ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Contributions to the Theory of Probabilistic Discrete Systems

Master's Thesis

David Lanzenberger

August 13, 2019

Supervisor and Advisor: Prof. Dr. Ueli Maurer
Department of Computer Science, ETH Zürich

Abstract

A probabilistic discrete system (PDS) is an abstract object operating in rounds. In every round, an environment (which is a complex object like a PDS) can input a value and the system responds with an output value. A PDS may be probabilistic and each round may depend on the previous rounds. Many cryptographic systems (which can be modeled as probabilistic discrete systems) have security definitions based on an environment interacting with the system, essentially modeling the adversary. For example, a system is defined to be secure if it is indistinguishable from a certain ideal system for any environment, leading to the notion of a distinguisher.

Recently, Maurer proposed a novel paradigm called *environment-less*, in which properties of systems are expressed as intrinsic properties of the systems as objects themselves, free of the notion of an environment or an adversary. The paradigm gives new insight into the very essence of the properties and enables more minimal and abstract reasoning about systems.

This work makes the first steps towards an environment-less (cryptographic) systems theory. We show that two key properties of cryptographic systems, namely the indistinguishability of two systems and the optimal winning probability of a game, can be stated *equivalently* and naturally within the environment-less paradigm. Our treatment is abstract: we merely assume that an object of a set $\mathcal{A}$ is observable by one function (or projection) of a set $\mathcal{F}$. As a consequence, our results are applicable to a broad class of system types and beyond.

Furthermore, we present a new variant of Maurer's theory of discrete systems. In contrast to Maurer's representation, we define discrete systems as inductive objects. We show how this new representation allows to prove elementary statements about systems in a rigorous and formalizable manner.

Finally, we use environment-less indistinguishability to prove a new indistinguishability amplification theorem in an elementary fashion, generalizing previous results. This demonstrates that the environment-less paradigm is not only of conceptual interest but a powerful technical tool as well.

Acknowledgements

This work has been supported in virtually every imaginable way by Ueli Maurer. His door was always open for me. He was incredibly generous and patient, contributing many ideas and investing an embarrassing amount of time into our meetings. His continuous striving for the *right view* as well as his ability to always ask the right questions resulted in numerous stimulating discussions. He always gave me a lot of freedom to explore what interested me the most and kept me highly motivated throughout the last months. All of this is absolutely incredible and I am deeply grateful to be given the opportunity of working with him.

Moreover, I thank Martin for the exchanges on various topics such as formal methods and discrete systems.

The enduring support and continuous encouragement of my family and friends has been essential for my years of study and enabled me to pursue what I enjoy. Thank you.

Contents

Chapter 1

Introduction

1.1 Motivation

In many areas of computer science and in particular in cryptography, we are interested in discrete systems. Informally, a discrete system is an abstract object which operates in rounds. In the i-th round, an input x_i is answered with some output y_i. A discrete system may keep state and may be probabilistic, i.e., the output y_i may depend probabilistically on the inputs and outputs of previous rounds.

One often discusses properties of discrete systems depending on what interaction is allowed with the system, leading to the notion of *environments* and, in cryptography, to the notion of distinguishers. A natural question is what kind of statements which classically involve environments can be expressed equivalently as intrinsic properties of the systems *themselves*, i.e., without the explicit concept of an environment.

We call this the *environment-less* paradigm and explain it informally by two examples.

(i) Imagine a game which is modeled by a probabilistic discrete system $\mathbf{G}$. A player (or a winner) can interact with $\mathbf{G}$ in multiple rounds (see above) and in the final round, $\mathbf{G}$ announces whether the player has won or lost the game. Naturally, we are interested in the maximal winning probability $\nu(\mathbf{G})$ of a given game $\mathbf{G}$, i.e., the probability that player $\mathbf{P}$ wins the game G, maximized over *all* players $\mathbf{P}$. We now claim that

> $\nu(\mathbf{G}) \leq \epsilon$ if and only if there exists $\mathbf{G}'$ equivalent to $\mathbf{G}$ such that with probability at least $1 - \epsilon$, $\mathbf{G}'$ is an always-lose game.

We explain this statement by a concrete example. Let $n \in \mathbb{N}_{\geq 1}$ be a constant and let $\mathbf{G}$ be the game which samples a uniform random number X from the set $\{1, \dots, n\}$. In the first and only round, the player has to input a number $X' \in \{1, \dots, n\}$. The player wins if and only if she guessed the number correctly, i.e., $X' = X$. It is easy to see that the maximum winning probability is $\nu(\mathbf{G}) = \frac{1}{n}$.

Now consider the game $\mathbf{G}'$ which samples a biased bit $win \in \{0, 1\}$ such that $win = 1$ with probability $\frac{1}{n}$. In the first and only round, the player has to input a number $X' \in \{1, \dots, n\}$. The player wins if and only if $win = 1$ (irrespective of the input X').

The games $\mathbf{G}$ and $\mathbf{G}'$ are equivalent: Whatever number is input, the game is won with probability $\frac{1}{n}$. Moreover, with probability $1 - \frac{1}{n}$, $\mathbf{G}'$ is an always-lose game: If $win = 0$, then the game is lost (without even considering the player). Thus, the above claim is satisfied.

Observe that the "if" direction of the claim is unsurprising. If $\mathbf{G}$ is always-lose with probability $1 - \epsilon$, clearly no player can win $\mathbf{G}$ with probability greater than ϵ, thus $\nu(\mathbf{G}) \leq \epsilon$. The "only if" direction, however, is not obvious. Interestingly, it is true even for much more complex games with multiple dependent rounds and even if we allow adaptive players.

Finally, note that a game being always-lose is an intrinsic property of the game as object: We merely need to verify whether it is possible (with non-zero probability) that the game announces the winning event. If this is not possible, the game is always-lose.

(ii) A distinguisher $\mathbf{D}$ is a probabilistic discrete system trying to distinguish between two systems $\mathbf{S}$ and $\mathbf{T}$. In particular, the goal of $\mathbf{D}$ is to interact with the connected system and to guess whether the system is $\mathbf{S}$ or $\mathbf{T}$. The *distinguishing advantage* measures how well a distinguisher can tell the two systems apart. As in the first example, we are interested in the *maximal distinguishing advantage* $\Delta(\mathbf{S}, \mathbf{T})$. We now claim that

$\Delta(\mathbf{S}, \mathbf{T}) \leq \epsilon$ if and only if there exist $\mathbf{S}'$ equivalent to $\mathbf{S}$ and $\mathbf{T}'$ equivalent to $\mathbf{T}$ such that $\mathbf{S}'$ and $\mathbf{T}'$ are *equal* with probability $1 - \epsilon$.

Loosely speaking, "$\mathbf{S}'$ and $\mathbf{T}'$ are equal with probability $1 - \epsilon$" means that there exists a (joint) distribution over pairs of deterministic discrete systems, such that the marginal distributions are $\mathbf{S}'$ and $\mathbf{T}'$, respectively, and both discrete systems are equal with probability $1 - \epsilon$. This can be equivalently expressed using the statistical distance of distributions.

The above claim generalizes the well-known fact that for systems which are just random variables, the maximal distinguishing advantage is simply the statistical distance of the corresponding distributions. As a consequence, the claimed equivalence enables reasoning about the distance of systems without mentioning environments at all, expressing a purely intrinsic property of systems as objects.

The motivation for the outlined way of thinking is at least two-fold:

- First, it gives new insight into what our systems actually are. This can be used to justify classical definitions which are often not very intuitive (e.g., the distinguishing advantage), and it allows to reason about discrete systems in a clearer and more minimal manner.
- Second, it can be understood as a technical lemma and therefore used as a technical tool to prove new statements about environments while completely avoiding environments (with their complexity) in the proof itself.

The long-term goal is to develop a complete (cryptographic) systems theory which is environment-less. This work makes the first steps in this direction by discussing the property of system indistinguishability and a basic variant of game winning probability.

1.2 Related Work

A first version of discrete systems, called *random systems*, is described by Maurer in [9]. An extended theory of deterministic systems has been introduced in [12].

Furthermore, we follow for the most part the paradigms introduced in the theories of Abstract Cryptography [14] and Constructive Cryptography [10]. In particular, we strive for a high level of abstraction. We follow the top-down paradigm and attempt to introduce only essential elements. This approach results in more minimal statements which are at the same time more general.

1.3 Contributions and Outline

This work makes three main contributions, each of which relies on the previous one. We remark, however, that each contribution has been developed to be of independent interest.

- In Chapter 3, we introduce the concept of the *intersection* of (sets of) abstract distributions. Based on this concept, we define *observation compatibility*, which relates the abstract indistinguishability of distributions $\mathbf{X}$ and $\mathbf{Y}$ under a set of functions (or projections) $\mathcal{F}$ with the intersection of corresponding equivalence classes $[\mathbf{X}]_{\mathcal{F}}$ and $[\mathbf{Y}]_{\mathcal{F}}$.
- In Chapter 4, we give a new representation of Maurer's theory of discrete systems. Furthermore, we define a new environment-less and natural distance $\widehat{\Delta}$ on probabilistic discrete systems. We then use the results on observation compatibility to show the Distance Lemma, which states that $\widehat{\Delta}$ is equivalent to the classical distinguishing advantage Δ. Finally,

we provide a new environment-less perspective on the optimal winning probability of a certain type of probabilistic discrete games and also show an equivalence to the conventional (environment-based) definition.

- In Chapter 5, we generalize the notion of neutralizing constructions of [13]. Using the new distance $\widehat{\Delta}$, we then show a new indistinguishability amplification theorem in an elementary fashion. Said theorem generalizes the Product Theorem of [13].

Chapter 2

Preliminaries

2.1 Notation

The natural numbers (or non-negative integers) are denoted by $\mathbb{N} = \{0, 1, \ldots\}$. The integers, rationals, and reals are denoted by $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, respectively, and $\mathbb{R}_{\geq 0} := \{x \mid x \in \mathbb{R}, x \geq 0\}$. For $n \in \mathbb{N}$, $\mathbb{Z}_n$ denotes the set $\{0, \ldots, n-1\}$ and $\oplus_n$ denotes addition modulo n in $\mathbb{Z}_n$. Moreover, for $n \in \mathbb{N}$, we write $[n]$ to denote the set $\{1, 2, \ldots, n\}$, with the convention $[0] = \varnothing$.

The *power set* (or, the set of subsets) of a set $\mathcal{X}$ is denoted by $\mathcal{P}(\mathcal{X})$. The *symmetric difference* of two sets A and B is defined as $A \triangle B := A \cup B - A \cap B$.

The *Cartesian product* of two sets A and B is denoted by $A \times B$. The n-fold Cartesian product $A \times \cdots \times A$ is denoted by A^n. The value at the i-th index of an element $a = (a_1, \ldots, a_n) \in A^n$ is denoted by a_i. For a set A with $0 \in A$, the *hamming weight* $\mathsf{hw}(a)$ of an element $a \in A^n$ is defined by the number of indices with non-zero values.

A (total) *function* from X to Y is a binary relation $f \subseteq X \times Y$ such that for every $x \in X$, there exists a unique $y \in Y$ with $(x, y) \in f$. We usually write $f(x) = y$ instead of $(x, y) \in f$. Anonymous functions are described by a mapping $x \mapsto f_x$ for an expression f_x involving x.

The set of (total) functions from X to Y is denoted by Y^X, and $Y^X_{=} \subseteq Y^X$ denotes the set of constant functions, i.e., functions $f : X \to Y$ for which there exists $y \in Y$ such that $f(x) = y$ for all $x \in X$. Moreover, we denote the domain of a function $f \in Y^X$ by $\mathsf{dom}(f)$. A *partial function* f from X to Y, denoted by $f : X \nrightarrow Y$, is a total function from X' to Y for some $X' \subseteq X$. $Y^{\subseteq X}$ denotes the set of all partial functions from X to Y. The *support* of a function $f : X \to Y$ with $0 \in Y$ is denoted by $\mathsf{supp}(f)$, i.e., $\mathsf{supp}(f) := \{x \mid x \in X, f(x) \neq 0\}$. The *preimage* (or inverse image) of a function $f : X \to Y$ for $y \in Y$ is the *set* of all elements that f maps to y, i.e., $f^{-1}(y) := \{x \mid x \in X, f(x) = y\}$. *Function composition* is denoted by the symbol $\circ$, i.e., for (partial) functions $f : Y \to Z$ and $g : X \to Y$, $f \circ g$ is the (partial) function $x \mapsto f(g(x))$ from X to Z. For three functions $f : A \to B$, $g : A \to B$, and $op : B \to C$, we define $op(f, g) : A \to C$ by $a \mapsto op(f(a), g(a))$. The functions $(a, b) \mapsto a$ and $(a, b) \mapsto b$ are denoted by left and right, respectively.

