathcal{A}_{\mathcal{I}}$ occurs for some $\mathcal{I}$ with $|\mathcal{I}| \leq 1$. Clearly, $\mathsf{P}[\mathcal{E}] = \varepsilon^m + m(1-\varepsilon)\varepsilon^{m-1}$ and $\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1 \,\middle|\, \mathcal{E}\right] = 1$, since $\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m$ under $\mathcal{E}$ behaves either as the identity or as Q', and in both cases the first $\log\left((1-\varepsilon)^{-1}\right)$ output bits are all 0.

Assume that $\mathcal{A}_{\mathcal{I}}$ occurs for $\mathcal{I}$ with $k := |\mathcal{I}| \geq 2$, and let $\mathbf{Q}'_1, \ldots, \mathbf{Q}'_k$ be independent random permutations answering according to Q'. Then $\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1 \,\middle|\, \mathcal{A}_{\mathcal{I}}\right] = \mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}'_1 \rhd \cdots \rhd \mathbf{Q}'_k) = 1\right]$. Note that for any input $x \neq 0^n$ the probability that the first $\log\left((1-\varepsilon)^{-1}\right)$ output bits of $\mathbf{Q}'_k(x)$ are all 0 is exactly $1-\varepsilon$, whereas the probability that $\mathbf{Q}'_k$ is invoked on 0^n is *at most* $\frac{1}{(1-\varepsilon)2^n}$ (as regardless of the input, the output $\mathbf{Q}'_{k-1}$ is uniformly distributed on a set of at least size $(1-\varepsilon)2^n$), and therefore

$$\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}'_1 \rhd \cdots \rhd \mathbf{Q}'_k) = 1\right] \geq \left(1 - \frac{1}{(1-\varepsilon)2^n}\right)\cdot(1-\varepsilon) = 1 - \varepsilon - \frac{1}{2^n},$$

and therefore we also have $\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1 \,\middle|\, \overline{\mathcal{E}}\right] \geq 1 - \varepsilon - \frac{1}{2^n}$, from which we conclude

$$\begin{aligned}
\Delta^{\mathbf{D}^*}(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m, \mathbf{P}) &\geq \mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1\right] - \mathsf{P}\left[\mathbf{D}^*(\mathbf{P}) = 1\right] \\
&= \mathsf{P}[\mathcal{E}]\cdot\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1 \,\middle|\, \mathcal{E}\right] + (1-\mathsf{P}[\mathcal{E}])\cdot\mathsf{P}\left[\mathbf{D}^*(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m) = 1 \,\middle|\, \overline{\mathcal{E}}\right] - (1-\varepsilon) \\
&= \mathsf{P}[\mathcal{E}]\cdot 1 + (1-\mathsf{P}[\mathcal{E}])\left(1 - \varepsilon - \frac{1}{2^n}\right) - (1-\varepsilon) \geq \mathsf{P}[\mathcal{E}]\cdot\varepsilon - \frac{1}{2^n} \\
&\geq \varepsilon^{m+1} + m(1-\varepsilon)\varepsilon^m - \frac{1}{2^m} = (m-(m-1)\varepsilon)\varepsilon^m - \frac{1}{2^n}.
\end{aligned}$$

5 Conclusions and Open Problems

This paper has presented the first tight analysis of the security amplification properties of the cascade of weak PRPs, both in the one- and two-sided cases. Our main tool is a hardcore lemma (Theorem 2) for computational indistinguishability of discrete interactive cc-stateless systems. It is our belief that the generality of this result makes it suitable to the solution of a number of other problems. For instance, an interesting problem is whether *parallel* and *deterministic* security-amplifying constructions for *arbitrarily* weak pseudorandom *functions* exist. To date, the best known constructions are either randomized [36, 33], or only work for moderately weak PRFs [7, 33]. Also, quantitative improvements of our results should also be of interest. One may try to minimize the length of the state output by the state sampler or to improve the bound of Theorem 3.

We stress that we have followed *one* possible path to generalize the hardcore lemma to the setting of interactive systems, tailored at applications in the context of secret-key primitives. It is a meaningful

question to investigate whether alternative approaches are possible which would be applicable to other areas, such as security amplification of two-party *protocols*, where in particular no repetition is allowed within the reduction.

Acknowledgments. The author is grateful to Peter Gaži, Thomas Holenstein, Russell Impagliazzo, Ueli Maurer, and Salil Vadhan for helpful discussions and insightful feedback. This work was done while the author was a graduate student at ETH Zurich, supported in part by the Swiss National Science Foundation (SNF), project no. 200020-113700/1. He is currently at UCSD, partially supported by NSF grant CNS-0716790.

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, 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.
4. K.-M. Chung and F.-H. Liu, "Parallel repetition theorems for interactive arguments," in *Theory of Cryptography — TCC 2010*, vol. 5978 of *Lecture Notes in Computer Science*, pp. 19–36, 2010.
5. T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.
6. 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.
7. 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.
8. Y. Dodis, K. Pietrzak, and P. Puniya, "A new mode of operation for block ciphers and length-preserving MACs," in *Advances in Cryptology — EUROCRYPT 2008*, vol. 4965 of *Lecture Notes in Computer Science*, pp. 198–219, 2008.
9. Y. Dodis and J. P. Steinberger, "Message authentication codes from unpredictable block ciphers," in *Advances in Cryptology — CRYPTO 2009*, vol. 5677 of *Lecture Notes in Computer Science*, pp. 267–285, 2009.
10. 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.
11. S. Even and O. Goldreich, "On the power of cascade ciphers," *ACM Transactions on Computer Systems*, vol. 3, no. 2, pp. 108–116, 1985.
12. P. Gaži and U. Maurer, "Cascade encryption revisited," in *Advances in Cryptology — ASIACRYPT 2009*, vol. 5912 of *Lecture Notes in Computer Science*, pp. 37–51, Dec. 2009.
13. P. Gaži and U. Maurer, "Free-start distinguishing: Combining two types of indistinguishability amplification," in *ICITS 2009*, Lecture Notes in Computer Science, 2010.
14. 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.
15. O. Goldreich, N. Nisan, and A. Wigderson, "On Yao's XOR-lemma," *Electronic Colloquium on Computational Complexity (ECCC)*, vol. 2, no. 50, 1995.
16. I. Haitner, "A parallel repetition theorem for any interactive argument," in *FOCS '09: Proceedings of the 50th IEEE Annual Symposium on Foundations of Computer Science*, pp. 241–250, 2009.
17. 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.
18. 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.
19. J. Håstad, R. Pass, D. Wikström, and K. Pietrzak, "An efficient parallel repetition theorem," in *Theory of Cryptography — TCC 2010*, vol. 5978 of *Lecture Notes in Computer Science*, pp. 1–18, 2010.
20. W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
21. 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.

22. T. Holenstein, "Pseudorandom generators from one-way functions: A simple construction for any hardness," in *Theory of Cryptography — TCC 2006*, vol. 3876 of *Lecture Notes in Computer Science*, pp. 443–461, 2006.
23. 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.
24. 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.
25. 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.
26. M. Luby and C. Rackoff, "How to construct pseudorandom permutations from pseudorandom functions," *SIAM Journal on Computing*, vol. 17, no. 2, pp. 373–386, 1988.
27. U. Maurer, "Indistinguishability of random systems," in *Advances in Cryptology — EUROCRYPT 2002*, vol. 2332 of *Lecture Notes in Computer Science*, pp. 110–132, 2002.
28. 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.
29. U. Maurer and K. Pietrzak, "Composition of random systems: When two weak make one strong," in *Theory of Cryptography — TCC 2004*, vol. 2951 of *Lecture Notes in Computer Science*, pp. 410–427, 2004.
30. 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.
31. 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.
32. U. Maurer and S. Tessaro, "Basing PRFs on constant-query weak PRFs: Minimizing assumptions for efficient symmetric cryptography," in *Advances in Cryptology — ASIACRYPT 2008*, vol. 5350 of *Lecture Notes in Computer Science*, pp. 161–178, 2008.
33. U. Maurer and S. Tessaro, "Computational indistinguishability amplification: Tight product theorems for system composition," in *Advances in Cryptology — CRYPTO 2009*, vol. 5677 of *Lecture Notes in Computer Science*, pp. 350–368, 2009.
34. U. Maurer and S. Tessaro, "A hardcore lemma for computational indistinguishability: Security amplification for arbitrarily weak prgs with optimal stretch," in *Theory of Cryptography — TCC 2010*, vol. 5978 of *Lecture Notes in Computer Science*, pp. 237–254, 2010.
35. 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.
36. S. Myers, "Efficient amplification of the security of weak pseudo-random function generators," *Journal of Cryptology*, vol. 16, pp. 1–24, 2003.
37. S. Myers, "Black-box composition does not imply adaptive security," in *Advances in Cryptology — EUROCRYPT 2004*, vol. 3027 of *Lecture Notes in Computer Science*, pp. 189–206, 2004.
38. 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.
39. 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.
40. K. Pietrzak, "Composition does not imply adaptive security," in *Advances in Cryptology — CRYPTO 2005*, vol. 3621 of *Lecture Notes in Computer Science*, pp. 55–65, 2005.
41. K. Pietrzak, "Composition implies adaptive security in minicrypt," in *Advances in Cryptology — EUROCRYPT 2006*, vol. 4004 of *Lecture Notes in Computer Science*, pp. 328–338, 2006.
42. 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.
43. D. Unruh, "Random oracles and auxiliary input," in *Advances in Cryptology — CRYPTO 2007*, vol. 4622 of *Lecture Notes in Computer Science*, pp. 205–223, 2007.
44. 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.
45. J. Wullschleger, "Oblivious-transfer amplification," in *Advances in Cryptology — EUROCRYPT 2007*, vol. 4515 of *Lecture Notes in Computer Science*, pp. 555–572, 2007.
46. 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 Tail Estimates

The following well-known result from probability theory [20] is repeatedly used throughout this paper.

Lemma 1 (Hoeffding's Inequalities). *Let $X_1, \ldots, X_\varphi$ be independent random variables with range $[0,1]$, and let $\overline{X} := \frac{1}{\varphi}\sum_{i=1}^{\varphi} X_i$. Then, for all $\varepsilon > 0$ we have*

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

In particular,

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

B An Example: The XOR Lemma for System-Bit Pairs

In this section we provide a self-contained proof of a generalization of Yao's XOR Lemma [46, 15] to system-bit pairs: Namely, given m multiple instances $(\mathbf{S}_1, B_1), \ldots, (\mathbf{S}_m, B_m)$ of a cc-stateless system-bit pair $(\mathbf{S}, B) \equiv (\mathbf{S}(S), S(S))$ with the property that $\mathrm{Guess}_{t,q}(B \,|\, \mathbf{S}) \leq \varepsilon$, we prove that given access to all of $\mathbf{S}_1, \ldots, \mathbf{S}_m$ in *parallel*[9] (we denote the parallel composition as $\mathbf{S}_1 \| \ldots \| \mathbf{S}_m$) the XOR $B_1 \oplus \cdots \oplus B_m$ can only be guessed with advantage $\mathrm{Guess}_{t',q',\ldots,q'}(B_1 \oplus \cdots \oplus B_m \,|\, \mathbf{S}_1 \| \ldots \| \mathbf{S}_m)$ at most $\varepsilon^m + \gamma$ by an adversary with related time complexity t' and making at most q' queries to each system $\mathbf{S}_i$. The obtained parameters are slightly worse than those of the direct proof of [33] (which in turn was based on Levin's approach [15]), but the given reduction is simpler (given Theorem 1) and nicely illustrates the use of state samplers with non-zero errors ζ_1 and ζ_2.

Theorem 5 (XOR Lemma). *If, for some integers $t, q > 0$, and $\varepsilon \in [0, 1)$,*

$$\mathrm{Guess}_{t,q}(B \,|\, \mathbf{S}) \leq \varepsilon,$$

then, for all $0 < \zeta_1, \zeta_2 < 1$ and $0 < \gamma \leq \frac{1}{2}$, we have

$$\mathrm{Guess}_{t',q',\ldots,q'}(B_1 \oplus \cdots \oplus B_m \,|\, \mathbf{S}_1 \| \ldots \| \mathbf{S}_m) \leq \varepsilon^m + 2m \cdot (\zeta_1 + \zeta_2) + \gamma,$$

where $t = \varphi_{\mathsf{hc}} \cdot \left[t' + (m-1) \cdot t_{A_{(\mathbf{S},B)}}(q, s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}} \cdot q'))\right]$ and $q = \varphi_{\mathsf{hc}} \cdot q'$.

Proof. Let $\mathbf{A}$ be an adversary with time complexity t issuing q queries to each subsystem $\mathbf{S}_1, \ldots, \mathbf{S}_m$ and outputting a guess B' for $B_1 \oplus \cdots \oplus B_m$. Further, using the HCL (Theorem 1), let $\mathcal{M}$ with $\mu(\mathcal{M}) \geq 1 - \varepsilon$ be such that

$$\mathrm{Guess}_{t'/\varphi_{\mathsf{hc}},q}(B(S') \,|\, \mathbf{S}(S')) \leq \gamma \tag{3}$$

for $S' \xleftarrow{\$} \mathcal{M}$, and let O be the corresponding (ζ_1, ζ_2)-sampler for $\mathcal{M}$ and $A_{(\mathbf{S},B)}$ with length $s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}} \cdot q)$.

We consider an adversary $\mathbf{A}'$ which, given access to $\mathbf{S}(S')$, first samples m pairs $(\Sigma_1, Z_1), \ldots, (\Sigma_m, Z_m)$ using O, and subsequently flips independent bits $U_1, \ldots, U_m \in \{0,1\}$, where $\mathsf{P}_{U_j}(1) := Z_j$ for $j = 1, \ldots, m$. Then, it picks an $i \in \{1, \ldots, m\}$ (provided it exists) such that $U_i = 1$, and simulates the interaction of $\mathbf{A}$ with

$$A_{\mathbf{S}}[\Sigma_1] \,\|\, \ldots \,\|\, A_{\mathbf{S}}[\Sigma_{i-1}] \,\|\, \mathbf{S}(S') \,\|\, A_{\mathbf{S}}[\Sigma_{i+1}] \,\|\, \cdots \,\|\, A_{\mathbf{S}}[\Sigma_m],$$

where $A_{\mathbf{S}}[\Sigma_j]$ is the instance of $\mathbf{S}$ simulated by $A_{(\mathbf{S},B)}[\Sigma_j]$ (with B'_j being the associated bit), for $j = 1, \ldots, m$, $j \neq i$. When $\mathbf{A}$ outputs a bit B', $\mathbf{A}'$ outputs $B' \oplus \bigoplus_{j \neq i} B'_j$. If no such i exists, $\mathbf{A}'$ terminates directly by returning a random bit. Clearly, $\mathbf{A}'$ can equivalently be seen as sampling m independent pairs $(S_1, Z(S_1)), \ldots, (S_m, Z(S_m))$, letting U_j be one with probability $Z(S_j)$, and replacing

[9] In particular, the adversary can query at any point in time any system $\mathbf{S}_i$, adaptively depending on the outcomes of previous queries to this and other systems $\mathbf{S}_j$ for $j \neq i$.

$A_{\mathbf{S}}[\Sigma_j]$ and B'_j with $\mathbf{S}(S_j)$ and $B(S_j)$) for all $j = 1, \ldots, m$, $j \neq i$. This does not alter the guessing advantage.

We also consider the adversary $\widetilde{\mathbf{A}}'$ where $Z(S_j)$ is replaced by $\mathcal{M}(S_j)$ for all $j = 1, \ldots, m$, and denote as $\widetilde{U}_1, \ldots, \widetilde{U}_m$ the corresponding independent bits. (Remark that $\mathsf{P}_{\widetilde{U}_j}(1) = \mu(\mathcal{M}) \geq 1 - \varepsilon$.) Analogously, in the experiment where $\mathbf{A}$ interacts with $\mathbf{S}_1 \| \ldots \| \mathbf{S}_m$, define the bits $\widetilde{U}_j$ accordingly. In both experiments, let $\mathcal{E}$ be the event that $\widetilde{U}_1 = \cdots \widetilde{U}_m = 0$. Note that $\mathsf{P}[\mathcal{E}] \leq \varepsilon^m$. Furthermore,

$$\mathsf{P}\left[\widetilde{\mathbf{A}}'(\mathbf{S}(S')) = B(S') \,\middle|\, \overline{\mathcal{E}}\right] = \mathsf{P}\left[\mathbf{A}(\mathbf{S}_1 \| \ldots \| \mathbf{S}_m) = B_1 \oplus \cdots \oplus B_m \,\middle|\, \overline{\mathcal{E}}\right],$$

from which, using the facts that $\mathsf{P}\left[\mathbf{A}(\mathbf{S}_1 \| \ldots \| \mathbf{S}_m) = B_1 \oplus \cdots \oplus B_m \,\middle|\, \mathcal{E}\right] \leq 1$ and that $\mathsf{P}\left[\widetilde{\mathbf{A}}'(\mathbf{S}(S')) = B(S') \,\middle|\, \mathcal{E}\right] = \frac{1}{2}$, we can upper bound

$$\begin{aligned}
\mathsf{P}[\mathbf{A}(\mathbf{S}_1 \| \ldots \| \mathbf{S}_m) = B_1 \oplus \cdots \oplus B_m] \leq \\
&\leq \mathsf{P}[\mathcal{E}] + (1 - \mathsf{P}[\mathcal{E}]) \cdot \mathsf{P}\left[\widetilde{\mathbf{A}}'(\mathbf{S}(S')) = B(S') \,\middle|\, \overline{\mathcal{E}}\right] \\
&\leq \frac{\mathsf{P}[\mathcal{E}]}{2} + \mathsf{P}[\mathcal{E}] \cdot \frac{1}{2} + (1 - \mathsf{P}[\mathcal{E}]) \cdot \mathsf{P}\left[\widetilde{\mathbf{A}}'(\mathbf{S}(S')) = B(S') \,\middle|\, \overline{\mathcal{E}}\right] \\
&\leq \frac{\varepsilon^m}{2} + \mathsf{P}\left[\widetilde{\mathbf{A}}'(\mathbf{S}(S')) = B(S')\right],
\end{aligned}$$

or, in other words

$$\mathrm{Guess}^{\mathbf{A}}(B_1 \oplus \cdots \oplus B_m \,|\, \mathbf{S}_1 \| \ldots \| \mathbf{S}_m) \leq \varepsilon^m + \mathrm{Guess}^{\widetilde{\mathbf{A}}'}(B(S') \,|\, \mathbf{S}(S')). \tag{4}$$

Furthermore, note that for all $j = 1, \ldots, m$ and $b \in \{0, 1\}$

$$d(\widetilde{U}_j, U_j) = \left|\mathsf{P}_{\widetilde{U}_j}(b) - \mathsf{P}_{U_j}(b)\right| \leq \zeta_1 + \mathsf{P}\left[\mathcal{E}'\right] \leq \zeta_1 + \zeta_2,$$

where $\mathcal{E}'$ is the event that an error larger than ζ_1 occurs in the estimate. Then, it is not hard to verify that

$$\begin{aligned}
\mathrm{Guess}^{\widetilde{\mathbf{A}}'}(B(S') \,|\, \mathbf{S}(S')) &\leq \mathrm{Guess}^{\mathbf{A}'}(B(S') \,|\, \mathbf{S}(S')) + \\
&\quad + 2 \cdot d((U_1, \ldots, U_m), (\widetilde{U}_1, \ldots, \widetilde{U}_m)) \\
&\leq \mathrm{Guess}^{\mathbf{A}'}(B(S') \,|\, \mathbf{S}(S')) + 2m(\zeta_1 + \zeta_2).
\end{aligned}$$

To conclude the proof, we notice that an adversary at least as good as $\mathbf{A}'$ can be implemented in time $t' + (m-1) \cdot t_{A_{(\mathbf{S},B)}}(q, s_{A_{(\mathbf{S},B)}}(q \cdot \psi_{\mathsf{hc}})) \leq t/\varphi_{\mathsf{hc}}$ (by fixing optimally chosen states $\sigma_1, \ldots, \sigma_m$), and by Equation 3 this adversary achieves advantage at most γ. □

C Full Proof of the Hardcore Lemma

C.1 High-Level Description of the Proof of Theorem 1

In this section, we present a full proof of the hardcore lemma for system-bit pairs (Theorem 1). At a high level, our proof follows Holenstein's proof of the tight *uniform*[10] hardcore lemma given in [21] (which, in turn, was inspired by one of the proofs given by Impagliazzo [24]). However, there are major differences (and difficulties) due to the fact that we are considering discrete interactive systems: In particular, for

[10] Interestingly, no simpler approach appears to work despite the statement being proven being non-uniform.