The *Z-restriction* of a function $f : X \to Y$, denoted by $f^{\cap[Z]}$, is the function from $X \cap Z$ to Y with $f^{\cap[Z]}(x) = f(x)$ for all $x \in X \cap Z$. We will sometimes describe a set Z for an Z-restriction by a pattern. For example, the Z-restriction for $Z = \{(x_1, \ldots, x_k) \mid (x_1, \ldots, x_k) \in X, x_2 = 3\}$ is denoted by $f^{\cap[(\_,3,\ldots)]}$. In such patterns, the underscore _ is used to denote an arbitrary symbol.

A *(U-indexed) tuple* over X is a partial function $f : U \nrightarrow X$ over an index set U. We will often denote a tuple by $\langle x_j \rangle_{j \in \Lambda}$ with the understanding that $\mathsf{dom}(f) = \Lambda$ and $f(j) = x_j$ for $j \in \Lambda$. We call a tuple *finite* or *countable* if its domain is finite or countable, respectively.

A *multiset over* $\mathcal{A}$ is a partial function $M : \mathcal{A} \nrightarrow \mathbb{N}$. We represent multisets in set notation, e.g., $M = \{(a, 2), (b, 7)\}$ denotes the multiset M with domain $\{a, b\}$, $M(a) = 2$, and $M(b) = 7$. The cardinality $|M|$ of a multiset is $\sum_{a \in \mathsf{dom}(M)} M(a)$. The union $\cup$, intersection $\cap$, sum $+$, and difference $-$ of two multisets is defined by the pointwise maximum, minimum, sum, and difference, respectively. Finally, the symmetric difference $M \triangle M'$ of two multisets is defined by $M \cup M' - M \cap M'$.

2.2 Discrete Probability and Statistical Distance

Definition 2.1. A *discrete random experiment* is a pair $\mathcal{E} = (\Omega, \Pr)$, for a non-empty set Ω called the *sample space*, and a probability measure $\Pr : \mathcal{P}(\Omega) \to \mathbb{R}_{\geq 0}$ such that $\Pr(\Omega) = 1$ and for every *event* $E \subseteq \Omega$ we have

$$\Pr(E) = \sum_{e \in E} \Pr(\{e\}).$$

Definition 2.2. For a given sample space Ω, a *random variable* over $\mathcal{X}$ is a function $\mathrm{X} : \Omega \to \mathcal{X}$.

Notation 2.3. A predicate involving random variables $\mathrm{X}_1, \ldots, \mathrm{X}_k$ defines the event $\mathcal{E} \subseteq \Omega$ containing all $\omega \in \Omega$ for which the predicate is true.

For example, we write $\mathrm{X} = \mathrm{Y}$ as a short-hand for the event

$$\{\omega \mid \omega \in \Omega,\ \mathrm{X}(\omega) = \mathrm{Y}(\omega)\}.$$

Definition 2.4. The *probability distribution* $\mathbf{X} : \mathcal{X} \to \mathbb{R}_{\geq 0}$ of a random variable X is defined as follows:

$$\mathbf{X}(x) := \Pr(\mathrm{X} = x) = \Pr(\{\omega \mid \omega \in \Omega, \mathrm{X}(\omega) = x\}) = \sum_{\substack{\omega\in\Omega \\ \mathrm{X}(\omega)=x}} \Pr(\{\omega\}).$$

One can understand a probability distribution $\mathbf{X}$ as a *specification* for a random variable (within some universe).

Notation 2.5. We write $\mathrm{X} \sim \mathbf{X}$ to denote that the distribution of the random variable X is $\mathbf{X}$.

Definition 2.6. The events $E_1, \dots, E_k \subseteq \Omega$ are *independent* if $\Pr(\cap_{i\in[k]} E_i') = \prod_{i\in[k]} \Pr(E_i')$ for all $E_1', \dots, E_k'$ with $E_i' \subseteq E_i$. Analogously, the random variables $\mathrm{X}_1, \dots, \mathrm{X}_k$ are independent if the events $\{X_1 = x_1\}, \dots, \{X_k = x_k\}$ are independent for all $x_1, \dots, x_k$.

Definition 2.7. The *statistical distance* (or total variation distance) of two probability distributions $\mathbf{X}$ and $\mathbf{Y}$ over the same countable set $\mathcal{X}$ is defined by

$$\delta(\mathbf{X}, \mathbf{Y}) := \frac{1}{2} \sum_{x\in\mathcal{X}} |\mathbf{X}(x) - \mathbf{Y}(x)|.$$

The following fact is easily verified and thus stated without proof.

Fact 2.8. *For any two probability distributions* $\mathbf{X}$ *and* $\mathbf{Y}$ *over the same countable set* $\mathcal{X}$ *we have*

$$\begin{aligned}\delta(\mathbf{X}, \mathbf{Y}) &= \sum_{x\in\mathcal{X}} \max(0, \mathbf{X}(x) - \mathbf{Y}(x)) \\ &= \sum_{x\in\mathcal{X}} \max(0, \mathbf{Y}(x) - \mathbf{X}(x)) \\ &= 1 - \sum_{x\in\mathcal{X}} \min\left(\mathbf{X}(x), \mathbf{Y}(x)\right).\end{aligned}$$

Definition 2.9. $\mathrm{Bernoulli}(p) : \{0, 1\} \to \mathbb{R}_{\geq 0}$ for $p \in [0, 1]$ denotes the probability distribution defined by

$$\mathrm{Bernoulli}(p)(x) := \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

2.3 Order Theory

Definition 2.10 (Preorder, Partial Order, Total Order)**.** On a given set $\mathcal{A}$, we define the following types of binary relations:

- A *preorder* is a binary relation $R \subseteq \mathcal{A} \times \mathcal{A}$ that is reflexive and transitive.
- A *partial order* is an antisymmetric preorder.
- A *total order* is a partial order such that for every $(a, b) \in \mathcal{A} \times \mathcal{A}$ either $(a, b) \in R$ or $(b, a) \in R$.

Often, we will denote a binary relation $R \subseteq A \times A$ by the symbol $\leq$ and write $a \leq b$ if and only if $(a, b) \in R$. Moreover, we write $a \geq b$ to denote $b \leq a$.

Definition 2.11. A partially ordered set $(\mathcal{X}, \leq)$ is a

- *meet-semilattice* if the greatest lower bound (or meet, or infimum) of any two-element set $\{a, b\} \subseteq \mathcal{X}$ exists. Unless stated otherwise, the greatest lower bound of $\{a, b\} \subseteq \mathcal{X}$ is denoted by $\inf(a, b)$ and the greatest lower bound of a set $A \subseteq \mathcal{X}$ is denoted by $\inf A$.
- *join-semilattice* if the least upper bound (or join, or supremum) of any two-element set $\{a, b\} \subseteq \mathcal{X}$ exists. Unless stated otherwise, the least upper bound of $\{a, b\} \subseteq \mathcal{X}$ is denoted by $\sup(a, b)$ and the least upper bound of a set $A \subseteq \mathcal{X}$ is denoted by $\sup A$.
- *lattice* if it is both a meet-semilattice and a join-semilattice.

2.4 Algebra

Definition 2.12. A set $\mathcal{M}$ with a (closed) binary operation $+$ is called a *magma*. Moreover,

- a magma is *commutative* if the binary operation $+$ is commutative, and
- a magma is *cancellative* if $a + b = a + c$ or $b + a = c + a$ implies $b = c$ for all $a, b, c \in \mathcal{M}$.

Definition 2.13. A *monoid* is a magma $\mathcal{M}$ with associative binary operation $+$ and an identity element $0 \in \mathcal{M}$.

Definition 2.14. A magma $\mathcal{M}$ is called a *quasigroup* if for every $x, y \in \mathcal{M}$ there exist unique $x_r, x_l \in \mathcal{M}$ such that $x + x_r = y$ and $x_l + x = y$.

Definition 2.15. A monoid $\mathcal{M}$ is called a *group* if for every element $a \in \mathcal{M}$ there exists $-a \in \mathcal{M}$ such that $a + (-a) = 0$. The element $-a$ is called the *inverse* of a.

Definition 2.16. For a monoid $(\mathcal{M}, +, 0)$ the *algebraic preorder* ${}_{+}{\leq}$ is defined by

$$a \,{}_{+}{\leq}\, b \quad \text{if and only if} \quad \exists c \in \mathcal{M} : a + c = b.$$

In the following, we define so-called *refinement monoids* which will be crucial for many statements made in Chapter 3. Refinement monoids have been introduced independently by Dobbertin [4] and Grillet [5]. Similar properties have been discussed even earlier, for example in Tarski's work on Cardinal algebras [16].

Definition 2.17. A commutative monoid $(\mathcal{M}, +, 0)$ has the *refinement property* if for any $a_1, a_2, b_1, b_2 \in \mathcal{M}$ with $a_1 + a_2 = b_1 + b_2$ there exist elements $c_{1,1}, c_{1,2}, c_{2,1}, c_{2,2} \in \mathcal{M}$ such that $a_i = c_{i,1} + c_{i,2}$ and $b_i = c_{1,i} + c_{2,i}$ for $i \in \{1, 2\}$.

A *refinement monoid* is a commutative monoid with the refinement property.

Example 2.18. The following magmas are refinement monoids:

(i) Any commutative group $(\mathcal{G}, +, 0)$, e.g., $(\mathbb{Z}_n, \oplus_n, 0)$ or $(\mathbb{Z}, +, 0)$.

(ii) $(\mathbb{N}, +, 0)$, $(\mathbb{Q}_{\geq 0}, +, 0)$, and $(\mathbb{R}_{\geq 0}, +, 0)$.

(iii) $(\{0, 1\}, \star, 0)$ with $a \star b = \max(a, b)$.

The magmas *(i)* and *(ii)* are cancellative. Observe moreover that for *(ii)* and *(iii)*, the algebraic preorder ${}_+\!\leq$ is a total order.

The following lemma is well-known and can be proved by induction.

Lemma 2.19. *Let $(\mathcal{M}, +, 0)$ be a refinement monoid. For any $a_1, \dots, a_m \in \mathcal{M}$ and $b_1, \dots, b_n \in \mathcal{M}$ with*

$$\sum_{i \in [m]} a_i = \sum_{n \in [n]} b_i,$$

there exists a set $\{c_{i,j}\}_{i \in [m], j \in [n]} \subseteq \mathcal{M}$ such that

$$\sum_{j \in [n]} c_{i,j} = a_i \text{ for all } i \in [m], \quad \text{and} \sum_{i \in [m]} c_{i,j} = b_i \text{ for all } j \in [n].$$

We prove three basic lemmas about commutative monoids with the algebraic preorder.

Lemma 2.20. *If $(\mathcal{M}, +, 0)$ is a commutative monoid, then for any $a, a', b, b' \in \mathcal{M}$ such that $a \,{}_+\!\leq a'$ and $b \,{}_+\!\leq b'$ we have*

$$a + b \,{}_+\!\leq a' + b'.$$

Proof. Since $a \,{}_+\!\leq a'$ there exists $a'' \in \mathcal{M}$ such that $a + a'' = a'$. Analogously, there exists $b'' \in \mathcal{M}$ such that $b + b'' = b'$. Thus, by commutativity and associativity we have $(a + b) + (a'' + b'') = (a + a'') + (b + b'') = a' + b'$, which implies $a + b \,{}_+\!\leq a' + b'$. □

Lemma 2.21. *If $(\mathcal{M}, +, 0)$ is a commutative monoid such that ${}_{+}\!\leq$ is a total order, then*

$$\inf(a, c) + \inf(b, d) \;{}_{+}\!\leq \inf(a + b, c + d)$$

for any $a, b, c, d \in \mathcal{M}$.

Proof. Since ${}_{+}\!\leq$ is a total order, we have $\inf(x, y) = x$ if and only if $x \;{}_{+}\!\leq y$ and otherwise $\inf(x, y) = y$. Assume w.l.o.g. that $a + b \;{}_{+}\!\leq c + d$.

$$\inf(a, c) + \inf(b, d) \;{}_{+}\!\leq a + \inf(b, d) \;{}_{+}\!\leq a + b = \inf(a + b, c + d).$$

The first step follows from the fact that either $\inf(a, c) = a$ or $\inf(a, c) = c$. In the former case, the first step is trivial. In the latter case we have $c \;{}_{+}\!\leq a$, implying that there exists $c' \in \mathcal{M}$ such that $c + c' = a$. The second step follows analogously. □

Lemma 2.22. *If $(\mathcal{M}, +, 0)$ is a commutative monoid such that ${}_{+}\!\leq$ is a total order, then for any finite sets $A, B \subseteq \mathcal{M}$ we have*

$$\inf A + \inf B = \inf(A + B),$$

where $A + B = \{(a + b) \mid (a, b) \in A \times B\}$.

Proof. First, observe that as ${}_{+}\!\leq$ is a total order, the infimum of finite sets always exists and is actually a minimum.

As $\inf A$ is a lower bound of A, and $\inf B$ is a lower bound of B, it is easy to see that $\inf A + \inf B$ is a lower bound of $A + B$. As $\inf(A + B)$ is the *greatest* lower bound, we have $\inf A + \inf B \;{}_{+}\!\leq \inf(A + B)$.

Since the infimum is a minimum, we have $\inf A + \inf B = a^* + b^*$ for $(a^*, b^*) \in A \times B$. Thus, $\inf A + \inf B \geq_{+} \inf(A + B)$, concluding the proof. □

Chapter 3

Intersection of Finite Distributions

In this chapter, we discuss elementary properties of (finite) distributions, in particular the similarity of distributions. We use a notion of distribution which is essentially an abstract finite measure: For a commutative monoid $\mathcal{M}$, we call any function assigning a weight $\omega \in \mathcal{M}$ to every element of some set $\mathcal{A}$ a distribution. The weight of a set $A \subseteq \mathcal{A}$ is simply defined as the sum of each element's weight.

We start by assuming very little structure on the monoid $\mathcal{M}$. For many statements, however, more structure is necessary. We attempt to assume only as much structure as required for each statement we make. We note that the monoids $(\mathbb{N}, +, 0)$, $(\mathbb{Q}_{\geq 0}, +, 0)$, and $(\mathbb{R}_{\geq 0}, +, 0)$ satisfy all assumptions made in this chapter.

The main contribution of this chapter is the introduction and discussion of a new property called observation compatibility, which relates the abstract indistinguishability of two distributions $\mathbf{X}$ and $\mathbf{Y}$ with the intersection of corresponding equivalence classes $[\mathbf{X}]_{\mathcal{F}}$ and $[\mathbf{Y}]_{\mathcal{F}}$. It is then shown how observation-compatible pairs $(\mathcal{A}, \mathcal{F})$ can be constructed.

We focus on the treatment of finite distributions and do not go into the realm of convergence, (continuous) measure theory and σ-algebras.

3.1 Finite Distributions

3.1.1 Definitions and Notation

Definition 3.1 (Finite Distribution)**.** For a commutative monoid $(\mathcal{M}, +, 0)$, a *finite $\mathcal{M}$-weighted distribution over $\mathcal{A}$* (or an $\mathcal{A}$-valued distribution) is a function

$$\mathbf{X} : \mathcal{A} \to \mathcal{M}$$

with finite support.

We call $\mathbf{X}(a)$ the *weight of an element* $a \in \mathcal{A}$. Moreover, for any distribution $\mathbf{X} : \mathcal{A} \to \mathcal{M}$ we define the function $\widehat{\mathbf{X}} : \mathcal{P}(\mathcal{A}) \to \mathcal{M}$ by[1]

$$\widehat{\mathbf{X}}(A) := \sum_{a \in A} \mathbf{X}(a),$$

and we call $\widehat{\mathbf{X}}(A)$ the *weight of the set* $A \subseteq \mathcal{A}$. Finally, we define the *weight of* $\mathbf{X}$ by $|\mathbf{X}| := \widehat{\mathbf{X}}(\mathcal{A}) = \sum_{a \in \mathcal{A}} \mathbf{X}(a)$ and we say that $\mathbf{X}$ is *weight-*ω if $\omega = |\mathbf{X}|$.

Definition 3.2. For a commutative monoid $(\mathcal{M}, +, 0)$ and a set $\mathcal{A}$, let $\Gamma_{\mathcal{A}\mathcal{M}} \subseteq \mathcal{M}^{\mathcal{A}}$ be the set of all $\mathcal{A}$-valued $\mathcal{M}$-weighted distributions.

Remark 3.3. The concept of abstract (non-numerical) measures appears in a similar form in Tarski's work on Cardinal algebras [16]. In [3], applications of such abstract measures are discussed, for example for probability theory. It is shown in [7] that some abstract measures are essentially equivalent (in an isomorphism-sense) to a numerical measure.

Many statements over distributions are only true if the monoid $\mathcal{M}$ satisfies further properties, such as

- $\mathcal{M}$ is cancellative or a refinement monoid.
- $(\mathcal{M}, {}_{+}\!\leq)$ is a partially ordered set (poset), a meet-semilattice, a lattice, a distributive (semi-)lattice, or a totally ordered set.

Definition 3.4 (Transformation)**.** For a distribution $\mathbf{X} : \mathcal{A} \to \mathcal{M}$ and a partial function $f : \mathcal{A} \nrightarrow \mathcal{B}$, the *$f$-transformation of* $\mathbf{X}$ is the $\mathcal{B}$-valued distribution defined by

$$f(\mathbf{X}) := \widehat{\mathbf{X}} \circ f^{-1}.$$

Note that the A-restriction $\mathbf{X}^{\cap[A]}$ of a distribution $\mathbf{X}$ is the id-transformation of $\mathbf{X}$, where id is the identity function $a \mapsto a$ with domain $\mathcal{A} \cap A$.

Notation 3.5. We often transform a distribution $\mathbf{X}$ by prepending or appending constant values to the elements of a distribution's domain. We write, for example, $(a, \mathbf{X}, b)$ to denote the distribution

$$(a, \mathbf{X}, b) := f_{(a,\cdot,b)}(\mathbf{X}), \qquad \text{where } f_{(a,\cdot,b)}(x) := (a, x, b).$$

Definition 3.6 (Evaluation)**.** For a function-valued distribution $\mathbf{X} : \mathcal{Z}^{\subseteq \mathcal{Y}} \to \mathcal{M}$ and $y \in \mathcal{Y}$, the *y-evaluation* of $\mathbf{X}$ is the $\mathcal{Z}$-valued distribution defined as

$$\mathbf{X}^{\downarrow y} := f_y(\mathbf{X}),$$

where $f_y : \mathcal{Z}^{\subseteq \mathcal{Y}} \nrightarrow \mathcal{Z}$ is the partial function $g \mapsto g(y)$.

[1] For an infinite set A, the sum $\sum_{a \in A} \mathbf{X}(a)$ is defined as $\sum_{a \in A \cap \mathsf{supp}(\mathbf{X})} \mathbf{X}(a)$.

3.1.2 Basic Properties of Distributions

We start by proving two elementary lemmas about distributions. The first lemma shows what structure of the monoid $\mathcal{M}$ is inherited by the set of distributions $\Gamma_{\mathcal{AM}}$. The second lemma states that the joint distribution of arbitrary distributions exist, as long as they all have the same weight.

Lemma 3.7. $(\Gamma_{\mathcal{AM}}, +, 0)$ *is a commutative monoid where* $+$ *is pointwise addition and* 0 *is the constant* 0 *function. Moreover,*

(i) the pointwise algebraic preorder is the algebraic preorder ${}_{+}{\leq}$ *over* $\Gamma_{\mathcal{AM}}$.

(ii) the relation $\mathbf{X} \leq_{|\cdot|} \mathbf{Y} :\Longleftrightarrow |\mathbf{X}| \,{}_{+}{\leq}\, |\mathbf{Y}|$ *is a preorder over* $\Gamma_{\mathcal{AM}}$.

(iii) for every $\mathbf{X} \in \Gamma_{\mathcal{AM}}$ *we have* $\omega \,{}_{+}{\leq}\, |\mathbf{X}|$ *if there exist a weight-*ω *distribution* $\mathbf{X}^{\omega} \in \Gamma_{\mathcal{AM}}$ *such that* $\mathbf{X}^{\omega} \,{}_{+}{\leq}\, \mathbf{X}$. *In case* $\mathcal{M}$ *is a refinement monoid the other direction ("only if") is true as well.*

Proof. We only prove *(iii)*, as the other claims are easy to verify. Assume there exists a weight-ω distribution $\mathbf{X}^{\omega} \in \Gamma_{\mathcal{AM}}$ such that $\mathbf{X}^{\omega} \,{}_{+}{\leq}\, \mathbf{X}$. Thus there exists $\mathbf{X}' \in \Gamma_{\mathcal{AM}}$ such that $\mathbf{X}^{\omega} + \mathbf{X}' = \mathbf{X}$, implying $|\mathbf{X}| = |\mathbf{X}^{w} + \mathbf{X}'| = |\mathbf{X}^{w}| + |\mathbf{X}'| = \omega + |\mathbf{X}'| \geq_{+} \omega$.

For the other direction, assume $\omega \,{}_{+}{\leq}\, |\mathbf{X}|$, which implies that there exists $\omega' \in \mathcal{M}$ such that $\omega + \omega' = |\mathbf{X}|$. By invoking Lemma 2.19 with (ω, ω') and $(\mathbf{X}(a_1), \ldots, \mathbf{X}(a_k))$ for $\{a_1, \ldots, a_k\} = \mathsf{supp}(\mathbf{X})$ we obtain a weight-ω distribution $\mathbf{X}^{\omega} \in \Gamma_{\mathcal{AM}}$ and a distribution $\mathbf{X}' \in \Gamma_{\mathcal{AM}}$ such that $\mathbf{X}^{\omega} + \mathbf{X}' = \mathbf{X}$, implying $\mathbf{X}^{\omega} \,{}_{+}{\leq}\, \mathbf{X}$. □

Lemma 3.8. *Assume* $\mathcal{M}$ *is a refinement monoid and let* $\langle \mathbf{X}_j \rangle_{j \in \Lambda}$ *be a finite tuple of weight-*ω *distributions over* $\mathcal{A}$, *i.e.,* $\langle \mathbf{X}_j \rangle_{j \in \Lambda} \in (\Gamma_{\mathcal{AM}})^{\Lambda}$ *and* $|\mathbf{X}_j| = \omega$ *for* $j \in \Lambda$. *There exists a tuple-valued weight-*ω *distribution* $\mathbf{X} : \mathcal{A}^{\Lambda} \to \mathcal{M}$, *such that* $\mathbf{X}^{\downarrow j} = \mathbf{X}_j$ *for all* $j \in \Lambda$.

Proof (sketch). The claim follows immediately from the refinement property via Lemma 2.19 and by induction over $|\Lambda|$. □

3.1.3 Intersection of Distributions

For $\mathcal{M}$-weighted distributions such that $(\mathcal{M}, {}_{+}{\leq})$ is a meet-semilattice, we define the *intersection* of two distributions as follows.

Definition 3.9. Assume $(\mathcal{M}, {}_{+}{\leq})$ is a meet-semilattice. The *intersection* $\mathbf{X} \sqcap \mathbf{Y}$ of two distributions $\mathbf{X} : \mathcal{A} \to \mathcal{M}$ and $\mathbf{Y} : \mathcal{A} \to \mathcal{M}$ is the distribution defined as the pointwise infimum, i.e.,

$$\mathbf{X} \sqcap \mathbf{Y} := \inf(\mathbf{X}, \mathbf{Y}).$$

Moreover, we call $|\mathbf{X} \sqcap \mathbf{Y}|$ the *intersection weight* of $\mathbf{X}$ and $\mathbf{Y}$, and use it as a similarity measure between two distributions.

The following proposition is easy to verify and thus stated without proof.

Proposition 3.10. *If $(\mathcal{M}, {}_{+}\!\leq)$ is a meet-semilattice, then $(\Gamma_{\mathcal{AM}}, {}_{+}\!\leq)$ is a meet-semilattice with pointwise meet $\sqcap$.*

3.1.4 Basic Lemmas about the Intersection Weight $\sqcap$

We prove four elementary lemmas about the intersection weight of distributions. The first two lemmas (Lemmas 3.11 and 3.12) are mostly of technical interest and used two prove more involved statements.