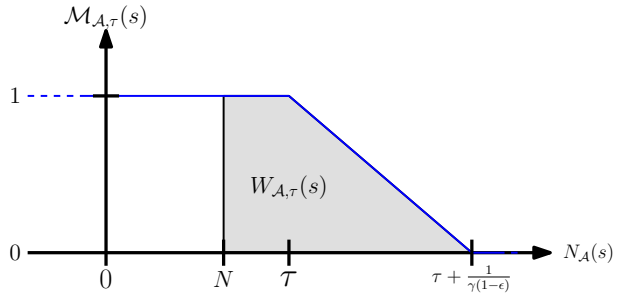

Fig. 1. Measure $\mathcal{M}_{\mathcal{A},\tau}(s)$ as a function of $N_{\mathcal{A}}(s)$. In particular, the gray area represents $W_{\mathcal{A},\tau}(s)$ given that $N_{\mathcal{A},\tau}(s) = N$.

a given $s \in \mathcal{S}$ the behavior of the system $\mathbf{S}(s)$ can be randomized (but stateless). Furthermore, we need to take into account the fact that we prove the existence of a measure for which a sufficiently good state sampler exists. Also, unlike the case of random variables, it is crucial that no quantity in the proof depends on the size of states S, as we cannot generally assume them to be small. Another final difference (which turns out to be rather easy to handle at the technical level) is that, contrary to the traditional hardcore lemma, we allow for the distribution P_S of the state S to be not uniformly distributed.

Initial definitions. For a collection $\mathcal{A}$ of (deterministic) adversaries, we define

$$N_{\mathcal{A}}(s) := \sum_{\mathbf{A} \in \mathcal{A}} \mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s)),$$

for all $s \in \mathcal{S}$. Note that $|N_{\mathcal{A}}(s)| \leq |\mathcal{A}|$. In particular, $N_{\mathcal{A}}(s)$ provides a quantitative measure of the overall quality of the collection $\mathcal{A}$ in guessing $B(s)$ given access to $\mathbf{S}(s)$. For instance, $N_{\mathcal{A}}(s)/|\mathcal{A}|$ gives the advantage of a randomly chosen $\mathbf{A} \in \mathcal{A}$ in guessing $B(s)$ while interacting with $\mathbf{S}(s)$. Moreover, for a non-negative integer $\tau \in \mathbb{N}$, we define the measure $\mathcal{M}_{\mathcal{A},\tau} : \mathcal{S} \to [0,1]$ associated with $\mathcal{A}$ and τ such that

$$\mathcal{M}_{\mathcal{A},\tau}(s) := \begin{cases} 1 & \text{if } N_{\mathcal{A}}(s) \leq \tau, \\ 1 - (N_{\mathcal{A}}(s) - \tau)\gamma(1-\varepsilon) & \text{if } \tau < N_{\mathcal{A}}(s) < \tau + \frac{1}{\gamma(1-\varepsilon)}, \\ 0 & \text{if } N_{\mathcal{A}}(s) \geq \tau + \frac{1}{\gamma(1-\varepsilon)}. \end{cases}$$

The value $\mathcal{M}_{\mathcal{A},\tau}(s)$ is plotted as a function of $N_{\mathcal{A}}(s)$ on Figure 1. In other words, elements $s \in \mathcal{S}$ on which many adversaries in $\mathcal{A}$ are bad are given more weight by the measure, whereas those for which many adversaries are good are given no measure. Note that $\mathcal{M}_{\emptyset,0}$ is the all-one measure, i.e., in particular the associated distribution $\mathsf{P}_{\mathcal{M}}$ equals P_S. Hence, $\mu(\mathcal{M}_{\emptyset,0}) \geq 1 - \varepsilon$.

Structure of the proof. The proof is by contradiction. It starts by assuming that Theorem 1 is false. In other words, we assume that for some γ, ζ_1 and ζ_2 the following holds:

> **Assumption ¬HC.** For all measures $\mathcal{M} : \mathcal{S} \to [0,1]$ with $\mu(\mathcal{M}) \geq 1 - \varepsilon$ such that a (ζ_1, ζ_2)-sampler for $\mathcal{M}$ and $A_{(\mathbf{S},B)}$ with length[11] $\ell := s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}} \cdot q')$ exists, there exists an adversary $\mathbf{A}$ with time complexity $t' := t/\varphi_{\mathsf{hc}}$, making $q' := q/\varphi_{\mathsf{hc}}$ queries, and such that
>
> $$\mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma.$$
>
> where $S' \xleftarrow{\$} \mathcal{M}$.

The core of proof consists of the procedure FindCollection, specified in Figure 2, which, under Assumption ¬HC, outputs a collection of *deterministic* adversaries $\mathcal{A}$ such that $|\mathcal{A}| \leq 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ and such that (at least) one of the following two properties is satisfied:

[11] In the case where $(\mathbf{S}(s), B(s))$ is deterministic for all $s \in \mathcal{S}$, we replace ψ_{hc} by $7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ and let $\zeta_1 = \zeta_2 = 0$.

Procedure GoodEnough($\mathcal{A}$): // collection of adversaries $\mathcal{A}$

1: **if** $\mathcal{A} = \emptyset$ **then return** `false`
2: $p := \min_{\mathcal{E}:\mathsf{P}[\mathcal{E}]=1-\varepsilon} \mathsf{P}\left[\mathbf{A}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}, \mathcal{E}\right]$
3: $\rho := \mathsf{P}[N_{\mathcal{A}}(S) > \gamma \cdot |\mathcal{A}|]$
4: **if** $p \geq \frac{1}{2} + \frac{1}{4\cdot(1-\varepsilon)\cdot|\mathcal{A}|\gamma} \vee \rho \geq 1 - \frac{1-\varepsilon}{4}$ **then**
5: **return** `true`
6: **else**
7: **return** `false`

Procedure FindCollection: // no input

1: $\mathcal{A} := \emptyset$, $\tau := 0$
2: **while not** GoodEnough($\mathcal{A}$) **do**
3: **if** $\mu(\mathcal{M}_{\mathcal{A},\tau}) < 1 - \varepsilon$ **then**
4: $\tau := \tau + 1$
5: get deterministic $\mathbf{A} \in \mathcal{A}_{t',q'}$ such that // see text

$$\mathsf{P}\left[\mathbf{A}(\mathbf{S}(S')) = B(S') \,\middle|\, S' \xleftarrow{\$} \mathcal{M}_{\mathcal{A},\tau}\right] \geq \frac{1+\gamma}{2}$$

6: $\mathcal{A} := \mathcal{A} \cup \{\mathbf{A}\}$
7: **return** $\mathcal{A}$

Fig. 2. Description of the procedure FindCollection used in the proof of Theorem 1.

(A) $\mathsf{P}[N_{\mathcal{A}}(S) > \gamma \cdot |\mathcal{A}|] \geq 1 - \frac{1}{4}(1-\varepsilon)$
(B) For all events $\mathcal{E}$ (defined by a probability distribution $\mathsf{P}_{\mathcal{E}|S}$) such that $\mathsf{P}[\mathcal{E}] = 1 - \varepsilon$, we have

$$\mathsf{P}\left[\mathbf{A}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}, \mathcal{E}\right] \geq \frac{1}{2} + \frac{1}{4 \cdot (1-\varepsilon) \cdot |\mathcal{A}|\gamma}.$$

We show how to derive, from a collection $\mathcal{A}$ satisfying one of the two conditions, corresponding adversaries $\mathbf{A}^{(A)}$ and $\mathbf{A}^{(B)}$ with time complexity t and making q queries which guess $B(S)$ given access to $\mathbf{S}(S)$ with advantage greater than ε, contradicting the assumed hardness of $(\mathbf{S}, B)$.

Without loss of generality, throughout this proof we assume that $B(s)$ is determined by s.

The procedure FindCollection. The procedure FindCollection (cf. Figure 2) starts with $\mathcal{A} := \emptyset$ and parameter $\tau = 0$, iteratively adds to $\mathcal{A}$ a *deterministic*[12] adversary $\mathbf{A}$ with time complexity t' and query complexity q' such that

$$\mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma,$$

for $S' \xleftarrow{\$} \mathcal{M}_{\mathcal{A},\tau}$. (Possibly, it first increments the value of τ by one to ensure that $\mu(\mathcal{M}_{\mathcal{A},\tau}) \geq 1 - \varepsilon$.) Finally, it terminates when the collection $\mathcal{A}$ satisfies (at least) one of conditions (A) or (B).

Note that the procedure FindCollection is to be seen as a process constructing a collection, where in particular adversaries at Line 5 are not found efficiently and explicitly, they are only guaranteed to exist by Assumption $\neg$HC.

[12] This can be assumed without loss of generality: Given a randomized adversary, an equally good deterministic adversary can be obtained by fixing its randomness to a good value.

Correctness of FindCollection. Note that in order for an adversary $\mathbf{A}$ as required in Line 5 to exist under Assumption $\neg\mathsf{HC}$, two conditions must be met immediately before executing Line 5: First, the measure $\mathcal{M}_{\mathcal{A},\tau}$ must have density at least $1-\varepsilon$. Second, there must exist a (ζ_1,ζ_2)-sampler for $\mathcal{M}_{\mathcal{A},\tau}$ and $A_{(\mathbf{S},B)}$ with length $\ell \le s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}}\cdot q')$. The following lemmas show that both conditions are always satisfied.

Lemma 2. *At every execution of Line 5 in* FindCollection*, the condition $\mu(\mathcal{M}_{\mathcal{A},\tau}) \ge 1-\varepsilon$ holds.*

Proof. The proof goes by induction. Clearly, it holds the first time Line 5 is executed, since $\mu(\mathcal{M}_{\emptyset,0}) = 1 \ge 1-\varepsilon$, and hence τ is not increased. Assume that the claim of the lemma holds up to the i-th iteration of the **while**-loop, and consider the beginning of the $(i+1)$-st iteration: Assume that $\mu(\mathcal{M}_{\mathcal{A},\tau}) < 1-\varepsilon$. This means that there exists $\mathcal{A}$ and $\mathbf{A}$ such that $\mu(\mathcal{M}_{\mathcal{A},\tau}) \ge 1-\varepsilon$, and $\mu(\mathcal{M}_{\mathcal{A}\cup\{A\},\tau}) < 1-\varepsilon$. Note that for all $s \in \mathcal{S}$ we have $N_{\mathcal{A}\cup\{\mathbf{A}\}}(s) \le N_{\mathcal{A}}(s)+1$, which in particular implies that $\mathcal{M}_{\mathcal{A}\cup\{\mathbf{A}\},\tau+1}(s) \ge \mathcal{M}_{\mathcal{A},\tau}(s)$ for all $s \in \mathcal{S}$, and hence $\mu(\mathcal{M}_{\mathcal{A}\cup\{\mathbf{A}\},\tau+1}) \ge \mu(\mathcal{M}_{\mathcal{A},\tau}) \ge 1-\varepsilon$. □

Lemma 3. *For all ζ_1,ζ_2, all collections of adversaries $\mathcal{A}$ and all $\tau \in \mathbb{N}$, there exists a (ζ_1,ζ_2)-sampler for $\mathcal{M}_{\mathcal{A},\tau}$ and $A_{(\mathbf{S},B)}$ with length $\ell := s_{A_{(\mathbf{S},B)}}(\psi\cdot q')$, where*

$$\psi := \frac{4}{\zeta_1^2}\cdot|\mathcal{A}|^2\cdot\gamma^2(1-\varepsilon)^2\cdot\ln\left(\frac{2}{\zeta_2}\right).$$

Furthermore, if $(\mathbf{S}(S),B(S))$ is deterministic for every value taken by S, an error-free state sampler for $\mathcal{M}_{\mathcal{A},\tau}$ and $A_{(\mathbf{S},B)}$ for length $\ell := s_{A_{(\mathbf{S},B)}}(|\mathcal{A}|\cdot q')$ exists.

A proof of Lemma 3 is postponed to Section C.2.

Termination of FindCollection. We prove the following lemma in Section C.3: It upper bounds the number of iterations after which FindCollection returns an appropriate collection.

Lemma 4. *Under the assumption that a good adversary is always found at Line 5, the procedure* FindCollection *terminates after at most $7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$ iterations of the* **while***-loop, outputting a collection $\mathcal{A}$ of deterministic adversaries, each with time complexity t' and query complexity q', such that $|\mathcal{A}| \le 7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$ and (at least) one of the two conditions (A) and (B) is satisfied.*

The final adversaries. In order to conclude the proof, we need to show that such a collection yields an adversary accessing $\mathbf{S}(B)$ and contradicting the assumed hardness of guessing B. This is implied by the following two lemmas, whose proofs are deferred to Section C.4.

Lemma 5. *Let $\mathcal{A}$ be a set of deterministic adversaries, each with time complexity t' and query complexity q', satisfying Condition (A) and such that $|\mathcal{A}| \le 7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$. Then, there exists an adversary $\mathbf{A}^{(A)}$ with time and query complexities*

$$t_1 := \frac{4}{\gamma^2}\cdot\ln\left(\frac{4}{1-\varepsilon}\right)\cdot t' \quad \text{and} \quad q_1 := \frac{4}{\gamma^2}\cdot\ln\left(\frac{4}{1-\varepsilon}\right)\cdot q'$$

such that $\mathrm{Guess}^{\mathbf{A}^{(A)}}(B\,|\,\mathbf{S}) > \varepsilon$.

Lemma 6. *Let $\mathcal{A}$ be a set of deterministic adversaries, each with time complexity t' and query complexity q', satisfying Condition (B) such that $|\mathcal{A}| \le 7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$. Then, there exists an adversary $\mathbf{A}^{(B)}$ with time and query complexities*

$$t_2 := \frac{6400}{\gamma^2(1-\varepsilon)^4}\cdot\ln\left(\frac{160}{\gamma(1-\varepsilon)^3}\right)\cdot t' \quad \text{and} \quad q_2 := \frac{6400}{\gamma^2(1-\varepsilon)^4}\cdot\ln\left(\frac{160}{\gamma(1-\varepsilon)^3}\right)\cdot q'$$

such that $\mathrm{Guess}^{\mathbf{A}^{(B)}}(B\,|\,\mathbf{S}) > \varepsilon$.

Sampler $\mathsf{O}_{\mathcal{A},\tau}$: // collection of adversaries $\mathcal{A}$, $\tau \in \mathbb{N}$

$b \xleftarrow{\$} A_{(\mathbf{S},B)}$ // run $A_{(\mathbf{S},B)}$ to get the bit B
$\sigma :=$ state of $A_{(\mathbf{S},B)}$ after outputting B
$\psi := \frac{4}{\zeta_1^2} \cdot |\mathcal{A}|^2 \cdot \gamma^2(1-\varepsilon)^2 \cdot \ln\left(\frac{2}{\zeta_2}\right)$
for all $i := 1, \ldots, \psi$ **do**
  $\mathbf{A} \xleftarrow{\$} \mathcal{A}$
  $v_i := \mathbf{A}(A_{(\mathbf{S},B)}[\sigma]) \oplus B \oplus 1$ // 1 if correct, 0 else
  $\sigma :=$ last state of $A_{(\mathbf{S},B)}$ after interacting with $\mathbf{A}$
$\overline{N} := 2|\mathcal{A}| \cdot \left(\frac{1}{\psi}\sum_{i=1}^{\psi} v_i - \frac{1}{2}\right)$
$z := \max\{0, \min\{1, 1 - (\overline{N} - \tau)\gamma(1-\varepsilon)\}\}$
return (σ, z)

Sampler $\mathsf{O}'_{\mathcal{A},\tau}$: // collection of adversaries $\mathcal{A}$, $\tau \in \mathbb{N}$

$s \xleftarrow{\$} \mathsf{P}_S$
$\psi := \frac{4}{\zeta_1^2} \cdot |\mathcal{A}|^2 \cdot \gamma^2(1-\varepsilon)^2 \cdot \ln\left(\frac{2}{\zeta_2}\right)$
for all $i := 1, \ldots, \psi$ **do**
  $\mathbf{A} \xleftarrow{\$} \mathcal{A}$
  $v_i(s) := \mathbf{A}(\mathbf{S}(s)) \oplus B(s) \oplus 1$ // 1 if correct, 0 else
$\overline{N}(s) := 2|\mathcal{A}| \cdot \left(\frac{1}{\psi}\sum_{i=1}^{\psi} v_i(s) - \frac{1}{2}\right)$
$z(s) := \max\{0, \min\{1, 1 - (\overline{N}(s) - \tau)\gamma(1-\varepsilon)\}\}$
return $(s, z(s))$

Fig. 3. Top: Construction of the state sampler $\mathsf{O}_{\mathcal{A},\tau}$ for $\mathcal{M}_{\mathcal{A},\tau}$ and $A_{(\mathbf{S},B)}$. Bottom: Idealized state sampler $\mathsf{O}'_{\mathcal{A},\tau}$ used in the proof of Lemma 3.