Lemma 3.11. *Assume $(\mathcal{M}, {}_{+}\!\leq)$ is a meet-semilattice. Let $\mathbf{X}$ and $\mathbf{Y}$ be distributions with the same domain $\mathcal{A}$. For an arbitrary partition $\langle \mathcal{A}_i \rangle_{i \in I}$ of $\mathcal{A}$ we have*

$$|\mathbf{X} \sqcap \mathbf{Y}| = \sum_{i \in I} \left| \mathbf{X}^{\cap[\mathcal{A}_i]} \sqcap \mathbf{Y}^{\cap[\mathcal{A}_i]} \right|.$$

Proof. This follows immediately from Definition 3.9.

$$\begin{aligned} |\mathbf{X} \sqcap \mathbf{Y}| &= \sum_{a \in \mathcal{A}} \inf(\mathbf{X}(a), \mathbf{Y}(a)) = \sum_{i \in I} \sum_{a \in \mathcal{A}_i} \inf(\mathbf{X}(a), \mathbf{Y}(a)) \\ &= \sum_{i \in I} \sum_{a \in \mathcal{A}_i} \inf \left(\mathbf{X}^{\cap[\mathcal{A}_i]}(a), \mathbf{Y}^{\cap[\mathcal{A}_i]}(a) \right) \\ &= \sum_{i \in I} \left| \mathbf{X}^{\cap[\mathcal{A}_i]} \sqcap \mathbf{Y}^{\cap[\mathcal{A}_i]} \right|. \end{aligned}$$

□

Lemma 3.12. *Assume $\mathcal{M}$ is a refinement monoid and $(\mathcal{M}, {}_{+}\!\leq)$ is a meet-semilattice. Let $\mathbf{X}$ and $\mathbf{Y}$ be distributions with the same domain $\mathcal{A}$. For any $\omega \in \mathcal{M}$ the following two statements are equivalent:*

(i) $\omega \, {}_{+}\!\leq |\mathbf{X} \sqcap \mathbf{Y}|$

(ii) There exists a weight-ω distribution $\mathbf{E}^{\omega} : \mathcal{A} \to \mathcal{M}$ such that $\mathbf{E}^{\omega} \, {}_{+}\!\leq \mathbf{X}$ and $\mathbf{E}^{\omega} \, {}_{+}\!\leq \mathbf{Y}$.

Proof. Recall Definition 3.9. We prove both directions separately.

- $\Longrightarrow$. Assume $\omega \, {}_{+}\!\leq |\mathbf{X} \sqcap \mathbf{Y}|$. As $\mathcal{M}$ is a refinement monoid, Lemma 3.7 implies that there exists a weight-ω distribution $\mathbf{E}^{\omega}$ with domain $\mathcal{A}$ such that $\mathbf{E}^{\omega} \, {}_{+}\!\leq \mathbf{X} \sqcap \mathbf{Y}$. Observe that $\mathbf{X} \sqcap \mathbf{Y} \, {}_{+}\!\leq \mathbf{X}$ and $\mathbf{X} \sqcap \mathbf{Y} \, {}_{+}\!\leq \mathbf{Y}$, which implies the claim by transitivity.

- $\Longleftarrow$. Let $\mathbf{E}^\omega$ be as described in *(ii)*. Since $\mathbf{E}^\omega$ is a lower bound of $\{\mathbf{X}, \mathbf{Y}\}$ we have $\mathbf{E}^\omega \,_+\!\!\leq \mathbf{X} \sqcap \mathbf{Y}$ and thus $\omega \,_+\!\!\leq |\mathbf{X} \sqcap \mathbf{Y}|$.

□

The following lemma states that the intersection weight can only increase under an f-transformation.

Lemma 3.13. *Assume $(\mathcal{M}, {}_+\!\!\leq)$ is totally ordered. Let $\mathbf{X}$ and $\mathbf{Y}$ be distributions with the same domain $\mathcal{A}$. Then, for any total function $f : \mathcal{A} \to \mathcal{B}$,*

$$|\mathbf{X} \sqcap \mathbf{Y}| \,_+\!\!\leq |f(\mathbf{X}) \sqcap f(\mathbf{Y})|\,,$$

and $|\mathbf{X} \sqcap \mathbf{Y}| = |f(\mathbf{X}) \sqcap f(\mathbf{Y})|$ if f is injective.

Proof.

$$\begin{aligned}|f(\mathbf{X}) \sqcap f(\mathbf{Y})| = \left|(\widehat{\mathbf{X}} \circ f^{-1}) \sqcap (\widehat{\mathbf{Y}} \circ f^{-1})\right| &= \sum_{b \in \mathcal{B}} \inf\left(\widehat{\mathbf{X}}(f^{-1}(b)), \widehat{\mathbf{Y}}(f^{-1}(b))\right)\\ &= \sum_{b \in \mathcal{B}} \inf\left(\sum_{a \in f^{-1}(b)} \mathbf{X}(a), \sum_{a \in f^{-1}(b)} \mathbf{Y}(a)\right)\\ &\geq_+ \sum_{b \in \mathcal{B}} \sum_{a \in f^{-1}(b)} \inf\left(\mathbf{X}(a), \mathbf{Y}(a)\right)\\ &\geq_+ \sum_{a \in \mathcal{A}} \inf\left(\mathbf{X}(a), \mathbf{Y}(a)\right) = |\mathbf{X} \sqcap \mathbf{Y}|\,,\end{aligned}$$

where we have used Lemma 2.21 for the first *in*equality and the fact that f is total for the last inequality. Observe that if f is injective, the inequalities can be replaced with equalities. □

Given multiple pairs $(\mathbf{X}_j, \mathbf{Y}_j)$ of distributions, the following lemma states that one can construct joint distributions $\mathbf{X}$ and $\mathbf{Y}$ with an intersection weight at least as large as the infimum of all intersection weights $|\mathbf{X}_j \sqcap \mathbf{Y}_j|$. If $(\mathcal{M}, {}_+\!\!\leq)$ is totally ordered, it is easy to see that the intersection weight of $\mathbf{X}$ and $\mathbf{Y}$ cannot possibly be larger, thus we have actually an equality.

Lemma 3.14. *Assume $\mathcal{M}$ is a cancellative refinement monoid and $(\mathcal{M}, {}_+\!\!\leq)$ is a meet-semilattice. Let $\langle(\mathbf{X}_j, \mathbf{Y}_j)\rangle_{j \in \Lambda}$ be a finite tuple of pairs of distributions, where weight-ω $\mathbf{X}_j$ and weight-ω' $\mathbf{Y}_j$ have domain $\mathcal{A}_j$. Let $\mathcal{A} := \cup_{j \in \Lambda} \mathcal{A}_j$.*

There exists a weight-ω distribution $\mathbf{X}$ with domain $\mathcal{A}^\Lambda$ where $\mathbf{X}^{\downarrow j} = \mathbf{X}_j$ and a weight-ω' distribution $\mathbf{Y}$ with domain $\mathcal{A}^\Lambda$ where $\mathbf{Y}^{\downarrow j} = \mathbf{Y}_j$, such that

$$|\mathbf{X} \sqcap \mathbf{Y}| \geq_+ \inf_{j \in \Lambda} |\mathbf{X}_j \sqcap \mathbf{Y}_j|$$

and if $(\mathcal{M}, {}_+\!\!\leq)$ is totally ordered we have

$$|\mathbf{X} \sqcap \mathbf{Y}| = \inf_{j \in \Lambda} |\mathbf{X}_j \sqcap \mathbf{Y}_j|\,.$$

Proof. Let $\omega_{\text{inf}} := \inf_{j\in\Lambda} |\mathbf{X}_j \sqcap \mathbf{Y}_j|$. For every $j \in \Lambda$, we split the distributions $\mathbf{X}_j$ and $\mathbf{Y}_j$ according to Lemma 3.12 such that

$$\mathbf{X}_j = \mathbf{E}_j{}^{\omega_{\text{inf}}} + \mathbf{X}_j{}' \quad \text{and} \quad \mathbf{Y}_j = \mathbf{E}_j{}^{\omega_{\text{inf}}} + \mathbf{Y}_j{}'.$$

Let $\mathbf{E}^{\omega_{\text{inf}}}$ be a weight-ω_{inf} distribution with domain $\mathcal{A}^\Lambda$ such that $\mathbf{E}^{\omega_{\text{inf}}\downarrow j} = \mathbf{E}_j{}^{\omega_{\text{inf}}}$ for every $j \in \Lambda$. Such a distribution exists due to Lemma 3.8. Moreover, let $\mathbf{X}'$ and $\mathbf{Y}'$ be analogous distributions[2].

Let $\mathbf{X}$ and $\mathbf{Y}$ be distributions defined by

$$\mathbf{X} := \mathbf{E}^{\omega_{\text{inf}}} + \mathbf{X}' \text{ and } \mathbf{Y} := \mathbf{E}^{\omega_{\text{inf}}} + \mathbf{Y}'.$$

Lemma 3.12 implies $|\mathbf{X} \sqcap \mathbf{Y}| \geq_+ \omega_{\text{inf}}$, and it is easy to verify that $\mathbf{X}^{\downarrow j} = \mathbf{X}_j$ as well as $\mathbf{Y}^{\downarrow j} = \mathbf{Y}_j$ as desired.

Observe that if $(\mathcal{M}, {}_+\!\leq)$ is totally ordered, Lemma 3.13 is applicable and we have $|\mathbf{X} \sqcap \mathbf{Y}| {}_+\!\leq \left|\mathbf{X}^{\downarrow j} \sqcap \mathbf{Y}^{\downarrow j}\right| = |\mathbf{X}_j \sqcap \mathbf{Y}_j|$ for all $j \in \Lambda$. As $(\mathcal{M}, {}_+\!\leq)$ is a meet-semilattice it follows that $|\mathbf{X} \sqcap \mathbf{Y}| {}_+\!\leq \inf_{j\in\Lambda} |\mathbf{X}_j \sqcap \mathbf{Y}_j| = \omega_{\text{inf}}$, which implies $|\mathbf{X} \sqcap \mathbf{Y}| = \omega_{\text{inf}}$ by antisymmetry. □

3.2 Observing Distributions of Objects

In this section, we consider the following setting. Objects of a set $\mathcal{A}$ may be *observed* by one function f of a set $\mathcal{F}$ of total functions from $\mathcal{A}$ to some set $\mathcal{B}$. The interpretation is that for any object $a \in \mathcal{A}$, an observer may choose a function $f \in \mathcal{F}$ arbitrarily and then observes $f(a)$.

Moreover, we assume an equivalence relation $\equiv_{\mathcal{F}}$ on $\Gamma_{\mathcal{AM}}$ such that

$$\mathbf{X} \equiv_{\mathcal{F}} \mathbf{X}' \iff \forall f \in \mathcal{F} : f(\mathbf{X}) = f(\mathbf{X}'), \tag{3.1}$$

and let $[\mathbf{X}]_{\mathcal{F}} := \{\mathbf{X}' \mid \mathbf{X}' \in \Gamma_{\mathcal{AM}},\ \mathbf{X} \equiv_{\mathcal{F}} \mathbf{X}'\}$ denote the equivalence class of $\mathbf{X} \in \Gamma_{\mathcal{AM}}$.

Remark 3.15. It is not necessary to fix the set $\mathcal{F}$ first and to *define* the equivalence relation $\equiv_{\mathcal{F}}$ such that (3.1) is satisfied with respect to $\mathcal{F}$. A more elementary approach appears to be to first define the equivalence relation and to let $\mathcal{F}$ be the induced set such that (3.1) is satisfied. In general, however, both approaches are possible.

3.2.1 Observation Compatibility

Definition 3.16. The intersection weight of two sets of distributions $\mathcal{X}, \mathcal{Y} \subseteq \Gamma_{\mathcal{AM}}$ is defined as

$$|\mathcal{X} \sqcap \mathcal{Y}| := \sup_{(\mathbf{X},\mathbf{Y})\in\mathcal{X}\times\mathcal{Y}} |\mathbf{X} \sqcap \mathbf{Y}|.$$

[2] Observe that all $\mathbf{X}_j'$ (and all $\mathbf{Y}_j'$) are of the same weight, because all $\mathbf{X}_j$ are weight-ω and $\mathcal{M}$ is cancellative.

Remark 3.17. The intersection weight $|\mathcal{X} \sqcap \mathcal{Y}|$ of two sets $\mathcal{X}, \mathcal{Y} \subseteq \Gamma_{\mathcal{A}\mathcal{M}}$ may not exist. However, in most cases of our interest, not only does the supremum exist, but the supremum is actually a maximum, i.e., there *exist* $(\mathbf{X}, \mathbf{Y}) \in \mathcal{X} \times \mathcal{Y}$ such that $|\mathbf{X} \sqcap \mathbf{Y}| = |\mathcal{X} \sqcap \mathcal{Y}|$.

Proposition 3.18. *Assume* $(\mathcal{M}, {}_+\!\leq)$ *is totally ordered and* $\mathcal{F} \subseteq \mathcal{B}^{\mathcal{A}}$. *Let* $(\mathbf{X}, \mathbf{Y}) \in \Gamma_{\mathcal{A}\mathcal{M}} \times \Gamma_{\mathcal{A}\mathcal{M}}$ *be arbitrary. For any* $(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}]_{\mathcal{F}} \times [\mathbf{Y}]_{\mathcal{F}}$ *we have*

$$|\mathbf{X}' \sqcap \mathbf{Y}'| \,{}_+\!\leq \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,,$$

and thus, if $|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}|$ *exists,* $|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| \,{}_+\!\leq \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})|$.