Concluding the proof. To conclude the proof, we observe that by Assumption $\neg\mathsf{HC}$, Lemmas 2, 3, and 4, $\mathsf{FindCollection}$ outputs a collection $\mathcal{A}$ with $|\mathcal{A}| \leq 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ satisfying one of conditions (A) and (B). Then either Lemma 5 or Lemma 6 yields an adversary contradicting the assumed hardness of $(\mathbf{S}, B)$, as by definition of φ_{hc}, for the time complexity we have $t_1 \leq t$ and $t_2 \leq t$, as well as $q_1 \leq q$ and $q_2 \leq q$.

C.2 Existence of State Samplers (Proof of Lemma 3)

The key observation is that given the collection $\mathcal{A}$ as well as *black-box* access to the system-bit pair $(\mathbf{S}(S), B(S))$, it is possible to compute a sufficiently good estimate of $N_{\mathcal{A}}(S)$ (and thus of $\mathcal{M}_{\mathcal{A},\tau}(S)$) by letting (randomly chosen) adversaries from $\mathcal{A}$ sequentially interact with $\mathbf{S}(S)$ and check whether their outputs equal B or not. In particular, we can use the algorithm $A_{(\mathbf{S},B)}$ in order to simulate $(\mathbf{S}(S), B(S))$. Here, for notational convenience, we assume that the algorithm $A_{(\mathbf{S},B)}$ first outputs the simulated bit B, and subsequently simulates $\mathbf{S}$ accordingly.

This is summarized by the random process $\mathsf{O}_{\mathcal{A},\tau}$, which outputs a valid state of $A_{(\mathbf{S},B)}$, and is described in Figure 3 on top. We additionally consider the *idealized* process $\mathsf{O}'_{\mathcal{A},\tau}$ (depicted in the lower part of Figure 3) which, instead of using the algorithm $A_{(\mathbf{S},B)}$, uses the actual system-bit pair $(\mathbf{S}(S), B(S))$ for a random $S \xleftarrow{\$} \mathsf{P}_S$, and produces an estimate $Z(S)$ of $\mathcal{M}_{\mathcal{A},\tau}(S)$ as in $\mathsf{O}_{\mathcal{A},\tau}$, and

finally outputs the pair $(S, Z(S))$. (From now on, we use the convention that variables appearing within the process descriptions are upper-case when seen as random variables.) Obviously, $(A_{(\mathbf{S},B)}[\Sigma], Z)$ for $(\Sigma, Z) \xleftarrow{\$} \mathsf{O}_{\mathcal{A},\tau}$ and $(\mathbf{S}(S), B(S), Z(S))$ for $(S, Z) \xleftarrow{\$} \mathsf{O}'_{\mathcal{A},\tau}$ are exactly the same, that is, the equivalence

$$(A_{(\mathbf{S},B)}[\Sigma], Z) \equiv (\mathbf{S}(S), B(S), Z(S))$$

holds, due to the fact that $A_{(\mathbf{S},B)}$ perfectly simulates $(\mathbf{S}, B)$. Therefore, in order to finish the proof of property (ii) in Definition 1, it suffices to analyze the quality of the estimate $Z(S)$ output by $\mathsf{O}'_{\mathcal{A},\tau}$: Let us fix $s \in \mathcal{S}$ and consider the random variable $V(s)$ obtained by sampling a random $\mathbf{A} \xleftarrow{\$} \mathcal{A}$ and then letting $V(s) := 1$ if $\mathbf{A}(\mathbf{S}(s)) = B(s)$, and $V(s) := 0$ otherwise. Then, note that

$$\mathsf{E}[V(s)] = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{A} \in \mathcal{A}} \mathsf{P}[\mathbf{A}(\mathbf{S}(s)) = B(s)] = \frac{1}{|\mathcal{A}|} \sum_{\mathbf{A} \in \mathcal{A}} \left(\frac{1}{2} + \frac{\mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s))}{2} \right) = \frac{1}{2} + \frac{N_{\mathcal{A}}(s)}{2|\mathcal{A}|}.$$

Since $(\mathbf{S}(S), B(S))$ is cc-stateless, the pair $(\mathbf{S}(s), B(s))$ is stateless (and in particular, $\mathbf{S}(s)$ is stateless), and the random variables $V_1(s), \ldots, V_\psi(s)$ are independent with the same distribution as $V(s)$. In particular, $\mathsf{E}[V_i(s)] = \mathsf{E}[V(s)]$ for all $i \in \{1, \ldots, \psi\}$. The event that $Z(s)$ is a bad estimate, i.e., that $|Z(s) - \mathcal{M}_{\mathcal{A},\tau}(s)| > \zeta_1$, implies in particular $\left|\overline{N}(s) - N_{\mathcal{A}}(s)\right| > \frac{\zeta_1}{\gamma(1-\varepsilon)}$ and in turn $\left|\psi^{-1} \sum_{i=1}^{\psi} V_i(s) - \mathsf{E}[V(s)]\right| > \frac{\zeta_1}{2|\mathcal{A}|\gamma(1-\varepsilon)}$. We conclude by applying Hoeffding's inequality (Lemma 1),

$$\begin{aligned}
\mathsf{P}\left[|Z(s) - \mathcal{M}_{\mathcal{A},\tau}(s)| > \zeta_1\right] &\leq \mathsf{P}\left[\left|\overline{N}(s) - N_{\mathcal{A}}(s)\right| > \frac{\zeta_1}{\gamma(1-\varepsilon)}\right] \\
&\leq \mathsf{P}\left[\left|\frac{1}{\psi} \sum_{i=1}^{\psi} V_i(s) - \mathsf{E}[V(s)]\right| > \frac{\zeta_1}{2|\mathcal{A}|\gamma(1-\varepsilon)}\right] \leq 2e^{-\frac{\zeta_1^2}{4|\mathcal{A}|^2\gamma^2(1-\varepsilon)^2}\psi} = \zeta_2,
\end{aligned}$$

by the choice of ψ.

In the case where $(\mathbf{S}(S), B(S))$ is *deterministic* given S, then a much simpler (and error-free) state sampler can be built: Namely, it is sufficient to let *all* adversaries $\mathbf{A} \in \mathcal{A}$ interact (sequentially) with $\mathbf{S}$ (simulated by $A_{(\mathbf{S},B)}$), rather then choosing them randomly, and set v_i as above. Performing the analysis as above (via the ideal process $\mathsf{O}'_{\mathcal{A},\tau}$), we see that $\overline{N}(s)$ indeed equals $N_{\mathcal{A}}(s)$, and thus the estimate is always correct.

C.3 The Procedure FindCollection Terminates (Proof of Lemma 4)

Similarly to the proofs in [24, 21], we define for all collections of adversaries $\mathcal{A}$, all $\tau \in \mathbb{N}$ and all $s \in \mathcal{S}$ the quantity

$$W_{\mathcal{A},\tau}(s) := \begin{cases} \tau - N_{\mathcal{A}}(s) + \frac{1}{2\gamma(1-\varepsilon)} & \text{if } N_{\mathcal{A}}(s) \leq \tau, \\ \frac{\mathcal{M}_{\mathcal{A},\tau}(s)}{2} \left[\tau + \frac{1}{\gamma(1-\varepsilon)} - N_{\mathcal{A}}(s)\right] & \text{if } \tau < N_{\mathcal{A}}(s) < \tau + \frac{1}{\gamma(1-\varepsilon)}, \\ 0 & \text{else.} \end{cases}$$

This quantity is to be interpreted as the area under $\mathcal{M}_{\mathcal{A},\tau}(s)$, starting from $N_{\mathcal{A}}(s)$, as depicted in Figure 1. We also define the weighted average of $W_{\mathcal{A},\tau}(s)$ as

$$W(\mathcal{A}, \tau) := \mathsf{E}[W_{\mathcal{A},\tau}(S)] = \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot W_{\mathcal{A},\tau}(s).$$

In the following, we study the evolution of this quantity in order to prove termination for the execution of FindCollection. In particular, there are two causes for a change in the value of W at each iteration:

First, an adversary $\mathbf{A}$ is always added to $\mathcal{A}$. Second, the value τ is possibly incremented by one before the adversary is added. The following two claims analyze these two cases. We point out that the analysis is more involved than the one of [21], due to the different termination condition and the underlying distribution being arbitrary and not necessarily uniform.

The first claim shows that if we add an adversary to $\mathcal{A}$ in Line 5, then $W(\mathcal{A} \cup \{\mathbf{A}\}, \tau)$ is smaller than $W(\mathcal{A}, \tau)$ by at least $\frac{\gamma(1-\varepsilon)}{2}$.

Claim. Let $\mathcal{A}$ be a collection of adversaries and let $\tau \in \mathbb{N}$ be such that $\mu(\mathcal{M}_{\mathcal{A},\tau}) \geq 1 - \varepsilon$. Moreover, let $\mathbf{A} \notin \mathcal{A}$ be an adversary such that

$$\mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma$$

for $S' \xleftarrow{\$} \mathcal{M}_{\mathcal{A},\tau}$. Then $W(\mathcal{A} \cup \{\mathbf{A}\}, \tau) \leq W(\mathcal{A}, \tau) - \frac{\gamma(1-\varepsilon)}{2}$.

Proof. For all $s \in \mathcal{S}$, the definition of $N_{\mathcal{A}}$ yields $N_{\mathcal{A}\cup\{\mathbf{A}\}}(s) = N_{\mathcal{A}}(s) + \mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s))$, and consequently we obtain

$$W_{\mathcal{A}\cup\{\mathbf{A}\},\tau}(s) \leq W_{\mathcal{A},\tau}(s) - \mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s)) \cdot \mathcal{M}_{\mathcal{A},\tau}(s) + \frac{\gamma(1-\varepsilon)}{2}.$$

Averaging over the choice of S yields

$$\begin{aligned} W(\mathcal{A} \cup \{\mathbf{A}\}, \tau) &= \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot W_{\mathcal{A}\cup\{\mathbf{A}\},\tau}(s) \\ &\leq W(\mathcal{A}, \tau) + \frac{\gamma(1-\varepsilon)}{2} - \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s)) \cdot \mathcal{M}_{\mathcal{A},\tau}(s). \end{aligned}$$

However, note that if $S' \xleftarrow{\$} \mathcal{M}_{\mathcal{A},\tau}$, then by definition $\mathsf{P}_{S'}(s) := \mathsf{P}_S(s) \cdot \mathcal{M}_{\mathcal{A},\tau}(s)/\mu(\mathcal{M}_{\mathcal{A},\tau})$, and hence

$$\sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathrm{Guess}^{\mathbf{A}}(B(s) \,|\, \mathbf{S}(s)) \cdot \mathcal{M}_{\mathcal{A},\tau}(s) = \mu(\mathcal{M}_{\mathcal{A},\tau}) \cdot \mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')),$$

and the claim follows from $\mu(\mathcal{M}_{\mathcal{A},\tau}) \geq 1 - \varepsilon$ and $\mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma$. □

The following claim additionally shows bounds on the variation of $W(\mathcal{A}, \tau)$ upon incrementing τ.

Claim. Let $\eta > 0$, let $\mathcal{A}$ be a collection of adversaries, and let $\tau \in \mathbb{N}$. Further assume that $\mathsf{P}\left[N_{\mathcal{A}}(S) > \eta\right] < 1 - \frac{1}{4}(1-\varepsilon)$ and $\mu(\mathcal{M}_{\mathcal{A},\tau}) < 1 - \varepsilon$. Then,

$$W(\mathcal{A}, \tau+1) \leq \begin{cases} W(\mathcal{A}, \tau) + (1-\varepsilon) + \frac{\gamma(1-\varepsilon)}{2} - \frac{\gamma(1-\varepsilon)^2}{8} & \text{if } \tau > \eta, \\ W(\mathcal{A}, \tau) + (1-\varepsilon) + \frac{\gamma(1-\varepsilon)}{2} & \text{if } \tau \leq \eta. \end{cases}$$

Proof. For all $s \in \mathcal{S}$ we have

$$W_{\mathcal{A},\tau+1}(s) \leq W_{\mathcal{A},\tau}(s) + \mathcal{M}_{\mathcal{A},\tau}(s) + \frac{\gamma(1-\varepsilon)}{2}.$$

In fact, if $N_\tau(s) \leq \tau$, then we even have $W_{\mathcal{A},\tau+1}(s) \leq W_{\mathcal{A},\tau}(s) + \mathcal{M}_{\mathcal{A},\tau}(s)$. From this we can infer

$$\begin{aligned} W(\mathcal{A}, \tau+1) &= \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot W_{\mathcal{A},\tau+1}(s) \\ &\leq W(\mathcal{A}, \tau) + \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathcal{M}_{\mathcal{A},\tau}(s) + \sum_{s : N_{\mathcal{A}}(s) > \tau} \mathsf{P}_S(s) \cdot \frac{\gamma(1-\varepsilon)}{2} \\ &\leq W(\mathcal{A}, \tau) + (1-\varepsilon) + \mathsf{P}[N_{\mathcal{A}}(S) > \tau] \cdot \frac{\gamma(1-\varepsilon)}{2}. \end{aligned}$$

If $\tau > \eta$, then $\mathsf{P}[N_{\mathcal{A}}(S) > \tau] \leq \mathsf{P}[N_{\mathcal{A}}(S) > \eta] \leq 1 - \frac{1-\varepsilon}{4}$, and thus

$$\begin{aligned} W(\mathcal{A}, \tau+1) &\leq W(\mathcal{A}, \tau) + (1-\varepsilon) + \left(1 - \frac{1-\varepsilon}{4}\right) \cdot \frac{\gamma(1-\varepsilon)}{2} \\ &= W(\mathcal{A}, \tau) + (1-\varepsilon) + \frac{\gamma(1-\varepsilon)}{2} - \frac{\gamma(1-\varepsilon)^2}{8} \end{aligned}$$

whereas if $\tau \leq \eta$, we can only conclude that $W(\mathcal{A}, \tau+1) \leq W(\mathcal{A}, \tau) + (1-\varepsilon) + \frac{\gamma(1-\varepsilon)}{2}$. □

In the following, let $\mathcal{A}(i)$ and $\tau(i)$ be the values of $\mathcal{A}$ and τ at the beginning of the i-th iteration, i.e., when GoodEnough is invoked. In particular, $|\mathcal{A}(i)| = i - 1$. We now show that under the assumption[13] that $\mathsf{P}\left[N_{\mathcal{A}(i)}(S) > \gamma \cdot (i-1)\right] < 1 - \frac{1}{4}(1-\varepsilon)$ holds, then FindCollection terminates satisfying Condition (B). To this aim, we define the *potential function* π such that the potential at the beginning of the i-th iteration is

$$\pi(i) := W(\mathcal{A}(i), \tau(i)) - \tau(i) \cdot (1-\varepsilon).$$

Note that initially $\pi(1) = W(\emptyset, 0) - 0 \cdot (1-\varepsilon) = W(\emptyset, 0) = \frac{1}{2\gamma(1-\varepsilon)}$. The above two claims imply the following for all i's:

$$\pi(i+1) \leq \begin{cases} \pi(i) - \frac{\gamma(1-\varepsilon)}{2} & \text{if } \tau(i) = \tau(i+1) \\ \pi(i) & \text{if } \tau(i+1) = \tau(i) + 1 \text{ and } \tau(i) \leq \gamma \cdot (i-1) \\ \pi(i) - \frac{\gamma(1-\varepsilon)^2}{8} & \text{if } \tau(i+1) = \tau(i) + 1 \text{ and } \tau(i) > \gamma \cdot (i-1). \end{cases}$$

In particular, it is important to note that the value $\pi(i)$ *never* increases. The following claim proves that it also decreases sufficiently fast, reaching a negative value after $7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ iterations. Below, we show that in this case, the corresponding collection $\mathcal{A}$ satisfies Condition (B).

Claim. For $\lambda := 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ we have $\pi(\lambda) < 0$.

Proof. Assume, towards a contradiction, that the claim is wrong, i.e., we have an execution of FindCollection such that after $\lambda := 7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ steps we have $\pi(\lambda) \geq 0$. Consider the partition $\mathcal{I}^{=} \cup \mathcal{I}^{+} = \{1, \ldots, \lambda\}$, where $\mathcal{I}^{=}$ are the iterations i where $\tau(i)$ was not increased, whereas in $\mathcal{I}^{+}$ the value of $\tau(i)$ is increased. We also let $\mathcal{I}^{*} := \{(1-\gamma)^{-1} |\mathcal{I}^{=}| + 2, \ldots, \lambda\}$. Hence, for all $i \in \mathcal{I}^{*}$ we have

$$\tau(i) \geq i - 1 - |\mathcal{I}^{=}| > i - 1 - (1-\gamma)(i-1) = \gamma(i-1),$$

since $\tau(i) \geq i - 1 - |\mathcal{I}^{=}|$ for all $i \geq 1$, whereas $(1-\gamma)^{-1} |\mathcal{I}^{=}| < i - 1$ for all $i \in \mathcal{I}^{*}$. Finally, note that

$$|\mathcal{I}^{*} \cap \mathcal{I}^{+}| = |\mathcal{I}^{*} \setminus \mathcal{I}^{=}| \geq |\mathcal{I}^{*}| - |\mathcal{I}^{=}| \geq \lambda - (1-\gamma)^{-1} |\mathcal{I}^{=}| - 1 - |\mathcal{I}^{=}| = \lambda - \frac{2-\gamma}{1-\gamma} |\mathcal{I}^{=}| - 1.$$

Moreover, we have $|\mathcal{I}^{=}| \leq \gamma^{-2}(1-\varepsilon)^{-2} < \gamma^{-2}(1-\varepsilon)^{-3}$, as otherwise this would contradict $\pi(i) \geq 0$ by the first claim. Using $\gamma \leq \frac{1}{2}$, we obtain $|\mathcal{I}^{*} \cap \mathcal{I}^{+}| > 4\gamma^{-2}(1-\varepsilon)^{-3}$. However, in every step $i \in \mathcal{I}^{*} \cap \mathcal{I}^{+}$, by the above claim the value of π has been decreased by at least $\frac{\gamma(1-\varepsilon)^2}{8}$, and this implies $\pi(\lambda) < 0$. □

Finally, we show that if $\pi(i) < 0$, then the collection $\mathcal{A}(i)$ satisfies Condition (B), and thus termination is achieved.

[13] If this assumption is not satisfied, then GoodEnough will detect this, and return `true`, which in particular implies termination of FindCollection, which returns the collection $\mathcal{A}(i)$ satisfying Condition (A).

Adversary $\mathbf{A}^{(A)}$: // expecting to interact with system $\mathbf{S}(s)$

$\varphi_A := \frac{4}{\gamma^2} \cdot \ln\left(\frac{4}{1-\varepsilon}\right)$
for all $i := 1, \ldots, \varphi_A$ **do**
  $\mathbf{A} \xleftarrow{\$} \mathcal{A}$
  $v_i \xleftarrow{\$} \mathbf{A}(\mathbf{S}(s))$
return $\mathsf{majority}\,\{v_i \,:\, i \in \{1, \ldots, \varphi_A\}\}$

Fig. 4. Adversary $\mathbf{A}^{(A)}$.

Claim. If $\pi(i) < 0$, then $\mathcal{A}(i)$ satisfies

$$\mathsf{P}\left[\mathbf{A}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}(i), \mathcal{E}\right] \geq \frac{1}{2} + \frac{1}{4 \cdot (1-\varepsilon) \cdot |\mathcal{A}(i)| \cdot \gamma}$$

for all events $\mathcal{E}$ (defined by $\mathsf{P}_{\mathcal{E}|S}$) such that $\mathsf{P}[\mathcal{E}] = 1 - \varepsilon$.

Proof. Let $\tau := \tau(i)$ and $\mathcal{A} := \mathcal{A}(i)$. Also let $\mathcal{E}$ be an arbitrary event (defined by $\mathsf{P}_{\mathcal{E}|S}$) such that $p := \mathsf{P}[\mathcal{E}] = 1 - \varepsilon$. Then, by the definition of $N_{\mathcal{A}}$,

$$\mathsf{P}\left[\mathbf{A}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}, S \in \mathcal{E}\right] = \frac{1}{2} + \frac{1}{2(1-\varepsilon)|\mathcal{A}|} \cdot \sum_{s \in \mathcal{S}} \mathsf{P}_{S\mathcal{E}}(s) \cdot N_{\mathcal{A}}(s).$$

Note that for all $s \in \mathcal{S}$ it is easy to verify that

$$W_{\mathcal{A},\tau}(s) \geq \tau + \frac{1}{2\gamma(1-\varepsilon)} - N_{\mathcal{A}}(s),$$

and therefore

$$\begin{aligned}\sum_{s \in \mathcal{S}} \mathsf{P}_{S\mathcal{E}}(s) \cdot N_{\mathcal{A}}(s) &\geq \sum_{s \in \mathcal{S}} \mathsf{P}_{S\mathcal{E}}(s) \cdot \left[\tau + \frac{1}{2\gamma(1-\varepsilon)} - W_{\mathcal{A},\tau}(s)\right] \\ &\geq \mathsf{P}[\mathcal{E}] \cdot \left(\tau + \frac{1}{2\gamma(1-\varepsilon)}\right) - \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) W_{\mathcal{A},\tau}(s) \geq (1-\varepsilon)\tau + \frac{1}{2\gamma} - W(\mathcal{A},\tau) \geq \frac{1}{2\gamma}.\end{aligned}$$

The lemma follows by substituting this into the above equation.

Therefore, after at most $7 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3} + 1$ executions of the **while**-loop, the procedure FindCollection outputs a set of adversaries which satisfies at least one of Condition (A) and Condition (B).

C.4 Constructing the Final Adversaries

Proof of Lemma 5. Recall that we assume that

$$\mathsf{P}[N_{\mathcal{A}}(S) > \gamma \cdot |\mathcal{A}|] \geq 1 - \tfrac{1}{4}(1-\varepsilon) > 1 - \tfrac{1}{2}(1-\varepsilon) = \tfrac{1+\varepsilon}{2}.$$

The adversary $\mathbf{A}^{(A)}$ proceeds as described in Figure 4.

Let us fix s such that $N_{\mathcal{A}}(s) > \gamma \cdot |\mathcal{A}|$. Then, we denote as $I(s) \in \{0,1\}$ the indicator random variable which is 1 if a randomly sampled $\mathbf{A} \xleftarrow{\$} \mathcal{A}$ interacting with $\mathbf{S}(s)$ is successful, outputting $B(s)$. (Recall

that $B(s)$ is fully determined by s.) Note that

$$\mathsf{E}\left[I(s)\right] = \mathsf{P}\left[\mathbf{A}(\mathbf{S}(s)) = B(s) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}\right] = \frac{1}{|\mathcal{A}|}\sum_{\mathbf{A}\in\mathcal{A}} \mathsf{P}\left[\mathbf{A}(\mathbf{S}(s)) = B(s)\right]$$
$$= \frac{1}{2} + \frac{1}{2\cdot|\mathcal{A}|}\sum_{\mathbf{A}\in\mathcal{A}} \mathrm{Guess}^{\mathbf{A}}(B(s)\,|\,\mathbf{S}(s)) = \frac{1}{2} + \frac{N_{\mathcal{A}}(s)}{2\cdot|\mathcal{A}|} > \frac{1+\gamma}{2}.$$

For $i = 1,\ldots,\varphi_A$, denote as $I_i(s)$ the random variable indicating whether v_i equals $B(s)$ when interacting with $\mathbf{S}(s)$: Since $\mathbf{S}(s)$ is stateless, the variables $I_1(s),\ldots,I_{\varphi_A}(s)$ are independent, with $\mathsf{E}\left[I_i(s)\right] = \mathsf{E}\left[I(s)\right]$ for all $i = 1,\ldots,\varphi_A$. We can use this (combined with Hoeffding's inequality) to derive an upper bound on the failure probability of $\mathbf{A}^{(A)}$ given access to $\mathbf{S}(s)$,

$$\mathsf{P}\left[\mathbf{A}^{(A)}(\mathbf{S}(s)) \neq B(s)\right] = \mathsf{P}\left[\frac{1}{\varphi_A}\sum_{i=1}^{\varphi_A} I_i(s) < \frac{1}{2}\right] \leq \mathsf{P}\left[\frac{1}{\varphi_A}\sum_{i=1}^{\varphi_A} I_i(s) < \mathsf{E}\left[I(s)\right] - \frac{\gamma}{2}\right] < e^{-\frac{\gamma^2}{4}\varphi_A} = \tfrac{1-\varepsilon}{4},$$

and therefore

$$\mathsf{P}[\mathbf{A}^{(A)}(\mathbf{S}(S)) = B(S)] > \mathsf{P}\left[N_{\mathcal{A}}(S) > \gamma\cdot|\mathcal{A}|\right]\cdot\left(1-\tfrac{1-\varepsilon}{4}\right) \geq \left(1-\tfrac{1-\varepsilon}{4}\right)\cdot\left(1-\tfrac{1-\varepsilon}{4}\right) \geq 1-\tfrac{1-\varepsilon}{2} = \tfrac{1+\varepsilon}{2}.$$

Note that we can fix the choice of φ_A adversaries $\mathbf{A}_1,\ldots,\mathbf{A}_{\varphi_A}$ for the random choices of $\mathbf{A}$ maximizing the success probability of the resulting adversary $\mathbf{A}^{(A)}$. The time and query complexity bounds t_1 and q_1 are then clear.

Proof of Lemma 6. We define the functions $\alpha, \alpha_1 : \mathcal{S} \to [-1,1]$ such that

$$\alpha(s) := 2\cdot\mathsf{P}\left[\mathbf{A}(\mathbf{S}(s)) = B(s) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}\right] - 1 \text{ and } \alpha_1(s) := 2\cdot\mathsf{P}\left[\mathbf{A}(\mathbf{S}(s)) = 1 \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}\right] - 1.$$

for all $s \in \mathcal{S}$. We order the elements of $\mathcal{S}$ (with positive probability mass) as $s_1, s_2, \ldots$ such that $\alpha(s_i) \leq \alpha(s_{i+1})$ for $i = 1, 2, \ldots$. Then, let i^* be the (unique) index such that $\sum_{j=1}^{i^*-1}\mathsf{P}_S(s_j) < 1-\varepsilon$, but $\sum_{j=1}^{i^*}\mathsf{P}_S(s_j) \geq 1-\varepsilon$. We define the event $\mathcal{E}$ such that

$$\mathsf{P}_{\mathcal{E}|S}(s_i) := \begin{cases} 1 & \text{if } i < i^* \\ \frac{(1-\varepsilon)-\sum_{j=1}^{i^*-1}\mathsf{P}_S(s_j)}{\mathsf{P}_S(s_{i^*})} & \text{if } i = i^* \\ 0 & \text{if } i > i^*. \end{cases}$$

It is easy to verify that $\mathsf{P}\left[\mathcal{E}\right] = 1-\varepsilon$. Further, let $\alpha^* := \alpha(s_{i^*})$. We hence have (recall that $|\mathcal{A}| \leq 7\cdot\gamma^{-2}\cdot(1-\varepsilon)^{-3}+1$)

$$\mathsf{P}\left[\mathbf{A}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}, \mathcal{E}\right] > \frac{1}{2} + \frac{1}{4(1-\varepsilon)|\mathcal{A}|\gamma} \geq \frac{1}{2} + \frac{\gamma(1-\varepsilon)^2}{40}.$$

Note that this implies that $\alpha^* > \frac{\gamma(1-\varepsilon)^2}{20}$, as otherwise this would contradict the above lower bound on the guessing probability of a randomly chosen $\mathbf{A}$ from the collection $\mathcal{A}$. We consider the adversary $\mathbf{A}^{(B)}$ specified in Figure 5.

In the following, we assume that the above adversary is run on $\mathbf{S}(s)$ for some fixed $s \in \mathcal{S}$. Denote as $V_1(s),\ldots,V_{\varphi_B}(s)$ and $\overline{\alpha_1}(s)$ the corresponding random variables. Consider the random variable $V(s)$ which is the output of a randomly chosen $\mathbf{A} \xleftarrow{\$} \mathcal{A}$ accessing $\mathbf{S}(s)$. Note that

$$\mathsf{E}\left[V(s)\right] = \mathsf{P}\left[\mathbf{A}(\mathbf{S}(s)) = 1 \,\middle|\, \mathbf{A} \xleftarrow{\$} \mathcal{A}\right] = \tfrac{1}{2} + \tfrac{\alpha_1(s)}{2}.$$

Adversary $\mathbf{A}^{(B)}$: // expecting to interact with system $\mathbf{S}(s)$

$\varphi_B := \frac{6400}{\gamma^2(1-\varepsilon)^4} \cdot \ln\left(\frac{160}{\gamma(1-\varepsilon)^3}\right)$

for all $i := 1, \ldots, \varphi_B$ **do**

$\mathbf{A} \xleftarrow{\$} \mathcal{A}$

$v_i \xleftarrow{\$} \mathbf{A}(\mathbf{S}(s))$

$\overline{\alpha_1} := 2 \cdot \left(\frac{1}{\varphi_B} \sum_{i=1}^{\varphi_B} v_i - 1\right)$

return 1 with probability $\max\left\{0, \min\left\{1, \frac{1}{2} + \frac{\overline{\alpha_1}}{2\left(\alpha^* - \frac{\gamma(1-\varepsilon)^2}{20}\right)}\right\}\right\}$

Fig. 5. Adversary $\mathbf{A}^{(B)}$.

In particular, since $\mathbf{S}(s)$ is stateless, the random variables $V_1(s), \ldots, V_{\varphi_B}(s)$ are independent, and all distributed according to $V(s)$. Let now $\mathcal{G}$ be the event that $|\overline{\alpha}_1(s) - \alpha_1(s)| \leq \frac{\gamma(1-\varepsilon)^2}{40}$: By Hoeffding's inequality,

$$\begin{aligned}\mathsf{P}\left[\overline{\mathcal{G}}\right] &= \mathsf{P}\left[|\overline{\alpha_1} - \alpha_1| > \frac{\gamma(1-\varepsilon)^2}{40}\right] \\ &= \mathsf{P}\left[\left|\frac{1}{\varphi_B}\sum_{i=1}^{\varphi_B} V_i(s) - \mathsf{E}\left[V(s)\right]\right| > \frac{\gamma(1-\varepsilon)^2}{80}\right] \leq 2e^{-\left(\frac{\gamma(1-\varepsilon)^2}{80}\right)^2 \varphi_B} = \frac{\gamma(1-\varepsilon)^3}{80}.\end{aligned}$$

For all $s \in \mathcal{S}$, define $\overline{\alpha}(s)$ to be $\overline{\alpha_1}(s)$ if $B(s) = 1$, and $-\overline{\alpha_1}(s)$ if $B(s) = 0$, and note that by inspection

$$\mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(s)) = B(s)\right] = \min\left\{1, \frac{1}{2} + \frac{\overline{\alpha}(s)}{2\left(\alpha^* - \frac{\gamma(1-\varepsilon)^2}{20}\right)}\right\}.$$

For any s such that $\mathsf{P}_{\overline{\mathcal{E}}|S}(s) > 0$, then $\alpha(s) \geq \alpha^*$, and conditioned on the event $\mathcal{G}$, we have $\overline{\alpha}(s) \geq \alpha^* - \frac{\gamma(1-\varepsilon)^2}{40}$, and thus

$$\mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(s)) = B(s) \,\middle|\, \mathcal{G}\right] = 1.$$

Moreover, for every other $s \in \mathcal{S}$, we have

$$\mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(s)) = B(s) \,\middle|\, \mathcal{G}\right] \geq \frac{1}{2} + \frac{\alpha(s) - \frac{\gamma(1-\varepsilon)^2}{40}}{2}.$$

Therefore,

$$\mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathcal{E}, \mathcal{G}\right] > \frac{1}{2} + \frac{\gamma(1-\varepsilon)^2}{40} - \frac{\gamma(1-\varepsilon)^2}{80} = \frac{1}{2} + \frac{\gamma(1-\varepsilon)^2}{80}.$$

We finally conclude that

$$\begin{aligned}\mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S)\right] &\geq \mathsf{P}\left[\mathcal{G}\right] \cdot \mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathcal{G}\right] \\ &\geq \mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathcal{G}\right] - \mathsf{P}\left[\overline{\mathcal{G}}\right] \\ &= \mathsf{P}\left[\mathcal{E}\right] \cdot \mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathcal{G}, \mathcal{E}\right] \\ &+ \mathsf{P}\left[\overline{\mathcal{E}}\right] \cdot \mathsf{P}\left[\mathbf{A}^{(B)}(\mathbf{S}(S)) = B(S) \,\middle|\, \mathcal{G}, \overline{\mathcal{E}}\right] - \mathsf{P}\left[\overline{\mathcal{G}}\right] \\ &> \varepsilon \cdot 1 + \frac{1}{2}(1-\varepsilon) + \frac{\gamma(1-\varepsilon)^3}{80} - \mathsf{P}\left[\overline{\mathcal{G}}\right] = \frac{1+\varepsilon}{2},\end{aligned}$$

as we wanted to show. The time complexity t_2 and the query complexity q_2 are clear by inspection.

D Proof of the Hardcore Lemma for Computational Indistinguishability (Theorem 2)

Let $\mathcal{S}$ and $\mathcal{T}$ be the respective ranges of S and T. We define $(\mathbf{F}, B)$ to be the cc-stateless system-bit pair with a uniform random bit B and where $\mathbf{F}$ behaves as $\mathbf{S}$ if $B = 0$ and as $\mathbf{T}$ if $B = 1$. In particular, $(\mathbf{F}, B)$ has state $(x, b) \in (\mathcal{S} \times \{0\}) \cup (\mathcal{T} \times \{1\})$, sampled by first letting $b \xleftarrow{\$} \{0, 1\}$, and then choosing $x \xleftarrow{\$} \mathsf{P}_S$ if $b = 0$ and $x \xleftarrow{\$} \mathsf{P}_T$ otherwise, and

$$(\mathbf{F}(x, b), B(x, b)) = \begin{cases} (\mathbf{S}(x), 0) & \text{if } b = 0, \\ (\mathbf{T}(x), 1) & \text{if } b = 1. \end{cases}$$

The canonical implementation $A_{(\mathbf{F},B)}$ of $(\mathbf{F}, B)$ is obtained by first choosing the random bit $b \xleftarrow{\$} \{0, 1\}$, and then using $A_{\mathbf{S}}$ or $A_{\mathbf{T}}$ to simulate the respective system: The state of $A_{(\mathbf{S},B)}$ is a pair (σ, b) consisting of the bit b and the current state σ of $A_{\mathbf{S}}$ or $A_{\mathbf{T}}$.

By a standard argument,

$$\Delta_{t,q}(\mathbf{S}, \mathbf{T}) = \mathrm{Guess}_{t,q}(B \,|\, \mathbf{F}) \leq \varepsilon,$$

and Theorem 1 thus implies that there exists a measure $\mathcal{M}$ on $(\mathcal{S} \times \{0\}) \cup (\mathcal{T} \times \{1\})$ such that $\mu(\mathcal{M}) \geq 1 - \varepsilon$, and

$$\mathrm{Guess}_{t',q'}(B' \,|\, \mathbf{F}(X')) \leq \gamma, \tag{5}$$

where $(X', B') \xleftarrow{\$} \mathcal{M}$, $t' := t/\varphi_{\mathsf{hc}}$, and $q' := q/\varphi_{\mathsf{hc}}$. Furthermore, a (ζ_1, ζ_2)-sampler O is associated with the measure $\mathcal{M}$ and $A_{(\mathbf{F},B)}$.

First, note that $\mathsf{P}_{B'}(0) \in \left[\frac{1-\gamma}{2}, \frac{1+\gamma}{2}\right]$, since otherwise there exists a fixed value $b' \in \{0, 1\}$ which can be output by an adversary (with constant time complexity and making no query) to achieve advantage higher than γ, contradicting (5) for any q' and reasonable t'. In the following, we assume without loss of generality that $\frac{1-\gamma}{2} \leq \mathsf{P}_{B'}(0) \leq \frac{1}{2}$, and $\frac{1}{2} \leq \mathsf{P}_{B'}(1) \leq \frac{1+\gamma}{2}$. (The other case is symmetric.)