Proof. Since the support of both $\mathbf{X}$ and $\mathbf{Y}$ is finite and ${}_+\!\leq$ is a total order, there exists $f^* \in \mathcal{F}$ such that $|f^*(\mathbf{X}) \sqcap f^*(\mathbf{Y})| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})|$.

Let $(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}]_{\mathcal{F}} \times [\mathbf{Y}]_{\mathcal{F}}$ be arbitrary. We have

$$|\mathbf{X}' \sqcap \mathbf{Y}'| \,{}_+\!\leq |f^*(\mathbf{X}') \sqcap f^*(\mathbf{Y}')| = |f^*(\mathbf{X}) \sqcap f^*(\mathbf{Y})| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,.$$

The first step follows from Lemma 3.13. The second step follows from the fact that $\mathbf{X} \equiv_{\mathcal{F}} \mathbf{X}'$ and $\mathbf{Y} \equiv_{\mathcal{F}} \mathbf{Y}'$ by assumption. □

Definition 3.19. Assume $(\mathcal{M}, {}_+\!\leq)$ is totally ordered. For a set $\mathcal{A}$ and a set $\mathcal{F} \subseteq \mathcal{B}^{\mathcal{A}}$, the pair $(\mathcal{A}, \mathcal{F})$ is called *observation-compatible* if for any $(\mathbf{X}, \mathbf{Y}) \in \Gamma_{\mathcal{A}\mathcal{M}} \times \Gamma_{\mathcal{A}\mathcal{M}}$ we have

$$|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,,$$

where the equivalence class $[\cdot]_{\mathcal{F}}$ on $\Gamma_{\mathcal{A}\mathcal{M}}$ is defined via the equivalence relation $\equiv_{\mathcal{F}}$ as above.

Moreover, we call $(\mathcal{A}, \mathcal{F})$ *existentially observation-compatible* if there *exists* a pair $(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}]_{\mathcal{F}} \times [\mathbf{Y}]_{\mathcal{F}}$ such that $|\mathbf{X}' \sqcap \mathbf{Y}'| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})|$.

Remark 3.20. It is easy to see that existential observation compatibility implies observation compatibility.

Example 3.21. Assume $(\mathcal{M}, {}_+\!\leq)$ is totally ordered. Let $\mathcal{A}$ be a non-empty set and let $\mathcal{F} \subseteq \mathcal{B}^{\mathcal{A}}$ such that there exists an injective function $f^* \in \mathcal{F}$. Then we have for any $(\mathbf{X}, \mathbf{Y}) \in \Gamma_{\mathcal{A}\mathcal{M}} \times \Gamma_{\mathcal{A}\mathcal{M}}$

$$|\mathbf{X} \sqcap \mathbf{Y}| = |f^*(\mathbf{X}) \sqcap f^*(\mathbf{Y})| \geq_+ \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,.$$

The first step follows via Lemma 3.13 since f^* is injective. Moreover, by Proposition 3.18 and antisymmetry we conclude

$$|\mathbf{X} \sqcap \mathbf{Y}| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,.$$

Hence, $(\mathcal{A}, \mathcal{F})$ is existentially observation-compatible.

Example 3.22. Let $\mathcal{A} = \{00, 01, 10\}$, $\mathcal{F} = \{\mathsf{left}, \mathsf{right}\}$, and $\mathcal{M} = \mathbb{R}_{\geq 0}$. Moreover, let $\mathbf{X}(00) = 0$ and $\mathbf{X}(01) = \mathbf{X}(10) = 1$. Finally, let $\mathbf{Y}(00) = 2$ and $\mathbf{Y}(01) = \mathbf{Y}(10) = 0$. It is easy to see that $\inf_{f\in\mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = 1$. However, we have $[\mathbf{X}]_{\mathcal{F}} = \{\mathbf{X}\}$ and $[\mathbf{Y}]_{\mathcal{F}} = \{\mathbf{Y}\}$, thus $|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| = 0$. Thus, $(\mathcal{A}, \mathcal{F})$ is not observation-compatible. Note that the pair $(\mathcal{A}', \mathcal{F})$ for the extended set $\mathcal{A}' = \mathcal{A} \cup \{11\}$ is observation-compatible (see Lemma 3.24 below).

3.2.2 Lifting Observation Compatibility

In this section, we describe natural ways to lift observation compatibility, i.e., to construct an (existentially) observation-compatible pair $(\mathcal{A}, \mathcal{F})$ from (existentially) observation-compatible pairs $(\mathcal{A}_i, \mathcal{F}_i)$.

The following lifting lemma states the simple fact that constant functions have no influence on observation compatibility.

Lemma 3.23. *Assume $(\mathcal{M}, {}_{+}\!\leq)$ is totally ordered. A pair $(\mathcal{A}, \mathcal{F})$ is (existentially) observation-compatible for a non-empty set $\mathcal{F} \subseteq \mathcal{B}^{\mathcal{A}}$ if and only if the pair $(\mathcal{A}, \mathcal{F} \cup \{c\})$ is (existentially) observation-compatible, where $c : \mathcal{A} \to \mathcal{B}$ is a constant function.*

Proof (sketch). It is easy to see that for any $f \in \mathcal{F}$ we have

$$|f(\mathbf{X}) \sqcap f(\mathbf{Y})| \,{}_{+}\!\leq |c(\mathbf{X}) \sqcap c(\mathbf{Y})| = \inf(|\mathbf{X}|, |\mathbf{Y}|).$$

Thus, $\inf_{f\in\mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = \inf_{f\in(\mathcal{F}\cup\{c\})} |f(\mathbf{X}) \sqcap f(\mathbf{Y})|$. The claim follows by observing that the equivalence relation $\equiv_{(\mathcal{F}\cup\{c\})}$ is (equal to) $\equiv_{\mathcal{F}}$. □

A natural way to lift observation compatibility given observation-compatible pairs $(\mathcal{A}_i, \mathcal{F}_i)$ is by constructing a new set of objects $\mathcal{A}$ as the cross product $\mathcal{A}_1 \times \cdots \times \mathcal{A}_k$ and a corresponding set $\mathcal{F}$ which allows to observe an arbitrary index $i \in [k]$ of a tuple with an arbitrary function $f_i \in \mathcal{F}_i$. The following lemma states this for indexed tuples.

Lemma 3.24. *Assume $\mathcal{M}$ is a cancellative refinement monoid and $(\mathcal{M}, {}_{+}\!\leq)$ is totally ordered. Let $\langle(\mathcal{A}_j, \mathcal{F}_j)\rangle_{j\in\Lambda}$ be a finite tuple of existentially observation-compatible pairs. Then the pair $(\mathcal{A}, \mathcal{F})$ is existentially observation-compatible for $\mathcal{A} = \{a \mid \mathsf{dom}(a) = \Lambda, \forall j \in \Lambda : a(j) \in \mathcal{A}_j\}$ and*

$$\mathcal{F} := \{f \mid j \in \Lambda, f_j \in \mathcal{F}_j, f(a) := f_j(a(j))\}.$$

Proof. By the definition of $\mathcal{F}$ we have

$$\inf_{f\in\mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = \inf_{j\in\Lambda} \inf_{f_j\in\mathcal{F}_j} \left| f_j(\mathbf{X}^{\downarrow j}) \sqcap f_j(\mathbf{Y}^{\downarrow j}) \right|.$$

Since $(\mathcal{A}_j, \mathcal{F}_j)$ is existentially observation-compatible, there exist for every $j \in \Lambda$ distributions $\mathbf{X}'_j \in [\mathbf{X}^{\downarrow j}]_{\mathcal{F}_j}$ and $\mathbf{Y}'_j \in [\mathbf{Y}^{\downarrow j}]_{\mathcal{F}_j}$ such that

$$\inf_{j \in \Lambda} \inf_{f_j \in \mathcal{F}_j} \left| f_j(\mathbf{X}^{\downarrow j}) \sqcap f_j(\mathbf{Y}^{\downarrow j}) \right| = \inf_{j \in \Lambda} \left| \mathbf{X}'_j \sqcap \mathbf{Y}'_j \right| .$$

Invoking Lemma 3.14 with the tuple $\langle (\mathbf{X}'_j, \mathbf{Y}'_j) \rangle_{j \in \Lambda}$, we obtain a pair of distributions $(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}] \times [\mathbf{Y}]$ such that $|\mathbf{X}' \sqcap \mathbf{Y}'| = \inf_{j \in \Lambda} \left| \mathbf{X}'_j \sqcap \mathbf{Y}'_j \right|$. Hence,

$$|\mathbf{X}' \sqcap \mathbf{Y}'| = \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| ,$$

which concludes the proof. □

One might conjecture that observation compatibility is inherited in the same fashion if we allow to project *all* indices at once, i.e., we define the set $\mathcal{F}$ by

$$\mathcal{F} := \left\{ f \mid f(a) := \langle f_j(a(j)) \rangle_{j \in \Lambda} \text{ for } f_j \in \mathcal{F}_j \right\} .$$

This conjecture is false, as the following example shows.

Example 3.25. Let $\mathcal{M} = \mathbb{R}_{\geq 0}$, $\mathcal{A}_1 = \{0, 1\}$, $\mathcal{F}_1 = \{\mathsf{id}\}$, $\mathcal{A}_2 = \{00, 01, 10, 11\}$, and $\mathcal{F}_2 = \{\mathsf{right}, \mathsf{left}\}$. $(\mathcal{A}_i, \mathcal{F}_i)$ is existentially observation-compatible for $i \in \{1, 2\}$ due to $\mathsf{id} \in \mathcal{F}_1$ being injective and Lemma 3.24. We define $\mathcal{A} := \mathcal{A}_1 \times \mathcal{A}_2$ and $\mathcal{F} := \{f \mid f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2, f(a_1, a_2) = (f_1(a_1), f_2(a_2))\}$.

Consider the distributions $\mathbf{X}$ and $\mathbf{Y}$ which have support $\{(0, 10), (1, 11)\}$ and $\{(0, 00), (1, 10)\}$, respectively, and weight 1 for every element in their support.

It is easy to verify that

$$\inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = 1.$$

Convince yourself that $[\mathbf{X}]_{\mathcal{F}} = \{\mathbf{X}\}$ and $[\mathbf{Y}]_{\mathcal{F}} = \{\mathbf{Y}\}$, thus

$$\sup_{(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}]_{\mathcal{F}} \times [\mathbf{Y}]_{\mathcal{F}}} |\mathbf{X}' \sqcap \mathbf{Y}'| = 0.$$

Hence, $(\mathcal{A}, \mathcal{F})$ is not observation-compatible.

Observe that the set $\mathcal{F}$ from Example 3.25 can be extended to include all $\mathcal{F}_1$-adaptive functions, i.e., functions mapping (a_1, a_2) to $(f_1(a_1), f_2(a_2))$, but where the choice of f_2 depends on the value of $f_1(a_1)$. If this new set is denoted by $\mathcal{F}'$, we have

$$\inf_{f \in \mathcal{F}'} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = 0.$$

Thus, $(\mathcal{A}, \mathcal{F}')$ might be observation-compatible. In particular, one might conjecture that observation compatibility is inherited recursively if we allow adaptive projections in one direction (e.g., from left to right). In general, however, this conjecture is false as well, as the following example shows.

Example 3.26. Let $\mathcal{M} = \mathbb{R}_{\geq 0}$, $\mathcal{A}_1 = \mathcal{A}_2 = \{00, 01, 10, 11\}$, and $\mathcal{F}_1 = \mathcal{F}_2 = \{\mathsf{right}, \mathsf{left}\}$. $(\mathcal{A}_i, \mathcal{F}_i)$ is existentially observation-compatible by Lemma 3.24 for $i \in \{1, 2\}$. We define $\mathcal{A} := \mathcal{A}_1 \times \mathcal{A}_2$ and

$$\mathcal{F} := \{f \mid f_1 \in \mathcal{F}_1, v_2 \in \mathcal{F}_2^{\{0,1\}}, f(a_1, a_2) := (f_1(a_1), (v_2 \circ f_1)(a_1)(a_2))\}.$$

Consider the distributions $\mathbf{X}$ and $\mathbf{Y}$ which have support $\{(00, 11), (10, 00)\}$ and $\{(00, 10), (01, 01)\}$, respectively, and weight 1 for every element in their support.