We define the measures $\mathcal{M}_0 : \mathcal{S} \to [0, 1]$ and $\mathcal{M}_1 : \mathcal{R} \to [0, 1]$ such that

$$\mathcal{M}_0(s) := \mathcal{M}(s, 0) \quad \text{and} \quad \mathcal{M}_1(t) := \mathcal{M}(t, 1)$$

for all $s \in \mathcal{S}$ and all $r \in \mathcal{R}$. Then note that

$$\mu(\mathcal{M}_0) = \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathcal{M}_0(s) = 2\mu(\mathcal{M}) \cdot \mathsf{P}_{B'}(0), \mu(\mathcal{M}_1) = \sum_{t \in \mathcal{T}} \mathsf{P}_T(t) \cdot \mathcal{M}_1(t) = 2\mu(\mathcal{M}) \cdot \mathsf{P}_{B'}(1),$$

which in particular imply $(1 - \gamma) \cdot \mu(\mathcal{M}) \leq \mu(\mathcal{M}_0) \leq \mu(\mathcal{M})$ and $\mu(\mathcal{M}) \leq \mu(\mathcal{M}_1) \leq (1 + \gamma)\mu(\mathcal{M})$. Consequently, we set $\mathcal{M}_{\mathbf{T}} := \mathcal{M}_1$, whereas, if $\mu(\mathcal{M}_0) < 1 - \varepsilon$, we define

$$\mathcal{M}_{\mathbf{S}}(s) := \frac{\varepsilon}{1 - \mu(\mathcal{M}_0)} \cdot \mathcal{M}_0(s) + \frac{1 - \varepsilon - \mu(\mathcal{M}_0)}{1 - \mu(\mathcal{M}_0)}.$$

Observe that $\mathcal{M}_{\mathbf{S}}(s) \in [0, 1]$, $\mathcal{M}_{\mathbf{S}}(s) \geq \mathcal{M}_0(s)$, and

$$\mu(\mathcal{M}_{\mathbf{S}}) = \sum_{s \in \mathcal{S}} \mathsf{P}_S(s) \cdot \mathcal{M}_{\mathbf{S}}(s) = \frac{\varepsilon}{1 - \mu(\mathcal{M}_0)} \cdot \mu(\mathcal{M}_0) + \frac{1 - \varepsilon - \mu(\mathcal{M}_0)}{1 - \mu(\mathcal{M}_0)} = 1 - \varepsilon.$$

Finally, consider the distribution (X'', B'') which chooses B'' uniformly at random, and then X'' according to $\mathcal{M}_{\mathbf{S}}$ or $\mathcal{M}_{\mathbf{T}}$ depending on whether $B'' = 0$ or $B'' = 1$. With $S' \xleftarrow{\$} \mathcal{M}_{\mathbf{S}}$ and $T' \xleftarrow{\$} \mathcal{M}_{\mathbf{T}}$,

$$\begin{aligned}\Delta_{t',q'}(\mathbf{S}(S'), \mathbf{T}(T')) &= \mathrm{Guess}_{t',q'}(B'' \,|\, \mathbf{F}(X'')) \\ &\leq \mathrm{Guess}_{t',q'}(B' \,|\, \mathbf{F}(X')) + 2 \cdot d((B', X'), (B'', X'')) \\ &\leq \gamma + 2 \cdot d((B', X'), (B'', X'')),\end{aligned}$$

since in general, for any (X_1, B_1) and (X_2, B_2), we have

$$\mathrm{Guess}^{\mathbf{A}}(B_1 \,|\, \mathbf{F}(X_1)) - \mathrm{Guess}^{\mathbf{A}}(B_2 \,|\, \mathbf{F}(X_2)) \leq 2 \cdot d((X_1, B_1), (X_2, B_2))$$

for all adversaries $\mathbf{A}$.[14] The following claim yields the desired upper bound.

Claim. $d((X', B'), (X'', B'')) \leq \frac{\gamma}{2}$.

Proof. On the one hand, we first note that for all $t \in \mathcal{T}$ we have

$$\mathsf{P}_{X''B''}(t, 1) = \frac{1}{2\mu(\mathcal{M}_1)} \cdot \mathsf{P}_T(t) \cdot \mathcal{M}(t, 1) \leq \frac{1}{2\mu(\mathcal{M})} \cdot \mathsf{P}_T(t) \cdot \mathcal{M}(t, 1) = \mathsf{P}_{X'B'}(t, 1),$$

where we have used the fact that $\mu(\mathcal{M}_1) \geq \mu(\mathcal{M})$. On the other hand, for all $s \in \mathcal{S}$ we have (using $\mu(\mathcal{M}_{\mathbf{S}}) = (1 - \varepsilon) \leq \mu(\mathcal{M})$ and $\mathcal{M}_0(s) \leq \mathcal{M}_{\mathbf{S}}(s)$)

$$\mathsf{P}_{X''B''}(s, 0) = \frac{1}{2\mu(\mathcal{M}_{\mathbf{S}})} \cdot \mathsf{P}_S(s) \cdot \mathcal{M}_{\mathbf{S}}(s) \geq \frac{1}{2\mu(\mathcal{M})} \cdot \mathsf{P}_S(s) \cdot \mathcal{M}(s, 0) = \mathsf{P}_{X'B'}(s, 0).$$

Therefore,

$$\begin{aligned}d((X', B'), (X'', B'')) &= \sum_{t \in \mathcal{T}} \mathsf{P}_{X'B'}(t, 1) - \mathsf{P}_{X''B''}(t, 1) \\ &= \frac{1}{2} \sum_{r \in \mathcal{R}} \mathsf{P}_T(t) \cdot \mathcal{M}(t, 1) \cdot \left(\frac{1}{\mu(\mathcal{M})} - \frac{1}{\mu(\mathcal{M}_1)} \right) = \frac{1}{2} \left(\frac{\mu(\mathcal{M}_1)}{\mu(\mathcal{M})} - 1 \right) \leq \frac{\gamma}{2}.\square\end{aligned}$$

Finally, the state sampler O for $\mathcal{M}$ and $A_{(\mathbf{F},B)}$ with length $\ell = s_{A_{(\mathbf{F},B)}}(\psi \cdot q')$ outputs a triple (b, σ, z), where σ is a state for $A_{\mathbf{S}}$ if $b = 0$ and a state of $A_{\mathbf{T}}$ when $b = 1$. In particular, for $(\Sigma, B, Z) \xleftarrow{\$} \mathsf{O}$,

$$(A_B[\Sigma], B, Z) \equiv (\mathbf{F}(X, B), B, Z(X, B)),$$

where $A_0 := A_{\mathbf{S}}$, $A_1 := A_{\mathbf{T}}$, and $Z(X, B)$ is at most ζ_1 off $\mathcal{M}(X, B)$, except with probability ζ_2, for *all* values x, b taken by X, B.

Consequently, the state sampler $\mathsf{O}_{\mathbf{T}}$ for $\mathcal{M}_{\mathbf{T}}$ outputs (σ, z) sampled according to the output distribution O conditioned on $B = 1$. The sampler $\mathsf{O}_{\mathbf{S}}$ outputs (σ, z') by taking (σ, z) sampled from O conditioned on $B = 0$, and then sets $z' := z$ if $\mu(\mathcal{M}_0) \geq 1 - \varepsilon$, and otherwise sets $z' := \frac{\varepsilon}{1-\mu(\mathcal{M}_0)} \cdot z + \frac{1-\varepsilon-\mu(\mathcal{M}_0)}{1-\mu(\mathcal{M}_0)}$. Clearly, it remains a (ζ_1, ζ_2)-sampler, due to $\frac{\varepsilon}{1-\mu(\mathcal{M}_0)} \leq 1$.

Note that Theorem 1 only guarantees that the length of the states output by both samplers is bounded by $\max\{s_{A_{\mathbf{S}}}(\psi_{\mathsf{hc}} \cdot q'), s_{A_{\mathbf{T}}}(\psi_{\mathsf{hc}} \cdot q')\}$. The fact that the individual bounds hold for each of the sampler can be inferred from a careful analysis of the proof of Theorem 1. (However, note that this more precise statement is in general not necessary in the following, but will produce slightly nicer statements.)

Also, if both $\mathbf{S}$ and $\mathbf{T}$ are cc-stateless random functions, then $(\mathbf{F}, B)$ has deterministic behavior for any value of its initial state, and thus O is error-less, which implies that $\mathsf{O}_{\mathbf{S}}$ and $\mathsf{O}_{\mathbf{T}}$ are also error-less.

[14] To see this, note that the distinguisher which given (X, B) simulates the interaction of $\mathbf{A}$ with $\mathbf{F}(X)$ and outputs 1 if and only if its output equals B achieves advantage

$$\frac{1}{2} \left[\mathrm{Guess}^{\mathbf{A}}(B_1 \,|\, \mathbf{F}(X_1)) - \mathrm{Guess}^{\mathbf{A}}(B_2 \,|\, \mathbf{F}(X_2)) \right] \leq d((X_1, B_1), (X_2, B_2)).$$

E Cascade of Large-Min-Entropy Permutations

E.1 Indistinguishability Proofs: A Primer and a New Lemma

Before we turn to the proof of Theorem 3, this section provides a self-contained introduction to selected tools from the random systems framework [27, 30] which are needed for information-theoretic indistinguishability proofs.

Monotone Event Sequences and Indistinguishability. Given a system $\mathbf{S}$, a *monotone event sequence (MES)* $\mathcal{A} = A_0, A_1, \ldots$ on $\mathbf{S}$ is a sequence of events[15] where A_i is defined after the i-th query has been answered by $\mathbf{S}$ and such that A_0 is defined 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 $j \geq i$. When A_i holds, and $\overline{A_{i+1}}$ occurs, then we say that $\mathcal{A}$ *fails*.

Definition 2. *Let $\mathbf{S}$ and $\mathbf{T}$ be systems, and let $\mathcal{A}$ be a MES on $\mathbf{S}$. We write $\mathbf{S}|\mathcal{A} \equiv \mathbf{T}$ if for all $i \geq 1$ and for all $y_i, x^i = [x_1, \ldots, x_i]$, and $y^{i-1} = [y_1, \ldots, y_{i-1}]$,*

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

Furthermore, if $\mathcal{B} = B_0, B_1, \ldots$ is a MES on $\mathbf{T}$, we write $\mathbf{S}^{\mathcal{A}} \equiv \mathbf{T}^{\mathcal{B}}$ if for all $i \geq 1$ and for all $y_i, x^i = [x_1, \ldots, x_i]$, and $y^{i-1} = [y_1, \ldots, y_{i-1}]$

$$\mathsf{p}^{\mathbf{S}}_{A_iY_i|X^iY^{i-1}A_{i-1}}(y_i, x^i, y^{i-1}) = \mathsf{p}^{\mathbf{T}}_{B_iY_i|X^iY^{i-1}B_{i-1}}(y_i, x^i, y^{i-1}).$$

The shorthand $\nu_q(\mathbf{S}, \overline{A_q})$ stands for the optimal probability that some q query distinguisher makes $\mathcal{A}$ fail while interacting with $\mathbf{S}$, i.e., that the event $\overline{A_q}$ holds. The following lemma [27, 30] shows the connection between the probability of making a MES fail and the distinguishing advantage.

Lemma 7. *Let $\mathbf{S}$ and $\mathbf{T}$ be systems, and let $\mathcal{A} = A_0, A_1, \ldots$ and $\mathcal{B} = B_0, B_1, \ldots$ be MES on $\mathbf{S}$ and $\mathbf{T}$, respectively. If $\mathbf{S}|\mathcal{A} \equiv \mathbf{T}$ or $\mathbf{S}^{\mathcal{A}} \equiv \mathbf{T}^{\mathcal{B}}$, then*

$$\Delta_q(\mathbf{S}, \mathbf{T}) \leq \nu_q(\mathbf{S}, \overline{A_q}).$$

We also employ a new lemma (a similar statement was given in [33]) that simplifies the problem of upper bounding $\nu_q(\mathbf{S}, \overline{A_q})$ and considers the setting where a MES $\mathcal{A} = A_0, A_1, \ldots$ is defined on a system $\mathbf{S} \equiv \mathbf{S}(S)$ (i.e., where S is some internal variable of $\mathbf{S}$) such that the behavior of $\mathbf{S}$ is independent of S as long as $\mathcal{A}$ does not fail. We show that if $\mathcal{A}$ only depends on the input-output behavior and the value of S, we can equivalently consider the maximal probability, over all compatible transcripts (X^q, Y^q), that a state $S \xleftarrow{\$} \mathsf{P}_S$ provokes $\overline{A_q}$. This will be very handy in the proof of the next section.

Lemma 8. *Let $\mathbf{S} \equiv \mathbf{S}(S)$ be a system depending on a state S, and let $\mathbf{T}$ be an additional system. Let $\mathcal{A}$ be a MES on $\mathbf{S}$ such that for all $i \geq 1$ there exists a sequence of sets $\mathcal{E}_0, \mathcal{E}_1, \mathcal{E}_2, \ldots$ with the property that $\overline{A_i}$ holds if and only if $(X^i, Y^i, S) \in \mathcal{E}_i$. If for all $i \geq 1$ and all $(x^i, y^i, s) \notin \mathcal{E}_i$*

$$\mathsf{p}^{\mathbf{S}}_{Y_iA_i|X^iY^{i-1}A_{i-1}S}(y_i, x^i, y^{i-1}, s) = \mathsf{p}^{\mathbf{S}}_{Y_i|X^iY^{i-1}A_{i-1}S}(y_i, x^i, y^{i-1}, s) = \mathsf{p}^{\mathbf{T}}_{Y_i|X^iY^{i-1}}(y_i, x^i, y^{i-1}), \tag{6}$$

then

$$\Delta_q(\mathbf{S}, \mathbf{T}) \leq \max_{x^q, y^q} \mathsf{P}\left[(x^q, y^q, S) \in \mathcal{E}_q\right],$$

where the maximum is taken over all sequence of inputs $x^q = [x_1, \ldots, x_q]$ and outputs $y^q = [y_1, \ldots, y_q]$ which are compatible with an interaction with $\mathbf{S}$, and the probability is over the choice of S.

[15] We do not follow our notational conventions by writing A_i instead of $\mathcal{A}_i$. This is in order to be consistent with existing literature on random systems.

Proof. We define the system $\mathbf{H} \equiv \mathbf{H}(S)$ which initially chooses the (secret) state S according to the distribution P_S, and then ignores it and behaves as $\mathbf{T}$. Obviously, $\mathbf{T} \equiv \mathbf{H}$. We define the MES $\mathcal{B}$ on $\mathbf{H}$ which fails after i queries if $(X^i, Y^i, S) \in \mathcal{E}_i$. Then, for all $i \geq 1$ and all $(x^i, y^i, s) \notin \mathcal{E}_i$

$$\mathsf{p}^{\mathbf{S}}_{A_i Y_i | X^i Y^{i-1} A_{i-1} S}(y_i, x^i, y^{i-1}, s) = \mathsf{p}^{\mathbf{T}}_{Y_i | X^i Y^{i-1}}(y_i, x^i, y^{i-1}) = \mathsf{p}^{\mathbf{H}}_{Y_i B_i | X^i Y^{i-1} B_{i-1} S}(y_i, x^i, y^{i-1}, s),$$

which in particular implies $\mathbf{S}^{\mathcal{A}} \equiv \mathbf{H}^{\mathcal{B}}$, and therefore by Lemma 7, this implies that

$$\Delta_q(\mathbf{S}, \mathbf{T}) = \Delta_q(\mathbf{H}, \mathbf{S}) \leq \nu_q(\mathbf{H}, \overline{B_q}).$$

Let $\mathbf{D}$ be a q query distinguisher interacting with $\mathbf{S}$. Then we have (with $\mathsf{P}^{\mathbf{D}\,\mathbf{H}}$ denoting probabilities in the random experiment where $\mathbf{D}$ interacts with $\mathbf{H}$)

$$\begin{aligned}
\mathsf{P}^{\mathbf{D}\,\mathbf{H}}\left[\overline{B_q}\right] = \sum_{x^q, y^q, s} \mathsf{P}^{\mathbf{D}\mathbf{H}}_{X^q Y^q S \overline{B_q}}(x^q, y^q, s) &= \sum_{x^q, y^q, s} \mathsf{P}_S(s) \cdot \mathsf{P}^{\mathbf{D}\,\mathbf{H}}_{X^q Y^q | S}(x^q, y^q, s) \cdot \mathsf{p}^{\mathbf{H}}_{\overline{B_q} | X^q Y^q S}(x^q, y^q, s) \\
&= \sum_{x^q, y^q, s} \mathsf{P}_S(s) \cdot \mathsf{P}^{\mathbf{D}\,\mathbf{H}}_{X^q Y^q}(x^q, y^q) \cdot \mathsf{p}^{\mathbf{H}}_{\overline{B_q} | X^q Y^q S}(x^q, y^q, s) \\
&= \sum_{x^q, y^q} \mathsf{P}^{\mathbf{D}\,\mathbf{H}}_{X^q Y^q}(x^q, y^q) \cdot \underbrace{\sum_s \mathsf{P}_S(s) \cdot \mathsf{p}^{\mathbf{H}}_{\overline{B_q} | X^q Y^q S}(x^q, y^q, s)}_{= \mathsf{P}_S[(x^q, y^q, S) \in \mathcal{E}_q]},
\end{aligned}$$

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

E.2 Proof of Theorem 3

In this section, we present a proof of the following theorem.

Theorem 3 (Cascade of Large-Min-Entropy Permutations). *For all $q, \Lambda \geq 1$,*

$$\Delta_q(\mathbf{Q}_1 \rhd \mathbf{Q}_2, \mathbf{P}) \leq \Delta_q(\langle \mathbf{Q}_1 \rhd \mathbf{Q}_2 \rangle, \langle \mathbf{P} \rangle) \leq \tfrac{4q\Lambda}{N} + \tfrac{2\Lambda(q+\Lambda)}{(1-\varepsilon)N} + 2\left(\tfrac{q \log\left((1-\varepsilon)^{-1}\right)}{\Lambda}\right)^{\frac{1}{2}}.$$

In the following, recall that the *Shannon entropy* of X is $\mathsf{H}(X) := -\sum_{x \in \mathcal{X}} \mathsf{P}_X(x) \cdot \log \mathsf{P}_X(x)$ (with $0 \cdot \log 0 = 0$). Also, throughout this section, we let $N := |\mathcal{X}|$ and for an integer N, we define $N^{(i)} := N \cdot (N-1) \cdots (N-i+1)$. In particular, $N^{(N)} = N!$. For every distinguisher $\overline{\mathbf{D}}$ and for a cc-stateless two-sided random permutation $\langle \mathbf{Q} \rangle$ with domain $\mathcal{X}$, we define the two-sided random permutation $\langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle$ with domain $\mathcal{X}$ which is initialized by letting $\overline{\mathbf{D}}$ interact with $\mathbf{Q}$: If $(x_1, y_1), \ldots, (x_\Lambda, y_\Lambda)$ are the resulting *input-output pairs*, i.e., for all i either a forward query $(x_i, +)$ returned y_i or a backward query $(y_i, -)$ returned x_i, then $\langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle$ answers backward and forward queries according to a randomly chosen permutation Q constrained to $Q(x_i) = y_i$ for all $i \in \{1, \ldots, \Lambda\}$. The following lemma states that if a cc-stateless two-sided permutation $\langle \mathbf{Q} \rangle$ behaves according to a permutation table whose distribution has large Shannon entropy, than there always exists a good *deterministic* $\overline{\mathbf{D}}$ such that $\langle \mathbf{Q} \rangle$ and $\langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle$ are indistinguishable.

Lemma 9. *For all $q, \Lambda > 1$ and for all cc-stateless two sided random permutations $\langle \mathbf{Q} \rangle$ with domain $\mathcal{X}$ behaving according to a function table Q such that $\mathsf{H}(Q) \geq \log(N!) - \delta$, there exists a deterministic distinguisher $\overline{\mathbf{D}}$ making at most Λ queries such that*

$$\Delta_q(\langle \mathbf{Q} \rangle, \langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle) \leq \left(\frac{q\delta}{\Lambda}\right)^{\frac{1}{2}}.$$

The proof of Lemma 9 follows the approach of a similar result by Unruh [43] for the simpler case of random functions, and is postponed to Section E.3. Lemma 9 implies that in our settings there exists $\overline{\mathbf{D}}_1$ and $\overline{\mathbf{D}}_2$ such that for $i = 1, 2$,

$$\Delta_q(\langle \mathbf{Q}_i \rangle, \langle \mathbf{Q}_i \rangle_{\overline{\mathbf{D}}_i}) \leq \left(\frac{q\delta}{\Lambda}\right)^{\frac{1}{2}}.$$

Let $\langle \mathbf{Q}'_i \rangle := \langle (\mathbf{Q}_i)_{\overline{\mathbf{D}}_i} \rangle$ for $i = 1, 2$ be the two corresponding systems.

The Hybrid Systems $\langle \mathbf{H} \rangle$. We introduce a hybrid two-sided (stateful) random permutation $\langle \mathbf{H} \rangle$, which simulates the cascade $\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle$, with some differences. The two-sided permutation $\langle \mathbf{H} \rangle$ is initialized by running $\overline{\mathbf{D}}_1$ and $\overline{\mathbf{D}}_2$ on $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively: Let

$$\mathcal{T}^{(1)} := \left\{(x_1^{(1)}, y_1^{(1)}), \ldots, (x_\Lambda^{(1)}, y_\Lambda^{(1)})\right\} \quad \text{and} \quad \mathcal{T}^{(2)} := \left\{(y_1^{(2)}, z_1^{(2)}), \ldots, (y_\Lambda^{(2)}, z_\Lambda^{(2)})\right\}$$

be the resulting input-output pairs. We also use the shorthands

$$\mathcal{X}^{(1)} := \{x_1^{(1)}, \ldots, x_\Lambda^{(1)}\},\ \mathcal{Y}^{(1)} := \{y_1^{(1)}, \ldots, y_\Lambda^{(1)}\},\ \mathcal{Y}^{(2)} := \{y_1^{(2)}, \ldots, y_\Lambda^{(2)}\},\ \mathcal{Z}^{(2)} := \{z_1^{(2)}, \ldots, z_\Lambda^{(2)}\}.$$

Additionally $\langle \mathbf{H} \rangle$ maintains a set of triples $\mathcal{Q}$ (which is initially empty). A triple $(x, y, z) \in \mathcal{Q}$ means that the first permutation in the cascade simulated by $\langle \mathbf{H} \rangle$ maps x to y, whereas the second one maps y to z. For notational convenience, we let $\mathcal{Q}_1$ be the set of elements x which appear as first component of a triple in $\mathcal{Q}$. Similarly, we define $\mathcal{Q}_2$ and $\mathcal{Q}_3$ as the projections on the second and third components, respectively.

Also, at every point in time in the interaction, let $\widetilde{\mathcal{T}}^{(1)} \subseteq \mathcal{T}^{(1)}$ be the subset of pairs which do not share a component with $\mathcal{Q}_1$ and $\mathcal{Q}_2$, and let $\widetilde{\mathcal{T}}^{(2)}$ be defined analogously. Also, let $\widetilde{\mathcal{X}}^{(1)}$, $\widetilde{\mathcal{Y}}^{(1)}$, $\widetilde{\mathcal{Y}}^{(2)}$ and $\widetilde{\mathcal{Z}}^{(2)}$ be the associated projections.

On input $(x, +)$ with $x \notin \mathcal{Q}_1$, a new triple (x, y, z) with $y \notin \mathcal{Q}_2$ and $z \notin \mathcal{Q}_3$ is added as follows:

(A) If there exists $(x, y') \in \widetilde{\mathcal{T}}^{(1)}$, then we set $y := y'$. Furthermore, if there exists $(y, z') \in \widetilde{\mathcal{T}}^{(2)}$, then we set $z := z'$. Otherwise we choose $z \xleftarrow{\$} \mathcal{X} \setminus \mathcal{Q}_3$.

(B) Otherwise, we set $y \xleftarrow{\$} \mathcal{X} \setminus \mathcal{Q}_2$ (i.e., the set of all values which have not been used yet). Furthermore, if some $(y, z') \in \widetilde{\mathcal{T}}^{(2)}$, then $z := z'$. Otherwise, we set $z \xleftarrow{\$} \mathcal{X} \setminus (\mathcal{Q}_3 \cup \widetilde{\mathcal{Z}}^{(2)})$.

Backward queries $(y, -)$ are answered symmetrically.

Bounding $\Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{H} \rangle)$. The only difference between $\langle \mathbf{H} \rangle$ and $\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle$ occurs in (A) when sampling $z \xleftarrow{\$} \mathcal{X} \setminus \mathcal{Q}_3$ rather than $z \xleftarrow{\$} \mathcal{X} \setminus (\mathcal{Q}_3 \cup \mathcal{Z}^{(2)})$ and in (B) when $y \xleftarrow{\$} \mathcal{X} \setminus \mathcal{Q}_2$ rather than $y \xleftarrow{\$} \mathcal{X} \setminus (\mathcal{Q}_2 \cup \mathcal{Y}^{(1)})$. Also, the symmetrical statement holds for backward queries.

Consequently, we define a MES $\mathcal{A} = A_1, A_2, \ldots$ on $\langle \mathbf{H} \rangle$ which fails as soon as an element $z \in \mathcal{Z}^{(2)}$, or $y \in \mathcal{Y}^{(1)}$ (in a forward query), or $x \in \mathcal{X}^{(1)}$, or $y \in \mathcal{Y}^{(2)}$ (in a backward query) is sampled. Clearly $\langle \mathbf{H} \rangle \,|\mathcal{A} \equiv \langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle$, and by Lemma 7,

$$\Delta_q(\langle \mathbf{H} \rangle, \langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle) \leq \nu_q(\langle \mathbf{H} \rangle, \overline{A_q}).$$

To upper bound $\nu_q(\langle \mathbf{H} \rangle, \overline{A_q})$, we think of the four assignments of interest to be equivalently executed as follows: Initially, three independent lists $x'_1, \ldots, x'_q$, $y'_1, \ldots, y'_q$, and $z'_1, \ldots, z'_q$, consisting each of uniform independent random *distinct* elements of $\mathcal{X}$, are sampled. Whenever we have to assign a value x', y', or z' in one of the four cases of interest, we pick the first element from the corresponding list, and delete it from the list. Also, whenever $\langle \mathbf{H} \rangle$ adds a triple (x', y', z') to $\mathcal{Q}$, the values x', y', and z' are removed from the corresponding lists, if they appear in them, after the query is answered.

The probability that $\overline{A_q}$ occurs is upper bounded by the probability that one of $\mathcal{X}^{(1)} \cap \{x'_1, \ldots, x'_q\}$, $\mathcal{Y}^{(1)} \cap \{y'_1, \ldots, y'_q\}$, $\mathcal{Y}^{(2)} \cap \{y'_1, \ldots, y'_q\}$, or $\mathcal{Z}^{(2)} \cap \{z'_1, \ldots, z'_q\}$ is non-empty. As every element of the three lists is, individually, uniformly distributed, the union bound yields

$$\Delta_q(\langle \mathbf{H} \rangle, \langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle) \leq \nu_q(\langle \mathbf{H} \rangle, \overline{A_q}) \leq \frac{4q\Lambda}{N}.$$

Bounding $\Delta_q(\langle \mathbf{H} \rangle, \langle \mathbf{P} \rangle)$. In the following, we consider the sets $\mathcal{X}^{(1)}_{\text{in}} \subseteq \mathcal{X}^{(1)}$ and $\mathcal{Y}^{(1)}_{\text{in}} \subseteq \mathcal{Y}^{(1)}$ of inputs of forward queries and backward queries of $\overline{\mathbf{D}}_1$, respectively. Analogously, we define $\mathcal{Y}^{(2)}_{\text{in}}$ and $\mathcal{Z}^{(2)}_{\text{in}}$. Furthermore, we let

$$\mathcal{X}^{(1)}_{\text{out}} := \mathcal{X}^{(1)} \setminus \mathcal{X}^{(1)}_{\text{in}}, \mathcal{Y}^{(1)}_{\text{out}} := \mathcal{Y}^{(1)} \setminus \mathcal{Y}^{(1)}_{\text{in}}, \mathcal{Y}^{(2)}_{\text{out}} := \mathcal{Y}^{(2)} \setminus \mathcal{Y}^{(2)}_{\text{in}}, \mathcal{Z}^{(2)}_{\text{out}} := \mathcal{Z}^{(2)} \setminus \mathcal{Z}^{(2)}_{\text{in}}.$$

Then, we define the MES $\mathcal{B} = B_1, B_2, \ldots$ on $\langle \mathbf{H} \rangle$ such that B_i is false if at initialization $\mathcal{Y}^{(1)}_{\text{out}} \cap \mathcal{Y}^{(2)} \neq \emptyset$, or $\mathcal{Y}^{(2)}_{\text{out}} \cap \mathcal{Y}^{(1)} \neq \emptyset$ occur, or one the following is true:

- A forward query $X_j = (x, +)$ (with $j \in \{1, \ldots, i\}$) with $x \in \mathcal{X}^{(1)}_{\text{out}}$ is issued;
- A backward query $X_j = (z, -)$ (with $j \in \{1, \ldots, i\}$) with $z \in \mathcal{Z}^{(2)}_{\text{out}}$ is issued.

Assume that the i-th query is a *forward* query $(x, +)$. (For a backward query, the argument is symmetric.) Then, given that B_i holds, we consider two cases:

- If (A) occurs, then the fact that B_i holds yields that $x = x^{(1)}_j \in \mathcal{X}^{(1)}_{\text{in}}$ (for some $j \in \{1, \ldots, \Lambda\}$), and that $y^{(1)}_j \notin \mathcal{Y}^{(2)}$, which in turn implies $y^{(1)}_j \notin \widetilde{\mathcal{Y}}^{(2)}$: Thus a random value from $\mathcal{X} \setminus \mathcal{Q}_3$ is returned.
- Otherwise, if (B) occurs, consider a particular value $z \in \mathcal{X} \setminus \mathcal{Q}_3$. If $z = z^{(2)}_j \in \widetilde{\mathcal{Z}}^{(2)}$ (for some $j \in \{1, \ldots, \Lambda\}$), the probability that this value is returned is $\frac{1}{|\mathcal{X} \setminus \mathcal{Q}_2|} = \frac{1}{|\mathcal{X} \setminus \mathcal{Q}_3|}$, i.e., the probability that $y^{(2)}_j \in \widetilde{\mathcal{Y}}^{(2)}$ is chosen.
  On the other hand, if $z \notin \widetilde{\mathcal{Z}}^{(2)}$, the probability that z is returned equals

$$\frac{\left|\mathcal{X} \setminus (\mathcal{Q}_2 \cup \widetilde{\mathcal{Y}}^{(2)})\right|}{|\mathcal{X} \setminus \mathcal{Q}_2|} \cdot \frac{1}{\left|\mathcal{X} \setminus \left(\mathcal{Q}_3 \cup \widetilde{\mathcal{Z}}^{(2)}\right)\right|} = \frac{1}{|\mathcal{X} \setminus \mathcal{Q}_2|} = \frac{1}{|\mathcal{X} \setminus \mathcal{Q}_3|},$$

  since $|\mathcal{Q}_2| = |\mathcal{Q}_3|$, $\widetilde{\mathcal{Y}}^{(2)} \cap \mathcal{Q}_2 = \emptyset$, $\widetilde{\mathcal{Z}}^{(2)} \cap \mathcal{Q}_3 = \emptyset$, and $\left|\widetilde{\mathcal{Y}}^{(2)}\right| = \left|\widetilde{\mathcal{Z}}^{(2)}\right|$.

Also, with the state S consisting of a description of $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$, note that B_i only depends (deterministically) on S and X^i, Y^{i-1}, and the conditions for applying Lemma 8 are fulfilled with $\mathbf{T}$ being a URP, since

$$\mathsf{p}^{\langle \mathbf{H} \rangle}_{Y_i B_i | X^i Y^{i-1} B_{i-1} S} = \mathsf{p}^{\langle \mathbf{H} \rangle}_{B_i | X^i Y^{i-1} B_{i-1} S} \cdot \mathsf{p}^{\langle \mathbf{H} \rangle}_{Y_i | X^i Y^{i-1} B_i S} = \mathsf{p}^{\langle \mathbf{H} \rangle}_{B_i | X^i Y^{i-1} B_{i-1} S} \cdot \mathsf{p}^{\langle \mathbf{P} \rangle}_{Y_i | X^i Y^{i-1}}.$$

Given input-output pairs $(x_1, z_1), \ldots, (x_q, z_q)$, we sample

$$(X^{(1)}_1, Y^{(1)}_1), \ldots, (X^{(1)}_\Lambda, Y^{(1)}_\Lambda) \quad \text{and} \quad (Y^{(2)}_1, Z^{(2)}_1), \ldots, (Y^{(2)}_\Lambda, Z^{(2)}_\Lambda)$$

by letting $\overline{\mathbf{D}}_1$ and $\overline{\mathbf{D}}_2$ interact with $\mathbf{Q}_1$ and $\mathbf{Q}_2$ independently, and upper bound the advantage by $\mathsf{P}[\mathcal{E}_1 \vee \mathcal{E}_2] \leq \mathsf{P}[\mathcal{E}_1] + \mathsf{P}[\mathcal{E}_2]$, where $\mathcal{E}_1$ is the event that *any* of the outputs of the queries of $\overline{\mathbf{D}}_1$ is in $\mathcal{Y}^{(2)} \cup \mathcal{X}^q$ and $\mathcal{E}_2$ is the event that any of the outputs of the queries of $\overline{\mathbf{D}}_2$ is in $\mathcal{Z}^q \cup \mathcal{Y}^{(1)}$. Note that the sequence of outputs of queries of $\overline{\mathbf{D}}_1$ and $\overline{\mathbf{D}}_2$ have both min-entropy at least $\log\left(N^{(\Lambda)}\right) - \log\left((1-\varepsilon)^{-1}\right)$, i.e., each

such sequence occurs with probability at most $\frac{1}{(1-\varepsilon)N^{(\Lambda)}}$. On the other hand, since $\left|\mathcal{X}^i \cup \mathcal{Y}^{(2)}\right| \leq q+\Lambda$ and $\left|\mathcal{Z}^i \cup \mathcal{Y}^{(1)}\right| \leq q+\Lambda$, there are (by a union bound) at most $\Lambda(q+\Lambda)\cdot(N-1)^{(\Lambda-1)}$ sequences provoking $\mathcal{E}_1$ and $\mathcal{E}_2$. Thus, we conclude that

$$\Delta_q(\langle \mathbf{H} \rangle, \langle \mathbf{P} \rangle) \leq 2 \cdot \frac{\Lambda(q+\Lambda)}{(1-\varepsilon)N}.$$

Wrapping up. We can now collect the previous bounds:

$$\begin{aligned}\Delta_q(\langle \mathbf{Q}_1 \rhd \mathbf{Q}_2 \rangle, \langle \mathbf{P} \rangle) &\leq \Delta_q(\langle \mathbf{Q}_1 \rhd \mathbf{Q}_2 \rangle, \langle \mathbf{Q}_1 \rhd \mathbf{Q}'_2 \rangle) + \\ &\quad + \Delta_q(\langle \mathbf{Q}_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle) + \Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{P} \rangle) \\ &\leq \Delta_q(\langle \mathbf{Q}_1 \rangle, \langle \mathbf{Q}'_1 \rangle) + \Delta_q(\langle \mathbf{Q}_2 \rangle, \langle \mathbf{Q}'_2 \rangle) + \Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{P} \rangle) \\ &\leq 2 \cdot \left(\frac{q \log\left((1-\varepsilon)^{-1}\right)}{\Lambda}\right)^{\frac{1}{2}} + \Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{P} \rangle),\end{aligned}$$

where we used the fact that removing the first (or the second) permutation only makes distinguishing easier. Furthermore,

$$\Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{P} \rangle) \leq \Delta_q(\langle \mathbf{Q}'_1 \rhd \mathbf{Q}'_2 \rangle, \langle \mathbf{H} \rangle) + \Delta_q(\langle \mathbf{H} \rangle, \langle \mathbf{P} \rangle) \leq \frac{4q\Lambda}{N} + \frac{2\Lambda(q+\Lambda)}{(1-\varepsilon)N},$$

which concludes the proof.

E.3 Proof of Lemma 9

Recall that we want to prove the following lemma.

Lemma 9. *For all $q, \Lambda > 1$ and for all cc-stateless two sided random permutations $\langle \mathbf{Q} \rangle$ with domain $\mathcal{X}$ behaving according to a function table Q such that $\mathsf{H}(Q) \geq \log(N!) - \delta$, there exists a deterministic distinguisher $\overline{\mathbf{D}}$ making at most Λ queries such that*

$$\Delta_q(\langle \mathbf{Q} \rangle, \langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle) \leq \left(\frac{q\delta}{\Lambda}\right)^{\frac{1}{2}}.$$

For a *deterministic* distinguisher $\mathbf{D}$ making q *non-redundant*[16] two-sided queries $X_1, \ldots, X_q$ such that $X_i \in \{(x,+),(y,-) : x, y \in \mathcal{X}\}$ to a two-sided cc-stateless random permutation $\langle \mathbf{Q} \rangle$ with domain $\mathcal{X}$, it is convenient to define the quantity

$$\mathsf{H}^{\mathbf{D}}(\mathbf{Q}) := \mathsf{H}(Y_1, Y_2, \ldots, Y_q),$$

where $Y_1, Y_2, \ldots, Y_q \in \mathcal{X}$ are the answers of these queries made by $\mathbf{D}$ (in the order they are issued). As all queries are non-redundant, $q \leq N$. Furthermore, note that the quantity $\mathsf{H}^{\mathbf{D}}(\mathbf{Q})$ makes only sense if $\mathbf{D}$ is deterministic, as otherwise additional randomness can be injected by $\mathbf{D}$ itself. Additionally, we define the function $d : \{1, \ldots, N\} \to \mathbb{R}$ such that

$$d(q) := \max_{\mathbf{D}} \left[\log N^{(q)} - \mathsf{H}^{\mathbf{D}}(\mathbf{Q})\right],$$

where the maximum is taken over all $\mathbf{D}$'s as above making exactly q non-redundant queries, i.e., $d(q)$ measure intuitively the best deviation from the maximal possible entropy which can be achieved by such a q-query $\mathbf{D}$, which is $\log N^{(q)}$. We will need the following properties for d.