It is easy to verify that

$$\inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| = 1$$

as well as $[\mathbf{Y}]_{\mathcal{F}} = \{\mathbf{Y}\}$ and $[\mathbf{X}]_{\mathcal{F}} = \{\mathbf{X}\}$. Thus,

$$\sup_{(\mathbf{X}', \mathbf{Y}') \in [\mathbf{X}]_{\mathcal{F}} \times [\mathbf{Y}]_{\mathcal{F}}} |\mathbf{X}' \sqcap \mathbf{Y}'| = 0.$$

Hence, $(\mathcal{A}, \mathcal{F})$ is not observation-compatible.

Recall that right denotes the mapping $(a_1, a_2) \mapsto a_2$.

Notation 3.27. We define the following notation for distributions $\mathbf{X}$ over a cross product $\mathcal{A}_1 \times \mathcal{A}_2$ for any $a_1 \in \mathcal{A}_1$:

$$\mathbf{X}^{\uparrow a_1} := \mathsf{right}\left(\mathbf{X}^{\cap[(a_1, \_)]}\right).$$

Thus, $\mathbf{X}^{\uparrow a_1}$ is a distribution over $\mathcal{A}_2$.

The following lemma demonstrates another lifting method. The intuition is that the constructed objects $\mathcal{A}$ are pairs of the form (l, a_l) such that $a_l \in \mathcal{A}_l$. The functions $\mathcal{F}$ are such that $(l, f_l(a_l))$ may be observed for an arbitrary function $f_l \in \mathcal{F}_l$. A possible interpretation is that l indicates the type of a_l, and an observer first observes this type, and then (depending on the value of l) chooses a function f_l fitting to type l to observe a_l.

Lemma 3.28. *Assume $(\mathcal{M}, {}_{+}\!\leq)$ is totally ordered. Let $\langle(\mathcal{A}_l, \mathcal{F}_l)\rangle_{l \in \mathcal{L}}$ be a tuple of existentially observation-compatible pairs. We define the sets $\mathcal{A} := \{(l, a_l) \mid l \in \mathcal{L}, a_l \in \mathcal{A}_l\}$, $\mathcal{F}_{\mathcal{A}_\mathcal{L}} := \{v \in (\cup_{l \in \mathcal{L}} \mathcal{F}_l)^{\mathcal{L}}, \forall l \in \mathcal{L} : v(l) \in \mathcal{F}_l\}$, as well as*

$$\mathcal{F} := \{f \mid v \in \mathcal{F}_{\mathcal{A}_\mathcal{L}}, f(l, a_l) := (l, v(l)(a_l))\}.$$

Then, $(\mathcal{A}, \mathcal{F})$ is existentially observation-compatible.

Proof.

$$\begin{aligned}\inf_{f\in\mathcal{F}}|f(\mathbf{X})\sqcap f(\mathbf{Y})| &= \inf_{f\in\mathcal{F}}\sum_{l\in\mathcal{L}}\left|f(\mathbf{X})^{\cap[(l,\_)]}\sqcap f(\mathbf{Y})^{\cap[(l,\_)]}\right| \\ &= \inf_{v\in\mathcal{F}_{\mathcal{A}_\mathcal{L}}}\sum_{l\in\mathcal{L}}\left|\left(l,v(l)(\mathbf{X}^{\uparrow l})\right)\sqcap\left(l,v(l)(\mathbf{Y}^{\uparrow l})\right)\right| \\ &= \inf_{v\in\mathcal{F}_{\mathcal{A}_\mathcal{L}}}\sum_{l\in\mathcal{L}}\left|v(l)(\mathbf{X}^{\uparrow l})\sqcap v(l)(\mathbf{Y}^{\uparrow l})\right| \\ &= \sum_{l\in\mathcal{L}}\inf_{f_l\in\mathcal{F}_l}\left|f_l(\mathbf{X}^{\uparrow l})\sqcap f_l(\mathbf{Y}^{\uparrow l})\right|.\end{aligned}$$

The first step follows from the Lemma 3.11. The second step follows from the definition of $\mathcal{F}$. For the third step we used that all values in the support of $\left(l,v(l)(\mathbf{X}^{\uparrow l})\right)$ and $\left(l,v(l)(\mathbf{Y}^{\uparrow l})\right)$ have the form $(l,\_)$, thus removing this first index does not change the intersection weight. The last step is due to Lemma 2.22.

For every $l\in\mathcal{L}$, the pair $(\mathcal{A}_l,\mathcal{F}_l)$ is existentially observation-compatible and $\mathbf{X}^{\uparrow l}$ as well as $\mathbf{Y}^{\uparrow l}$ are distributions over $\mathcal{A}_l$. Thus, for every $l\in\mathcal{L}$ there exist $(\mathbf{X}'_l,\mathbf{Y}'_l)\in[\mathbf{X}^{\uparrow l}]_{\mathcal{F}_l}\times[\mathbf{Y}^{\uparrow l}]_{\mathcal{F}_l}$ such that

$$\sum_{l\in\mathcal{L}}\inf_{f_l\in\mathcal{F}_l}\left|f_l(\mathbf{X}^{\uparrow l})\sqcap f_l(\mathbf{Y}^{\uparrow l})\right| = \sum_{l\in\mathcal{L}}|\mathbf{X}'_l\sqcap\mathbf{Y}'_l|.$$

Finally, we define[3] $\mathbf{X}' := \bigcup_{l\in\mathcal{L}}(l,\mathbf{X}'_l)$ and $\mathbf{Y}' := \bigcup_{l\in\mathcal{L}}(l,\mathbf{Y}'_l)$. It is easy to check that $(\mathbf{X}',\mathbf{Y}')\in[\mathbf{X}]_\mathcal{F}\times[\mathbf{Y}]_\mathcal{F}$. Invoking Lemma 3.11 we obtain

$$|\mathbf{X}'\sqcap\mathbf{Y}'| = \sum_{l\in\mathcal{L}}|(l,\mathbf{X}'_l)\sqcap(l,\mathbf{Y}'_l)| = \sum_{l\in\mathcal{L}}|\mathbf{X}'_l\sqcap\mathbf{Y}'_l| = \inf_{f\in\mathcal{F}}|f(\mathbf{X})\sqcap f(\mathbf{Y})|,$$

which concludes the proof. □

Finally, we state the following conjecture, describing a more general adaptive lifting of observation compatibility.

Conjecture 3.29. *Let $(\mathcal{A}_1,\mathcal{F}_1)$ and $(\mathcal{A}_2,\mathcal{F}_2)$ be existentially observation-compatible, respectively, and $\mathcal{F}_1\subseteq\mathcal{B}_1^{\mathcal{A}_1},\mathcal{F}_2\subseteq\mathcal{B}_2^{\mathcal{A}_2}$. We define*

$$\begin{aligned}\mathcal{A} &:= \mathcal{A}_1\times\mathcal{A}_2, \quad \textit{and} \\ \mathcal{F} &:= \left\{f \mid f_1\in\mathcal{F}_1, v_2\in\mathcal{F}_2^{\mathcal{B}_1}, f(a_1,a_2) := (f_1(a_1),(v_2\circ f_1)(a_1)(a_2))\right\} \\ &\quad\cup\left\{f \mid v_1\in\mathcal{F}_1^{\mathcal{B}_2}, f_2\in\mathcal{F}_2, f(a_1,a_2) := ((v_1\circ f_2)(a_2)(a_1),f_2(a_2))\right\}.\end{aligned}$$

Then, $(\mathcal{A},\mathcal{F})$ is existentially observation-compatible.

[3] Recall that functions are defined as sets, thus taking the union of functions with disjoint domain yields another function.

3.3 Meet Anti Triangle Inequality

In this section, we briefly discuss a meet anti triangle inequality over equivalence classes $[\cdot]_{\mathcal{F}}$ of distributions. Such an inequality is motivated by the fact that it would imply an actual triangle inequality for a natural distance notion defined over sets of distributions, proving that the distance is a pseudo-metric.

For the following lemma, we write $x \wedge y$ and $x \vee y$ instead of $\inf(x, y)$ and $\sup(x, y)$ to improve readability.

Lemma 3.30. *If $(\mathcal{M}, {}_{+}\!\leq)$ is totally ordered, we have for any $a, b, c \in \mathcal{M}$:*

$$(a \wedge c) + b \geq_{+} (a \wedge b) + (b \wedge c).$$

Proof. It is easy to see that $a \wedge c \geq_{+} (a \wedge b) \wedge (b \wedge c)$ and $b \geq_{+} (a \wedge b) \vee (b \wedge c)$. Thus,

$$\begin{aligned}(a \wedge c) + b &\geq_{+} ((a \wedge b) \wedge (b \wedge c)) + ((a \wedge b) \vee (b \wedge c)) \\ &= (a \wedge b) + (b \wedge c).\end{aligned}$$

The last step follows from the fact that $(x \wedge y) + (x \vee y) = x + y$ for any $x, y \in \mathcal{M}$, as ${}_{+}\!\leq$ is a total order. □

Lemma 3.31. *Assume $(\mathcal{M}, {}_{+}\!\leq)$ is totally ordered. Let the pair $(\mathcal{A}, \mathcal{F})$ be observation-compatible. Then for arbitrary $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \Gamma_{\mathcal{AM}}$ we have*

$$|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| + |\mathbf{Z}| \geq_{+} |[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Z}]_{\mathcal{F}}| + |[\mathbf{Z}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| .$$

Proof.

$$\begin{aligned}|[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| + |\mathbf{Z}| &= \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Y})| + |\mathbf{Z}| \\ &= \inf_{f \in \mathcal{F}} (|f(\mathbf{X}) \sqcap f(\mathbf{Y})| + |f(\mathbf{Z})|) \\ &\geq_{+} \inf_{f \in \mathcal{F}} (|f(\mathbf{X}) \sqcap f(\mathbf{Z})| + |f(\mathbf{Z}) \sqcap f(\mathbf{Y})|) \\ &\geq_{+} \inf_{f \in \mathcal{F}} |f(\mathbf{X}) \sqcap f(\mathbf{Z})| + \inf_{f \in \mathcal{F}} |f(\mathbf{Z}) \sqcap f(\mathbf{Y})| \\ &= |[\mathbf{X}]_{\mathcal{F}} \sqcap [\mathbf{Z}]_{\mathcal{F}}| + |[\mathbf{Z}]_{\mathcal{F}} \sqcap [\mathbf{Y}]_{\mathcal{F}}| .\end{aligned}$$

The first *in*equality follows from Lemma 3.30. □

An interesting problem is to generalize the above statement for general sets $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, and to replace observation compatibility with a weaker assumption.

3.4 Interpreting the Intersection Weight

The following lemma provides an interpretation of the intersection weight of distributions by relating it to the maximum weight of constant tuples over all joint distributions.

Lemma 3.32. *Assume $\mathcal{M}$ is a cancellative refinement monoid and $(\mathcal{M}, {}_+{\leq})$ is a meet-semilattice. Let $\langle \mathbf{X}_j \rangle_{j\in\Lambda}$ be a finite tuple of weight-ω distributions $\mathbf{X}_j \in \Gamma_{\mathcal{A}\mathcal{M}}$. Then*[4]*,*

$$\sup_{\substack{\mathbf{X}\in\Gamma_{\mathcal{A}^\Lambda\mathcal{M}}\\ \mathbf{X}^{\downarrow j}=\mathbf{X}_j}} \widehat{\mathbf{X}}\left(\mathcal{A}^{\Lambda}_{=}\right) = \left|\sqcap_{j\in\Lambda}\mathbf{X}_j\right|,$$

and there exists $\mathbf{X}^* \in \Gamma_{\mathcal{A}^\Lambda\mathcal{M}}$ with $\mathbf{X}^{*\downarrow j} = \mathbf{X}_j$ such that $\widehat{\mathbf{X}^*}(\mathcal{A}^{\Lambda}_{=}) = \left|\sqcap_{j\in\Lambda}\mathbf{X}_j\right|$.

Proof. Let $\mathbf{X} \in \Gamma_{\mathcal{A}^\Lambda\mathcal{M}}$ with $\mathbf{X}^{\downarrow j} = \mathbf{X}_j$ be arbitrary. For any $j \in \Lambda, a \in \mathcal{A}$ we have[5]

$$\mathbf{X}(a) \,{}_+{\leq}\, \mathbf{X}^{\downarrow j}(a) = \mathbf{X}_j(a).$$

Thus, $\mathbf{X}(a) \,{}_+{\leq} (\sqcap_{j\in\Lambda}\mathbf{X}_j)(a)$, which implies $\widehat{\mathbf{X}}(\mathcal{A}^{\Lambda}_{=}) \,{}_+{\leq} \left|\sqcap_{j\in\Lambda}\mathbf{X}_j\right|$.