[16] This means that the distinguisher $\mathbf{D}$ does not query *any* value for which it knows the output, for instance, if a query $(x,+)$ returns y, it will never issue a query $(y,-)$ at a later point in time.

Claim. The function d satisfies the following two properties:

(i) $d(N) \le \delta$.
(ii) $d(i) \le d(i+1)$ for all $i \in \{0, \dots, N-1\}$.

Proof. For (i), fix an arbitrary N-query deterministic $\mathbf{D}$ interacting with $\mathbf{Q}$. Note that with knowledge of $\mathbf{D}$, given answers $Y_1, \dots, Y_N$ we can uniquely reconstruct the function table Q, since we can simulate an execution of $\mathbf{D}$, answering each query X_i with Y_i, to obtain the corresponding queries $X_1, \dots, X_N$, and from these Q is fully determined by the assumption on the queries of $\mathbf{D}$ being non-redundant. Also given the function table of $\mathbf{Q}$, the variables $Y_1, \dots, Y_q$ are fully determined. Hence,

$$\mathsf{H}^{\mathbf{D}}(\mathbf{Q}) = \mathsf{H}(Y_1, Y_2, \dots, Y_N) = \mathsf{H}(Q) \ge \log N! - \delta,$$

which implies $d(N) \le \delta$.

To prove (ii), fix some $i > 0$, and let $\mathbf{D}$ be such that it makes i queries $X_1, \dots, X_i$, and maximizes $\log N^{(i)} - \mathsf{H}_{\mathbf{D}}(\mathbf{Q})$. Then, for any $\mathbf{D}'$ which acts as $\mathbf{D}$ for its first i queries, and makes an additional non-redundant query $X_{i+1} \notin \{X_1, \dots, X_i\}$ we have (with $Y_1, \dots, Y_{i+1} \in \mathcal{X}$ being the corresponding answers)

$$\begin{aligned} d(i+1) &\ge \log N^{(i+1)} - \mathsf{H}(Y_1, \dots, Y_i, Y_{i+1}) \\ &= \underbrace{\log N^{(i)} - \mathsf{H}(Y_1, \dots, Y_i)}_{=d(i)} + \underbrace{\log(N-i) - \mathsf{H}(Y_{i+1} \,|\, Y_1, \dots, Y_i)}_{\ge 0}, \end{aligned}$$

since $\mathsf{H}(Y_{i+1} \,|\, Y_1, \dots, Y_i) \le \log(N-i)$, due to the fact that i input-output pairs for $\mathbf{Q}$ are determined by $Y_1, \dots, Y_i$ (and $\mathbf{D}$), and thus at most $N-i$ values are possible for Y_{i+1}. □

The following claim shows that since d grows overall at most by δ on its domain (and it is monotone), then there must be a portion of length q which is almost flat among the first Λ values i.

Claim. There exists $i^* \in \{0, \dots, \Lambda\}$ such that

$$d(i^* + q) - d(i^*) \le \frac{q\delta}{\Lambda}.$$

Proof. Define $\lambda := \left\lfloor \frac{\Lambda}{q} \right\rfloor$ as well as the set $\mathcal{S} := \{i \cdot q \,:\, i = 0, \dots, \lambda\} \subseteq \{0, \dots, \Lambda\}$. Clearly, $d(\lambda \cdot q + q) \le d(N) \le \delta$, and hence there must be an i^* in $\mathcal{S}$ such that $d(i^* + q) - d(i^*) \le \frac{\delta}{\lambda+1}$, as otherwise this would contradict $d(\lambda q + q) \le \delta$. The claim follow by the fact that $\lambda + 1 \ge \frac{\Lambda}{q}$. □

In the following, let i^* as guaranteed to exist by Claim E.3, and take $\overline{\mathbf{D}}$ such that it maximizes $\log N^{(i^*)} - \mathsf{H}^{\overline{\mathbf{D}}}(\mathbf{Q}(X_1) \dots \mathbf{Q}(X_{i^*}))$. Consider now the (two-sided) random permutation $\langle \mathbf{Q} \to \mathbf{P} \rangle$ which behaves as $\langle \mathbf{Q} \rangle$ for the first i^* queries, but then behaves as a randomly chosen permutation consistent with the first i^* queries. Additionally, let $\mathbf{D}$ be an arbitrary q-query deterministic distinguisher issuing non-redundant queries: We define as $\overline{\mathbf{D}} \to \mathbf{D}$ be the distinguisher making $i^* + q$ non-redundant queries, which runs $\overline{\mathbf{D}}$ and then runs $\mathbf{D}$, but does not repeat queries for which the answers is known due to a query of $\overline{\mathbf{D}}$. In particular, it asks possibly extra (dummy) queries in order to ensure that $i^* + q$ queries are always asked.

Claim (A). For all deterministic distinguishers $\mathbf{D}$ issuing q queries, we have

$$\Delta^{\mathbf{D}}(\langle \mathbf{Q} \rangle, \langle \mathbf{Q}_{\overline{\mathbf{D}}} \rangle) \le d\left(\mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}} \to \mathbf{D})\langle \mathbf{Q} \rangle}, \mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}} \to \mathbf{D})\langle \mathbf{Q} \to \mathbf{P} \rangle} \right),$$

where $\mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}} \to \mathbf{D})\langle \mathbf{Q} \rangle}$ and $\mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}} \to \mathbf{D})\langle \mathbf{Q} \to \mathbf{P} \rangle}$ are the distributions of the answers of the queries of $\overline{\mathbf{D}} \to \mathbf{D}$ to the given systems when interacting with the systems $\langle \mathbf{Q} \rangle$ and $\langle \mathbf{Q} \to \mathbf{P} \rangle$, respectively.

Proof. Given the value of Y^{i^*+q} for a distinguisher attempting to distinguish $\mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}}\to\mathbf{D})\langle\mathbf{Q}\rangle}$ and $\mathsf{P}_{Y^{i^*+q}}^{(\overline{\mathbf{D}}\to\mathbf{D})\langle\mathbf{Q}\to\mathbf{P}\rangle}$, one possible strategy is to mimic the behavior of $\mathbf{D}$ in the corresponding interaction with $\langle\mathbf{Q}_{\overline{\mathbf{D}}}\rangle$ or $\langle\mathbf{Q}\rangle$. □

Recall that the *Kullback-Leibler divergence* of P_X and P_Y is defined as

$$D(\mathsf{P}_X\|\mathsf{P}_Y) := \sum_x \mathsf{P}_X(x)\cdot\log\left(\frac{P_X(x)}{P_Y(x)}\right).$$

Using the shorthands $\mathsf{P}^{(1)}$ and $\mathsf{P}^{(2)}$ for $\mathsf{P}^{(\overline{\mathbf{D}}\to\mathbf{D})\langle\mathbf{Q}\rangle}$ and $\mathsf{P}^{(\overline{\mathbf{D}}\to\mathbf{D})\langle\mathbf{Q}\to\mathbf{P}\rangle}$, respectively, we have

$$\begin{aligned} D\left(\mathsf{P}^{(1)}_{Y^{i^*+q}}\middle\|\mathsf{P}^{(2)}_{Y^{i^*+q}}\right) &= \sum_{y^{i^*+q}} \mathsf{P}^{(1)}_{Y^{i^*+q}}(y^{i^*+q})\cdot\log\left(\frac{\mathsf{P}^{(1)}_{Y^{i^*+q}}(y^{i^*+q})}{\mathsf{P}^{(2)}_{Y^{i^*+q}}(y^{i^*+q})}\right) \\ &= \sum_{y^{i^*}} \mathsf{P}^{(1)}_{Y^{i^*}}(y^{i^*})\cdot\sum_{y^q}\mathsf{P}^{(1)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})\cdot\log\left(\frac{\mathsf{P}^{(1)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})}{\mathsf{P}^{(2)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})}\right) \end{aligned}$$

where we have used the fact that $\mathsf{P}^{(1)}_{Y^{i^*}} = \mathsf{P}^{(2)}_{Y^{i^*}}$, since the first i^* queries are answered by $\langle\mathbf{Q}\rangle$ in both cases. Furthermore, we have $\log(\mathsf{P}^{(2)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})) = -\log\left[(N-i^*)\cdots(N-i^*-q+1)\right]$. Also note that

$$\sum_{y^{i^*}} \mathsf{P}^{(1)}_{Y^{i^*}}(y^{i^*})\cdot\sum_{y^q}\mathsf{P}^{(1)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})\cdot\log\left(\mathsf{P}^{(1)}_{Y_{i^*+1}\cdots Y_{i^*+q}|Y^{i^*}}(y^q,y^{i^*})\right) = \mathsf{H}(Y_{i^*+1},\ldots,Y_{i^*+q}|Y^{i^*}).$$

Therefore, replacing this in the above, and using that $\mathsf{H}(Y_{i^*+1},\ldots,Y_{i^*+q}|Y^{i^*}) = \mathsf{H}(Y^{i^*+q}) - \mathsf{H}(Y^{i^*})$,

$$\begin{aligned} D\left(\mathsf{P}^{(1)}_{Y^{i^*+q}}\middle\|\mathsf{P}^{(2)}_{Y^{i^*+q}}\right) &= \log\left[(N-i^*)\cdots(N-i^*-q+1)\right] - \mathsf{H}(Y_{i^*+1},\ldots,Y_{i^*+q}|Y^{i^*}) \\ &= \log N^{(i^*+q)} - \mathsf{H}(Y^{i^*+q}) - \left(\log N^{(i^*)} - \mathsf{H}(Y^{i^*})\right) \le d(i^*+q) - d(i^*) \le \frac{q\delta}{\Lambda}. \end{aligned}$$

It is a well-known fact that any two distributions P_X and P_Y satisfy $d(\mathsf{P}_X,\mathsf{P}_Y) \le \sqrt{D(\mathsf{P}_X\|\mathsf{P}_Y)}$ (cf. e.g. [5]). This in particular (combined with Claim (A)) implies that for all deterministic $\mathbf{D}$ issuing q queries we have $\Delta^{\mathbf{D}}(\langle\mathbf{Q}\rangle,\langle\mathbf{Q}_{\overline{\mathbf{D}}}\rangle) \le \sqrt{\frac{q\delta}{\Lambda}}$, which is a bound for *any* distinguisher issuing q queries, since a randomized distinguisher can never outperform a deterministic one.

F Uniform Hardcore Lemmas

All results of this paper are given in the non-uniform setting. On the one hand, this considerably simplifies the presentation of the paper and permits to convey the main ideas. On the other hand, we feel that the non-uniform model has established itself as the more relevant adversarial model in cryptography, as in particular it allows to make *concrete* statements, and it is questionable whether the higher complexity involved in uniform reductions is justified. Still, it is an interesting question to find out whether a uniform reduction is indeed possible. This section discusses uniform versions of the results of this paper. We omit full proofs of the statements, and only briefly illustrate the main modifications of the existing proofs in order to derive the corresponding uniform statements.

In the following, we say that a (uniform) algorithm is *efficient* (or polynomial-time) if its running time is polynomial in the (understood) security parameter n *and* the total length of the inputs it obtains.

In particular, we use poly as the placeholder for a polynomial function in both these parameters. (In line with this, $\Delta_{\mathsf{poly}}(\mathbf{S}, \mathbf{T})$ and $\mathrm{Guess}_{\mathsf{poly}}(B \mid \mathbf{S})$ denote the best advantages for polynomial-time distinguishers / adversaries.) Finally, we remark that even though all statements are asymptotic in nature, we omit security parameters to a large extent in order to keep the presentation simpler. This should not cause any confusion.

F.1 The Uniform Hardcore Lemma for System-Bit Pairs

In the following, we redefine

$$\psi_{\mathsf{hc}} := \frac{7200}{\gamma^2(1-\varepsilon)^4\zeta_1^2} \cdot \ln\left(\frac{2}{\zeta_2}\right).$$

The statement of Theorem 1 can be translated to the uniform setting following the same lines of Holenstein's uniform hardcore lemma [21].

Theorem 6 (Uniform Hardcore Lemma for System-Bit Pairs). *Let $(\mathbf{S}, B) \equiv (\mathbf{S}(S), B(S))$ be a cc-stateless system-bit pair with an efficient implementation $A_{(\mathbf{S},B)}$ with space complexity $s_{A_{(\mathbf{S},B)}}$. Furthermore, for some $\varepsilon \in [0,1)$ (with $1-\varepsilon$ noticeable)*

$$\mathrm{Guess}_{\mathsf{poly}}(B \mid \mathbf{S}) \leq \varepsilon.$$

For all noticeable $\zeta_1 > 0$, all $\zeta_2 = 2^{-\mathsf{poly}(k)} > 0$, and all $0 < \gamma \leq \frac{1}{2}$ (such that $\frac{2\zeta_1}{1-\varepsilon} + \zeta_2 \leq \frac{\gamma}{4}$ and $1-\varepsilon-\zeta_1-\zeta_2$ is noticeable), and for all polynomial-time q'-query (uniform) oracle adversaries $\mathbf{A}^{(\cdot)}$, there exists a measure $\mathcal{M}$ for $(\mathbf{S}, B)$ with $\mu(\mathcal{M}) \geq 1-\varepsilon$ such that the following two properties are satisfied:

(i) There exists a (ζ_1, ζ_2)-sampler O for $\mathcal{M}$ and $A_{(\mathbf{S},B)}$ with length $\ell := s_{A_{(\mathbf{S},B)}}(\psi_{\mathsf{hc}} \cdot q')$. In particular, if $(\mathbf{S}(s), B(s))$ is deterministic for all s, then O is a $(0,0)$-sampler for $\mathcal{M}$ and $A_{(\mathbf{S},B)}$ with length $s_{A_{(\mathbf{S},B)}}((256 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3}) \cdot q')$.

(ii) For $S' \xleftarrow{\$} \mathcal{M}$, [17]

$$\mathrm{Guess}^{\mathbf{A}^{\mathsf{O}}}(B(S') \mid \mathbf{S}(S')) \leq \gamma.$$

We note that the statement is weaker than the one of Theorem 1 in that we only provide one good measure for every adversary. This is however sufficient for most applications. Note that the length of the states returned by the sampler only depends on the number of queries q', but is independent of the time complexity of $\mathbf{A}$ (and the length of oracle queries $\mathbf{A}$ expects). Otherwise, it would be possible that the statement only holds for adversaries which cannot retrieve the output of the sampler.

Rather than giving a full proof, we opt for pointing out the modifications to be made to the proof of Theorem 1 to obtain a uniform reduction. We postpone a full analysis of the success probability (which relies on standard techniques similar to the ones of [21]) of the reduction to a later version of this paper. We only consider the general case where $(\mathbf{S}(s), B(s))$ is possibly randomized. (The deterministic statement will follow in the same way as in the non-uniform case.) Once again, we proceed by contradiction, and the corresponding initial assumption, which we tacitly make from now on, is modified as follows:

> **Assumption ¬HC'.** There exist noticeable ζ_1, γ, and $\zeta_2 = 2^{-\mathsf{poly}(n)}$ (with the property that $\frac{2\zeta_1}{1-\varepsilon} + \zeta_2 \leq \frac{\gamma}{4}$ and $1-\varepsilon-\zeta_1-\zeta_2$ is noticeable), as well as a polynomially bounded q', and an efficient q'-query oracle adversary $\mathbf{A}^*$ which satisfies
>
> $$\mathrm{Guess}^{\mathbf{A}^{\mathsf{O}}}(B(S') \mid \mathbf{S}(S')) > \gamma$$

[17] That is, $\mathbf{A}$ is given access to an oracle which upon each invocation ignores its input and returns a state sampled according to the state sampler.

Procedure GoodEnough($\mathcal{A}$): *// collection of adversaries $\mathcal{A}$*

1: **if** $\mathcal{A} = \emptyset$ **then return** `false`
2: $\overline{\rho} :=$ Estimate-$\rho(\mathcal{A})$
3: $\overline{p} :=$ Estimate-$p(\mathcal{A})$
4: **if** $\overline{\rho} \geq 1 - \frac{1-\varepsilon}{3} \vee \overline{p} \geq \frac{1}{2} + \frac{5}{32 \cdot (1-\varepsilon) \cdot |\mathcal{A}| \gamma}$ **then**
5: **return** `true`
6: **else**
7: **return** `false`

Fig. 6. Procedure GoodEnough in the proof of the uniform Hardcore Lemma.

for all measures $\mathcal{M}$ for $(\mathbf{S}, B)$ with $\mu(\mathcal{M}) \geq 1 - \varepsilon$ admitting a (ζ_1, ζ_2) state-sampler O for $A_{(\mathbf{S},B)}$ with length ℓ, and $S' \xleftarrow{\$} \mathcal{M}$.

The first step is to transform $\mathbf{A}^*$ into a more convenient form. (We omit the proof of the following lemma, which follows the lines of a similar statement in [21].)

Lemma 10. *There exists an efficient oracle adversary $\mathbf{B}^*$ which, for all measures $\mathcal{M}$ for $(\mathbf{S}, B)$ with $\mu(\mathcal{M}) \geq 1 - \varepsilon$ admitting a (ζ_1, ζ_2)-sampler O for $A_{(\mathbf{S},B)}$ with length ℓ, outputs with overwhelming probability, given oracle access to O, an efficient deterministic q'-query adversary $\mathbf{A}$ (making no oracle queries) such that*

$$\mathrm{Guess}^{\mathbf{A}}(B(S') \,|\, \mathbf{S}(S')) > \gamma/6$$

for $S' \xleftarrow{\$} \mathcal{M}$.

To ensure a consistent notation with the non-uniform case, in the remainder of this section the parameter γ will equal $\gamma/6$ for the "actual" γ that appears in Assumption ¬HC'. The security reduction implements a polynomial-time adversary $\mathbf{A}'$ with $\mathrm{Guess}^{\mathbf{A}'}(B(S) \,|\, \mathbf{S}(S)) > \varepsilon$ by initially running the procedure FindCollection, using the oracle adversary $(\mathbf{B}^*)^{\mathsf{O}}$ at Line 5 to obtain an adversary $\mathbf{A}$ to be added to the collection, and the state-sampler O for $\mathcal{M}_{\mathcal{A},\tau}$ accessed by $\mathbf{B}^*$ is simulated (also in polynomial-time) using the procedure given in the proof of Lemma 3. The final collection $\mathcal{A}$ output when GoodEnough returns `true` is used according to $\mathbf{A}^{(A)}$ or $\mathbf{A}^{(B)}$ depending on whether (a variation of) Condition (A) or Condition (B) is satisfied.

However, the parameters ρ, p (both in GoodEnough), and α^* (in $\mathbf{A}^{(B)}$) cannot be computed exactly: The procedure GoodEnough is consequently modified, as in Figure 6, so that these parameters are estimated and the conditions in the **if**-statement are adapted accordingly. The associated procedures Estimate-ρ and Estimate-p are described in detail in Figure 7. In both descriptions, we assume without loss of generality that $A_{(\mathbf{S},B)}$ first outputs the bit B, and then simulates $\mathbf{S}$. It can be shown that both procedures ensure that GoodEnough returns `true` whenever one of Conditions (A) or (B) is satisfied, which implies that the analysis of termination for FindCollection remains unchanged. Also, to take into account possible estimation errors, $\mathbf{A}^{(A)}$ can simply be modified to work under the condition that $\mathsf{P}\left[N_{\mathcal{A}}(S) > \frac{\gamma}{2} |\mathcal{A}|\right] \geq 1 - \frac{5(1-\varepsilon)}{12}$ by setting φ_A appropriately. Also, in the same way, we can ensure that its advantage is noticeably higher than ε in order to compensate for the negligible error in the reduction.

Additionally, we need to adapt $\mathbf{A}^{(B)}$ to only use the guarantee that $p \geq \frac{1}{2} + \frac{1}{16 \cdot (1-\varepsilon) \cdot |\mathcal{A}| \cdot \gamma}$ (which is unproblematic), and furthermore, we need to additionally estimate α^*: This can be done as the computation of $\overline{p}_{i^*}$ in Estimate-p, we the difference that we compute

$$\overline{\alpha}_i := 2 \cdot \left(\frac{1}{r'} \sum_{j=1}^{r'} b_j - 1 \right),$$

Procedure Estimate-$\rho(\mathcal{A})$: *// collection of adversaries $\mathcal{A}$*

$r := \left(\frac{12}{1-\varepsilon}\right)^2 \cdot n$
$r' := \left(\frac{64}{\gamma}\right)^2 \cdot n$
for all $i = 1, \ldots, r$ **do**
  $b \xleftarrow{\$} A_{(\mathbf{S},B)}[\perp]$
  $\sigma_0 :=$ state of $A_{(\mathbf{S},B)}$ after outputting b
  for all $j = 1, \ldots, r'$ **do**
    $\mathbf{A}_{i,j} \xleftarrow{\$} \mathcal{A}$
    $b_j := \mathbf{A}_{i,j}(A_{(\mathbf{S},B)}[\sigma_{j-1}]) \oplus b \oplus 1$
    $\sigma_j :=$ last state of $A_{(\mathbf{S},B)}[\sigma_{j-1}]$.
  $\overline{N}_i := 2 \cdot |\mathcal{A}| \cdot \left(\frac{1}{r'} \sum_{j=1}^{r'} b_j - \frac{1}{2}\right)$.
$\overline{\rho} := \frac{\left\{i : \overline{N}_i > \frac{3\gamma}{4} \cdot |\mathcal{A}|\right\}}{r}$
return $\overline{\rho}$

Procedure Estimate-$p(\mathcal{A})$: *// collection of adversaries $\mathcal{A}$*

$r := 1024 \cdot \gamma^2 \cdot |\mathcal{A}|^2 \cdot n$
$r' := 1024 \cdot (1-\varepsilon)^2 \cdot \gamma^2 \cdot |\mathcal{A}|^2 \cdot n$
for all $i = 1, \ldots, r$ **do**
  $b \xleftarrow{\$} A_{(\mathbf{S},B)}[\perp]$
  $\sigma_0 :=$ state of $A_{(\mathbf{S},B)}$ after outputting b
  for all $j = 1, \ldots, r'$ **do**
    $\mathbf{A}_{i,j} \xleftarrow{\$} \mathcal{A}$
    $b_j := \mathbf{A}_{i,j}(A_{\mathbf{S}}[\sigma_{j-1}]) \oplus b \oplus 1$
    $\sigma_j :=$ last state of $A_{(\mathbf{S},B)}[\sigma_{j-1}]$
  $\overline{p}_i := \frac{1}{r'} \sum_{j=1}^{r'} b_j$.
reorder $\overline{p}_1, \ldots, \overline{p}_r$ such that $\overline{p}_i \leq \overline{p}_{i+1}$ for all $i = 1, \ldots, r-1$
$i^* := (1-\varepsilon)r$
$\overline{p} := \frac{1}{i^*-1} \sum_{i=1}^{i^*-1} \overline{p}_i$
return $\overline{p}$

Fig. 7. Procedures Estimate-ρ and Estimate-p in the proof of the uniform Hardcore Lemma.

and take α_{i^*} as an approximation of α^*. Also, again we have to ensure that the advantage is noticeably higher than ε to compensate for the negligible error probability. We omit the details.

F.2 The Uniform Hardcore Lemma for Computational Indistinguishability

We state a uniform version of the hardcore lemma for computational indistinguishability.

Theorem 7 (Uniform Hardcore Lemma for Computational Indistinguishability). *Let $\mathbf{S} \equiv \mathbf{S}(S)$ and $\mathbf{T} \equiv \mathbf{T}(T)$ be cc-stateless systems, with respective efficient implementations $A_{\mathbf{S}}$ and $A_{\mathbf{T}}$. Furthermore, assume that for some $\varepsilon \in [0,1)$ (such that $1-\varepsilon$ is noticeable),*

$$\Delta_{\mathsf{poly}}(\mathbf{S}, \mathbf{T}) \leq \varepsilon.$$

For all noticeable $\zeta_1 > 0$, all $\zeta_2 = 2^{-\mathsf{poly}(k)} > 0$, and all $0 < \gamma \leq \frac{1}{2}$ (such that $\frac{2\zeta_1}{1-\varepsilon} + \zeta_2 \leq \frac{\gamma}{4}$ and $1-\varepsilon-\zeta_1-\zeta_2$ is noticeable), and for all polynomial-time q'-query (uniform) oracle distinguishers $\mathbf{D}^{(\cdot,\cdot)}$, there exist measures $\mathcal{M}_{\mathbf{S}}$ and $\mathcal{M}_{\mathbf{T}}$ such that $\mu(\mathcal{M}_{\mathbf{S}}) \geq 1-\varepsilon$ and $\mu(\mathcal{M}_{\mathbf{T}}) \geq 1-\varepsilon$ and the following properties hold:

(i) There exist (ζ_1,ζ_2)-samplers $\mathsf{O}_{\mathbf{S}}$ and $\mathsf{O}_{\mathbf{T}}$ for $\mathcal{M}_{\mathbf{S}}$ and $A_{\mathbf{S}}$ as well as for $\mathcal{M}_{\mathbf{T}}$ and $A_{\mathbf{T}}$, respectively, with length $\ell := \max\{s_{A_{\mathbf{S}}}(\psi_{\mathsf{hc}} \cdot q'), s_{A_{\mathbf{T}}}(\psi_{\mathsf{hc}} \cdot q')\}$. Furthermore, if both $\mathbf{S}$ and $\mathbf{T}$ are random functions, then both samplers can be made error-less with $\psi_{\mathsf{hc}} := 256 \cdot \gamma^{-2} \cdot (1-\varepsilon)^{-3}$.

(ii) For $S' \xleftarrow{\$} \mathcal{M}_{\mathbf{S}}$ and $T' \xleftarrow{\$} \mathcal{M}_{\mathbf{T}}$,

$$\Delta^{\mathbf{D}^{\mathsf{O}_{\mathbf{S}},\mathsf{O}_{\mathbf{T}}}}(\mathbf{S}(S'),\mathbf{T}(T')) \leq 13\gamma.$$

Assume that Theorem 7 is false, that is:

> There exist noticeable $\zeta_1 > 0$, $\zeta_2 = 2^{-\mathsf{poly}(n)} > 0$, and $0 < \gamma \leq \frac{1}{2}$ (such that $\frac{2\zeta_1}{1-\varepsilon} + \zeta_2 \leq \frac{\gamma}{4}$ and $1-\varepsilon-\zeta_1-\zeta_2$ is noticeable), a polynomially bounded q', and an efficient q'-query oracle distinguisher $\mathbf{D}^* = (\mathbf{D}^*)^{(\cdot,\cdot)}$ which satisfies
>
> $$\Delta^{(\mathbf{D}^*)^{\mathsf{O}_{\mathbf{S}},\mathsf{O}_{\mathbf{T}}}}(\mathbf{S}(S'),\mathbf{T}(T')) > 13\gamma$$
>
> *for all* measures $\mathcal{M}_{\mathbf{S}}$ and $\mathcal{M}_{\mathbf{T}}$, both with density at least $1-\varepsilon$, admitting (ζ_1,ζ_2)-state-samplers $\mathsf{O}_{\mathbf{S}}$ and $\mathsf{O}_{\mathbf{T}}$ for the corresponding implementations $A_{\mathbf{S}}$ and $A_{\mathbf{T}}$ with length ℓ, and for $S' \xleftarrow{\$} \mathcal{M}_{\mathbf{S}}$, $T' \xleftarrow{\$} \mathcal{M}_{\mathbf{T}}$.

Then, we prove that for the cc-stateless system-bit pair $(\mathbf{F},B)$ defined as in the proof of Theorem 2, there exists an efficient q'-query oracle adversary $\mathbf{A}^* = (\mathbf{A}^*)^{(\cdot)}$ such that *for all* measures $\mathcal{M}$ for $(\mathbf{F},B)$ with density at least $1-\varepsilon$ admitting a (ζ_1,ζ_2)-state-sampler O for $A_{(\mathbf{F},B)}$ with length ℓ,

$$\mathrm{Guess}^{(\mathbf{A}^*)^{\mathsf{O}}}(B' \,|\, \mathbf{F}(X',B')) > \gamma$$

for $(X',B') \xleftarrow{\$} \mathcal{M}$, except with negligible probability. As this suffices to contradict the uniform HCL for the system-bit pair $(\mathbf{F},B)$ (Theorem 6),[18] there exists a uniform polynomial-time distinguisher $\mathbf{D}$ such that

$$\Delta^{\mathbf{D}}(\mathbf{S},\mathbf{T}) \geq \mathrm{Guess}^{\mathbf{D}}(B \,|\, \mathbf{F}) > \varepsilon,$$

contradicting the indistinguishability of $\mathbf{S}$ and $\mathbf{T}$.

The adversary $\mathbf{A}^*$, given oracle access to the sampler O for $\mathcal{M}$, proceeds as follows: It first produces an estimate $\overline{p}_0$ of the probability p_0 that $B' = 0$ using O. Then, if B' takes a value $b \in \{0,1\}$ with probability substantially larger than $\frac{1}{2}$, it outputs b. Otherwise, it simply lets the distinguisher $\mathbf{D}^*$ interact with the given system $\mathbf{F}(X',B')$, with oracles $\mathsf{O}_{\mathbf{S}}$ and $\mathsf{O}_{\mathbf{T}}$ which are (ζ_1,ζ_2)-samplers for the measures $\mathcal{M}_{\mathbf{S}}$ and $\mathcal{M}_{\mathbf{T}}$ defined as

$$\mathcal{M}_{\mathbf{S}}(s) := \frac{\varepsilon}{\varepsilon+3\gamma(1-\varepsilon)} \cdot \mathcal{M}_0(s) + \frac{3\gamma(1-\varepsilon)}{\varepsilon+3\gamma(1-\varepsilon)}$$
$$\mathcal{M}_{\mathbf{T}}(t) := \frac{\varepsilon}{\varepsilon+3\gamma(1-\varepsilon)} \cdot \mathcal{M}_1(t) + \frac{3\gamma(1-\varepsilon)}{\varepsilon+3\gamma(1-\varepsilon)}.$$

where $\mathcal{M}_0$ and $\mathcal{M}_1$ as in the proof of Theorem 2. The sampler $\mathsf{O}_{\mathbf{S}}$ samples $((\sigma_1,b_1),z_1),((\sigma_2,b_2),z_2),\ldots$ from O until $b_i = 0$ is satisfied, and then $(\sigma_i, \frac{\varepsilon}{\varepsilon+3\gamma(1-\varepsilon)} z_i + \frac{3\gamma(1-\varepsilon)}{\varepsilon+2\delta(1-\varepsilon)})$ is output. If after n attempts

[18] As discussed in the proof of Theorem 6, the final adversary can be modified to compensate for a negligible error probability.

no pair with $b_i = 0$ is returned, then it outputs $(\perp, 1)$. (For $\mathcal{M}_{\mathbf{T}}$ the sampler $\mathsf{O}_{\mathbf{T}}$ symmetrically uses $b_i = 1$.) This yields, along the same lines as in the discussion at the end of the proof of Theorem 2, (ζ_1, ζ_2)-samplers for both measures, provided that a pair with $b_i = 0$ (for $\mathsf{O}_{\mathbf{S}}$) or $b_i = 1$ (for $\mathsf{O}_{\mathbf{T}}$) occurs within the n samples from O.

Finally, $\mathbf{A}^*$ determines (using an estimate with an appropriate error) whether the probability that $(\mathbf{D}^*)^{\mathsf{O}_{\mathbf{S}},\mathsf{O}_{\mathbf{T}}}$ outputs the correct bit is at least $\frac{1}{2}$. In the affirmative case, $\mathbf{A}^*$ returns the output bit of $(\mathbf{D}^*)^{\mathsf{O}_{\mathbf{S}},\mathsf{O}_{\mathbf{T}}}$, whereas otherwise its output bit is flipped before its returned.

F.3 Security Amplification for Weak PRPs

A uniform version of Theorem 4 can be given for the slightly weaker result that the cascade of m instances of a (two-sided) ε-PRP $E = \{E_k\}_{k\in\{0,1\}^n}$ is an $(\varepsilon^m + m(1-\varepsilon)\varepsilon^{m-1} + \nu)$-(two-sided) PRP for a negligible function ν in the security parameter n, as long as $\varepsilon \leq 1 - \frac{1}{n^{\mathcal{O}(1)}}$.

More concretely, assume that we are given a cc-stateless random permutation $\mathbf{Q} \equiv \mathbf{Q}(Q)$ (with efficient implementation $A_{\mathbf{Q}}$) and a URP $\mathbf{P} \equiv \mathbf{P}(P)$ (with canonical implementation $A_{\mathbf{P}}$), both with the same domain. Furthermore, let $\mathbf{D}$ be a polynomial-time distinguisher such that, for m independent instances $\mathbf{Q}_1, \ldots, \mathbf{Q}_m$ of $\mathbf{Q}$,

$$\Delta^{\mathbf{D}}(\mathbf{Q}_1 \rhd \cdots \rhd \mathbf{Q}_m, \mathbf{P}) > \varepsilon^m + m(1-\varepsilon)\varepsilon^{m-1} + \gamma(n),$$

for a noticeable function γ and for infinitely many values of the security parameter n. Then, we provide a polynomial-time (oracle) distinguisher $\mathbf{D}'$ with the following guarantee: For all measures $\mathcal{M}_{\mathbf{Q}}$ for $\mathbf{Q}$ and $\mathcal{M}_{\mathbf{P}}$ for $\mathbf{P}$, both with density at least $1-\varepsilon$, and both admitting error-less samplers $\mathsf{O}_{\mathbf{Q}}$ and $\mathsf{O}_{\mathbf{P}}$ with polynomial-length for the respective efficient implementations, we have

$$\Delta^{\mathbf{D}'^{\mathsf{O}_{\mathbf{Q}},\mathsf{O}_{\mathbf{P}}}}(\mathbf{Q}(Q'), \mathbf{P}(P')) > \gamma'(n),$$

for a noticeable function γ', $Q' \xleftarrow{\$} \mathcal{M}_{\mathbf{Q}}$, and $P' \xleftarrow{\$} \mathcal{M}_{\mathbf{P}}$. It is not hard to see that this contradicts Theorem 7, since (by an appropriate choice of parameters) there must be measures for $\mathbf{Q}$ and $\mathbf{P}$, with density $1-\varepsilon$ and with corresponding error-less samplers, for which $\Delta^{\mathbf{D}'^{\mathsf{O}_{\mathbf{Q}},\mathsf{O}_{\mathbf{P}}}}(\mathbf{Q}(Q'), \mathbf{P}(P')) \leq \gamma'(n)$.

The distinguisher $\mathbf{D}'$ operates as follows, given $\mathbf{S} \in \{\mathbf{Q}(Q'), \mathbf{P}(P')\}$. It starts by setting $\mathcal{G} := \emptyset$, issuing m queries to $\mathsf{O}_{\mathbf{Q}}$, obtaining pairs (σ_i, z_i), and for each $i \in \{1, \ldots, m\}$ it adds i to $\mathcal{G}$ with probability z_i. Furthermore, it picks an index i^* uniformly from $\{1, \ldots, m\}$, Then, if $i^* \in \mathcal{G}$, it behaves as

$$\mathbf{D}(\mathbf{C}(S_1, \ldots, S_{i^*-i}, \cdot, S_{i^*+1}, \ldots, S_m))$$

and otherwise just outputs $\mathbf{D}(\mathbf{C}(S_1, \ldots, S_m))$, where, for all $i \in \{1, \ldots, m\}$, we have $S_i := \mathbf{P}(P'_i)$ for $P'_i \xleftarrow{\$} \mathsf{P}_{\mathcal{M}_{\mathbf{P}}}$ if $i \in \mathcal{G}$ and $i < i^*$, and $S_i := A_{\mathbf{Q}}[\sigma_i]$ otherwise.

In particular, $\mathbf{P}(P'_i)$ is simulated as $A_{\mathbf{P}}[\Sigma]$, where Σ is generated by repeating at most $(n + \log(m))/\log\left(\frac{1}{\varepsilon}\right)$ the following: First $(\sigma, z) \xleftarrow{\$} \mathsf{O}_{\mathbf{P}}$ is sampled, and then $\Sigma := \sigma$ is output with probability z. If no attempts outputs Σ, we set $\Sigma := \perp$. This procedure outputs a state with the correct distribution, except with probability $\frac{2^{-n}}{m}$, and hence the probability that it ever outputs a state with the wrong distribution is at most 2^{-n}.

Remark 4. We finally remark that our argument for decreasing the bound by an multiplying with an additional factor ε seems to be inherently non-uniform, and we are not aware how to achieve this (slightly) stronger statement in the uniform setting.

.