By Lemma 3.12 there exists a distribution $\mathbf{X}'_j$ such that $\mathbf{X}_j = (\sqcap_{j\in\Lambda}\mathbf{X}_j) + \mathbf{X}'_j$ for every $j \in \Lambda$. Let $\mathbf{X}^= \in \Gamma_{\mathcal{A}^\Lambda\mathcal{M}}$ be with $\mathsf{supp}(\mathbf{X}^=) \subseteq \mathcal{A}^{\Lambda}_{=}$ and $\mathbf{X}^=(a) := (\sqcap_{j\in\Lambda}\mathbf{X}_j)(a)$, and let $\mathbf{X}'$ be the joint distribution of $\langle \mathbf{X}'_j \rangle_{j\in\Lambda}$ that exists[6] by Lemma 3.8. We define $\mathbf{X}^* := \mathbf{X}^= + \mathbf{X}'$. It is easy to see that $\mathbf{X}^{*\downarrow j} = \mathbf{X}_j$ for every $j \in \Lambda$ and we have $\widehat{\mathbf{X}}(\mathcal{A}^{\Lambda}_{=}) \geq_+ \left|\sqcap_{j\in\Lambda}\mathbf{X}_j\right|$, which concludes the proof. □

Hence, a possible interpretation of the intersection weight in classical probability theory is the following. For every random experiment $\mathcal{E}$ with a random variable $(\mathrm{X}, \mathrm{Z}_1, \mathrm{Z}_2, \ldots)$, where $\mathrm{X} \sim \mathbf{X}$, there exists another random experiment $\mathcal{E}'$ with a random variable $(\mathrm{X}', \mathrm{Y}', \mathrm{Z}'_1, \mathrm{Z}'_2, \ldots)$, such that $\mathrm{Y}' \sim \mathbf{Y}$, $(\mathrm{X}', \mathrm{Z}'_1, \mathrm{Z}'_2, \ldots)$ is distributed exactly like $(\mathrm{X}, \mathrm{Z}_1, \mathrm{Z}_2, \ldots)$ in $\mathcal{E}$, and

$$\mathrm{Pr}^{\mathcal{E}'}(X' = Y') = \left|\mathbf{X} \sqcap \mathbf{Y}\right|.$$

Thus, whenever we are considering a random experiment with random variable X distributed according to (non-ideal) $\mathbf{X}$, we can consider there being an additional ”shadow” random variable Y that is distributed according to (ideal) $\mathbf{Y}$, such that with probability $|\mathbf{X} \sqcap \mathbf{Y}|$, the random variables X and Y are *equal.*

[4] Recall that $\mathcal{A}^{\Lambda}_{=}$ denotes the set of constant functions from Λ to $\mathcal{A}$.

[5] We let $a \in \mathcal{A}$ also denote the constant function $j \mapsto a$ (with domain Λ).

[6] Due to cancellativity, all $\mathbf{X}'_j$ have the same weight.

An analogous interpretation is often given in terms of statistical distance and the probability that two random variables are *not* equal. For probability distributions $\mathbf{X}$ and $\mathbf{Y}$ over the same set we have $\delta(\mathbf{X}, \mathbf{Y}) = 1 - |\mathbf{X} \sqcap \mathbf{Y}|$ due to Fact 2.8. Thus, Lemma 3.32 implies the following lemma which is known in the literature as *Coupling Lemma.*

Lemma 3.33 (Coupling Lemma, Lemma 3.6 of [1])**.** *Let* $\mathbf{X}, \mathbf{Y}$ *be probability distributions over the same set.*

(i) For any *random experiment* $\mathcal{E}$ *with random variables* $\mathrm{X} \sim \mathbf{X}, \mathrm{Y} \sim \mathbf{Y}$ *we have*

$$\delta(\mathbf{X}, \mathbf{Y}) \leq \Pr^{\mathcal{E}}(\mathrm{X} \neq \mathrm{Y}).$$

(ii) There exists *a (joint) distribution* $\mathcal{C}_\delta(\mathbf{X}, \mathbf{Y})$ *such that for any random experiment* $\mathcal{E}'$ *with a random variable* $(\mathrm{X}, \mathrm{Y}) \sim \mathcal{C}_\delta(\mathbf{X}, \mathbf{Y})$ *we have* $\mathrm{X} \sim \mathbf{X}, \mathrm{Y} \sim \mathbf{Y}$ *and*

$$\delta(\mathbf{X}, \mathbf{Y}) = \Pr^{\mathcal{E}'}(\mathrm{X} \neq \mathrm{Y}).$$

We note that the Coupling Lemma is also the main tool of the so-called *Coupling Method*, a powerful proof technique which is used for example to show that certain Markov chains are rapidly mixing (see [1]).

It is tempting to think that with probability $|\mathbf{X} \sqcap \mathbf{Y}|$, the random variable X is distributed like $\mathbf{Y}$. This is wrong, as the following example shows. Let $\mathbf{X} \in \Gamma_{\{H,T\}\mathbb{R}_{\geq 0}}$ be the always-heads probability distribution, i.e., $\mathbf{X}(H) = 1$ and $\mathbf{X}(T) = 0$, and let $\mathbf{Y} \in \Gamma_{\{H,T\}\mathbb{R}_{\geq 0}}$ be the uniform coin, i.e., $\mathbf{X}(H) = \mathbf{X}(T) = \frac{1}{2}$. Clearly we have $|\mathbf{X} \sqcap \mathbf{Y}| = \frac{1}{2}$. However, since a random variable $\mathrm{X} \sim \mathbf{X}$ never takes on the value T, it is distributed like $\mathbf{Y}$ with probability 0.

3.5 Conclusions

We have introduced finite distributions as well as their intersection and proved various elementary properties of them. Furthermore, we have defined observation compatibility and shown how one can construct a large class of observation-compatible pairs via two lifting methods (Lemmas 3.24 and 3.28).

We conclude by discussing some open questions and future work.

- Can some of the statements be further generalized? For example, it is imaginable that some of the discussed properties are not actually properties of the distributions over some set $\mathcal{A}$, but can be expressed equivalently as concise and natural properties of the *objects* $\mathcal{A}$ themselves.

- What role does adaptivity play for observation compatibility? The conjecture is that while adaptivity does not influence the equivalence relation $\equiv_{\mathcal{F}}$, it is crucial for the indistinguishability. In particular, unrestricted adaptivity seems to be necessary to obtain observation compatibility.

- Is it possible to define (in a meaningful and natural fashion) not only the intersection weight $|\mathcal{X} \sqcap \mathcal{Y}|$ of two sets of distributions, but actually the intersection $\mathcal{X} \sqcap \mathcal{Y}$ of sets of distributions together with a weight $|\cdot|$ of a set of distributions? Moreover, a naturally defined (pseudo-)metric on sets of distributions might allow to elevate the distance of discrete systems (see Chapter 4) to a more abstract level.

- What algebraic structure does the intersection weight $|\mathcal{X} \sqcap \mathcal{Y}|$ over pairs of distribution sets inherit from the meet $\sqcap$? In particular, it seems that properties similar to a distributive lattice would imply the meet anti triangle-inequality. Moreover, a property similar to the inclusion-exclusion principle might be technically useful to prove abstract indistinguishability bounds.

- Is it necessary to generalize the statements made in this chapter for distributions with infinite support? It seems that at least within the realm of countable infinity one could reinterpret such distributions as a sequence of finite distributions.

Chapter 4

Discrete Systems

In this chapter, we give a new representation of Maurer's theory of discrete systems [12] (an extension of the random system framework [9]). Our representation differs in two main aspects.

- First, we define discrete systems as inductive objects. Not only is this inductive view very natural, it seems necessary to take this perspective to prove many elementary statements over discrete systems. While it is in principle possible to prove said statements over non-inductive definitions, the taken *perspective* will still be inductive and thus complicate the proof unnecessarily.
- Second, we represent probabilistic discrete systems as *distributions* as opposed to random variables. This defers the discussion of probability theory and random experiments to a later point and allows making more abstract statements over systems. Note that even though we use the term *probabilistic*, the distributions are a priori not probability distributions, i.e., they do not need (unless explicitly stated) to sum up to one.

We then use the results of Chapter 3 to show that discrete systems are observation-compatible with respect to a natural set of functions $\mathcal{F}$ that captures how a system may be observed by environments. This directly implies the Distance Lemma, which states that a newly defined environment-less pseudo-metric $\widehat{\Delta}$ is equal to the classical distinguishing advantage Δ.

Finally, we introduce an environment-less way to reason about the optimal game winning probability of a certain type of discrete games.

4.1 Deterministic Discrete Systems

We start by introducing *recursive domains*, which capture the *type* of a system (and the type of environments that are compatible to said system). A recursive

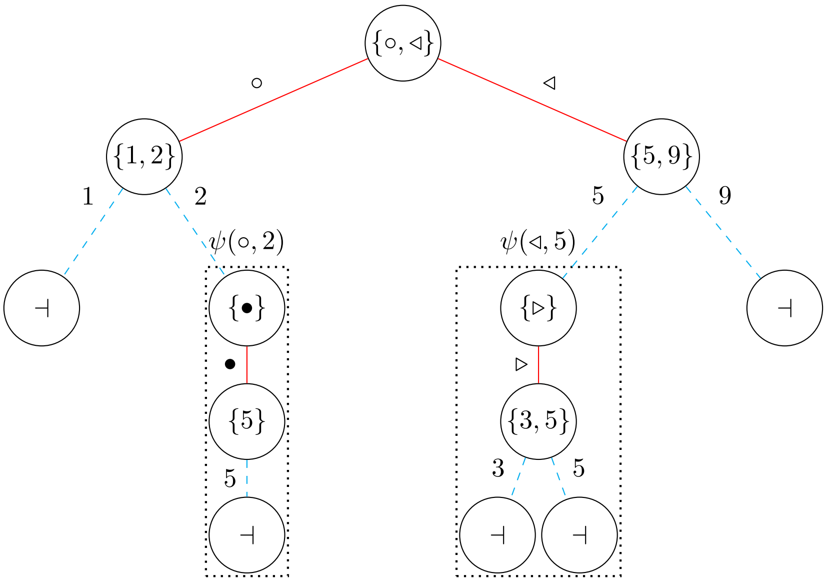

Figure 4.1: Visual representation of a finite recursive domain $\mathcal{R} = (\{\circ,\triangleleft\},\mathcal{Y},\psi)$ within input-output universe $(\mathcal{U}_\mathcal{I},\mathcal{U}_\mathcal{O}) = (\{\circ,\bullet,\triangleright,\triangleleft\},\mathbb{N})$. We have $\mathcal{Y}(\circ) = \{1,2\}$, $\mathcal{Y}(\triangleleft) = \{5,9\}$, and $\psi(\circ,1) = \psi(\triangleleft,9) = \dashv$.

domain describes the input and output alphabets at the different states of a system. Formally, we define it as a triple $(\mathcal{X},\mathcal{Y},\psi)$, where $\mathcal{X}$ is the set of allowed values for the next input and $\mathcal{Y}(x)$ is the set of possible values for the next output under the input $x \in \mathcal{X}$. For any input $x \in \mathcal{X}$ and corresponding output $y \in \mathcal{Y}(x)$, $\psi(x,y)$ is another recursive domain which describes the subsequent input and output alphabets.

Definition 4.1. For a given input-output universe $(\mathcal{U}_\mathcal{I},\mathcal{U}_\mathcal{O})$, the set $\mathcal{A}$ of *finite recursive domains* is the smallest set closed under the following rules[1]:

(i) The empty domain $\dashv$ is a finite recursive domain, i.e., $\dashv \in \mathcal{A}$.

(ii) Any triple $(\mathcal{X},\mathcal{Y},\psi)$ is a finite recursive domain, i.e., $(\mathcal{X},\mathcal{Y},\psi) \in \mathcal{A}$, if $\mathcal{X} \subseteq \mathcal{U}_\mathcal{I}$ is non-empty finite, $\mathcal{Y} \in \mathcal{P}(\mathcal{U}_\mathcal{O})^{\mathcal{X}}$ with non-empty finite $\mathcal{Y}(x)$ for every $x \in \mathcal{X}$, and $\psi \in \mathcal{A}^{\subseteq \mathcal{X}\times\mathcal{U}_\mathcal{O}}$ such that $\mathsf{dom}(\psi) = \{(x,y) \mid x \in \mathcal{X}, y \in \mathcal{Y}(x)\}$.

We call $\mathcal{R}$ *non-adaptive* if it is either the empty domain or if $\mathcal{R} = (\mathcal{X},\mathcal{Y},\psi)$ and for all $x \in \mathcal{X}$, the unary function $\psi(x,\cdot)$ is constant (i.e., is equal for all $y \in \mathcal{Y}(x)$) and maps to a non-adaptive $\mathcal{R}'$. Otherwise (if $\mathcal{R}$ is not non-adaptive) we call $\mathcal{R}$ *adaptive*.

[1] The two rules can be expressed equivalently by a single rule if we allow $\mathcal{X}$ to be empty in *(ii)* and define $\dashv := (\varnothing,\varnothing,\varnothing)$. We choose this representation to make the base case explicit.

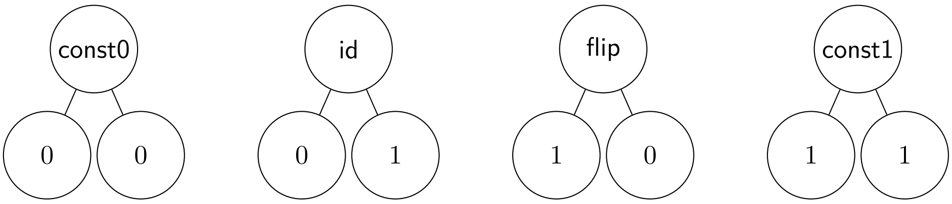

Figure 4.2: The four single-query DDS const0, id, flip, and const1.

Example 4.2. Figure 4.1 shows an example of a finite recursive domain $\mathcal{R}$. Observe that $\mathcal{R}$ is adaptive since (for example) $\psi(\circ, 1) \neq \psi(\circ, 2)$.

In the following, the input-output universe $(\mathcal{U}_{\mathcal{I}}, \mathcal{U}_{\mathcal{O}})$ is assumed to be given and is not made explicit.

For a non-empty domain $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$, we call every domain $\psi(x, y)$ for $x \in \mathcal{X}, y \in \mathcal{Y}(x)$ a *subdomain* of $\mathcal{R}$. It is easy to see that the subdomain relation is well-founded (within a given universe $(\mathcal{U}_{\mathcal{I}}, \mathcal{U}_{\mathcal{O}})$). Thus, one can prove statements of the form $\forall \mathcal{R} : \phi(\mathcal{R})$ for a predicate ϕ simply by induction based on the subdomain relation. In the base case of such a proof, we show $\phi(\dashv)$. In the induction step, we prove $\phi(\mathcal{R})$ for an arbitrary non-empty $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$ assuming $\phi(\mathcal{R}')$ is true for all subdomains $\mathcal{R}'$ of $\mathcal{R}$, i.e., all $\mathcal{R}' = \psi(x, y)$ with $x \in \mathcal{X}, y \in \mathcal{Y}(x)$.

We define a deterministic discrete system (inductively) as a function $\mathbf{s}$, such that for any input x, we have $\mathbf{s}(x) = (y, \mathbf{s}')$ for an output y and another deterministic discrete system $\mathbf{s}'$.

Definition 4.3. For a finite recursive domain $\mathcal{R}$, the set $\mathcal{S}_{\mathcal{R}}$ of (finite) *deterministic discrete $\mathcal{R}$-systems* (or $\mathcal{R}$-DDS) is the smallest set closed under the following rules:

(i) If $\mathcal{R}$ is the empty domain, i.e., $\mathcal{R} = \dashv$, then $\mathcal{S}_{\mathcal{R}} = \{\langle\,\rangle\}$, where $\langle\,\rangle$ denotes the empty (always-undefined) system.

(ii) If $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$, we have $\mathbf{s} \in \mathcal{S}_{\mathcal{R}}$ for any function $\mathbf{s} : \mathcal{X} \to \mathcal{V}_{\mathcal{R}}^{x}$, where

$$\mathcal{V}_{\mathcal{R}}^{x} := \left\{(y, \mathbf{s}') \mid y \in \mathcal{Y}(x), \mathbf{s}' \in \mathcal{S}_{\psi(x,y)}\right\}.$$

Example 4.4. Figure 4.2 depicts all four (single-query) DDS for the recursive domain $\mathcal{R} = (\{0,1\}, x \mapsto \{0,1\}, (x,y) \mapsto \dashv)$, i.e., all functions from $\{0,1\}$ to $\{0,1\} \times \{\langle\,\rangle\}$. For example, flip is defined such that

$$\mathsf{flip}(0) = (1, \langle\,\rangle) \text{ and } \mathsf{flip}(1) = (0, \langle\,\rangle).$$

A deterministic discrete environment is similar to a deterministic discrete system. However, it starts by giving an input x and then expects an output y.

We define it (inductively) as a pair $\mathbf{e} = (x, \tau)$ for an input x and a transition function τ such that $\tau(y) = \mathbf{e}'$ for another environment $\mathbf{e}'$.

Definition 4.5. The set $\mathcal{E}_{\mathcal{R}}$ of (finite) *deterministic discrete $\mathcal{R}$-environments* (or $\mathcal{R}$-DDE) is the smallest set closed under the following rules:

(i) $(\ ) \in \mathcal{E}_{\mathcal{R}}$, where $(\ )$ denotes the empty environment.

(ii) If $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$, we have $\mathbf{e} \in \mathcal{E}_{\mathcal{R}}$ for any $\mathbf{e} = (x, \tau) \in \mathcal{X} \times \mathcal{W}^x_{\mathcal{R}}$, where

$$\mathcal{W}^x_{\mathcal{R}} := \{\tau \mid \tau \in \left(\cup_{y \in \mathcal{Y}(x)} \mathcal{E}_{\psi(x,y)}\right)^{\mathcal{Y}(x)}, \forall y \in \mathcal{Y}(x) : \tau(y) \in \mathcal{E}_{\psi(x,y)}\}.$$

We call an environment $\mathbf{e}$ *non-adaptive* if either $\mathbf{e} = (\ )$ or $\mathbf{e} = (x, \tau)$ for a constant function τ which maps to a non-adaptive environment $\mathbf{e}'$.

An $\mathcal{R}$-DDE $\mathbf{e}$ can be connected to an $\mathcal{R}$-DDS $\mathbf{s}$, leading to a natural execution semantics and the notion of a *transcript.*

Definition 4.6. The *transcript* $tr(\mathbf{s}, \mathbf{e})$ between $\mathcal{R}$-DDS $\mathbf{s}$ and $\mathcal{R}$-DDE $\mathbf{e}$ is defined inductively by[2]

$$tr(\mathbf{s}, \mathbf{e}) := \begin{cases} (\ ) & \text{if } \mathbf{e} = (\ ) \\ (x, y), tr\left(\mathbf{s}', \tau(y)\right) & \text{if } \mathbf{e} = (x, \tau) \text{ and } \mathbf{s}(x) = (y, \mathbf{s}') \end{cases}$$

Moreover, we let $\mathcal{T}_{\mathcal{R}}$ denote the set of all transcripts $tr(\mathbf{s}, \mathbf{e})$ for any $\mathcal{R}$-DDS $\mathbf{s}$ and $\mathcal{R}$-DDE $\mathbf{e}$.

4.2 Probabilistic Discrete Systems

In this section, we introduce probabilistic discrete systems as *distributions* over deterministic discrete systems. Note that even though we use the term *probabilistic*, we do not demand actual probability distributions, i.e., the distributions do not need to sum up to one (unless explicitly stated).

Definition 4.7. A *probabilistic discrete $\mathcal{R}$-system* $\mathbf{S} : \mathcal{S}_{\mathcal{R}} \to \mathbb{R}_{\geq 0}$ (or $\mathcal{R}$-PDS) is an $\mathbb{R}_{\geq 0}$-weighted distribution over $\mathcal{R}$-DDS, i.e., $\mathbf{S} \in \Gamma_{\mathcal{S}_{\mathcal{R}}\mathbb{R}_{\geq 0}}$.

Analogously, a *probabilistic discrete $\mathcal{R}$-environment* $\mathbf{E} : \mathcal{E}_{\mathcal{R}} \to \mathbb{R}_{\geq 0}$ (or $\mathcal{R}$ -PDE) is an $\mathbb{R}_{\geq 0}$-weighted distribution over $\mathcal{R}$-DDE, i.e., $\mathbf{S} \in \Gamma_{\mathcal{E}_{\mathcal{R}}\mathbb{R}_{\geq 0}}$.

Recall the notation for the x-evaluation of function-valued distributions introduced in Definition 3.6. For $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$ and an $\mathcal{R}$-PDS $\mathbf{S}$, $\mathbf{S}^{\downarrow x}$ denotes the induced distribution over $\mathcal{Y} \times \cup_{y \in \mathcal{Y}(x)} \mathcal{S}_{\psi(x,y)}$ for $x \in \mathcal{X}$. Often, we combine this with Notation 3.27. For example, we write $\mathbf{S}^{\downarrow x \uparrow y}$ instead of $\left(\mathbf{S}^{\downarrow x}\right)^{\uparrow y}$, which is an $\psi(x, y)$-PDS.

[2] We omit the outer parentheses in the second case of the definition, i.e., $(x, y), tr\left(\mathbf{s}', \tau(y)\right)$ formally denotes the pair $\left((x, y), tr\left(\mathbf{s}', \tau(y)\right)\right)$.

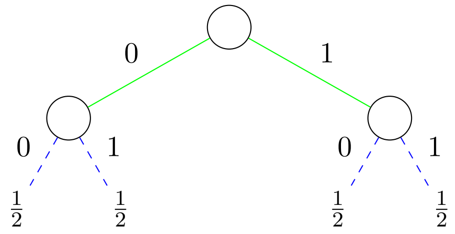

Figure 4.3: Behavior of the single-query uniform random $\{0,1\}$-function $\mathbf{V}$.

Definition 4.8. The *behavior* of an $\mathcal{R}$-PDS $\mathbf{S}$, denoted by $b(\mathbf{S})$, is defined by[3]

$$b(\mathbf{S}) := \begin{cases} |\mathbf{S}| & \text{if } \mathcal{R} = \dashv \\ \langle b(\mathbf{S}^{\downarrow x \uparrow y}) \rangle_{x \in \mathcal{X}, y \in \mathcal{Y}(x)} & \text{if } \mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi). \end{cases}$$

Interestingly, there exist different $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{S}'$ with the same behavior, i.e., $b(\mathbf{S}) = b(\mathbf{S}')$. This leads to the following equivalence relation.

Definition 4.9. $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{T}$ are equivalent, denoted $\mathbf{S} \equiv \mathbf{T}$, if they have the same behavior, i.e., if $b(\mathbf{S}) = b(\mathbf{T})$.

The equivalence class of an $\mathcal{R}$-PDS $\mathbf{S}$ is denoted by $[\mathbf{S}]$, i.e.,

$$[\mathbf{S}] := \{\mathbf{S}' \mid \mathbf{S}' \in \Gamma_{\mathcal{S}_{\mathcal{R}} \mathbb{R}_{\geq 0}}, \mathbf{S} \equiv \mathbf{S}'\}.$$

By Definition 4.8, $\mathbf{S}$ and $\mathbf{T}$ are equivalent if either $\mathcal{R} = \dashv$ and $|\mathbf{S}| = |\mathbf{T}|$ or if $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$ and for every $x \in \mathcal{X}$

$$\mathbf{S}^{\downarrow x \uparrow y} \equiv \mathbf{T}^{\downarrow x \uparrow y} \text{ for all } y \in \mathcal{Y}(x).$$

Example 4.10. Figure 4.3 depicts the behavior of the single-query uniform random $\{0,1\}$-function $\mathbf{V}$ defined by

$$\mathbf{V} := \left\{ \left(\mathsf{const0}, \frac{1}{4}\right), \left(\mathsf{const1}, \frac{1}{4}\right), \left(\mathsf{id}, \frac{1}{4}\right), \left(\mathsf{flip}, \frac{1}{4}\right) \right\}.$$

Consider the PDS $\mathbf{V}_\alpha$ for $\alpha \in [0, \frac{1}{2}]$ defined by

$$\mathbf{V}_\alpha := \left\{ (\mathsf{const0}, \alpha), (\mathsf{const1}, \alpha), \left(\mathsf{id}, \frac{1}{2} - \alpha\right), \left(\mathsf{flip}, \frac{1}{2} - \alpha\right) \right\}.$$

It is easy to verify that $\{\mathbf{V}_\alpha \mid \alpha \in [0, \frac{1}{2}]\} = [\mathbf{V}]$.

[3] Observe that in the first case ($\mathcal{R} = \dashv$) we have $|\mathbf{S}| = \mathbf{S}(\langle\,\rangle)$.

The following proposition can be easily proved by induction over $\mathcal{R}$.

Proposition 4.11. *For any two $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{T}$, $\mathbf{S} \equiv \mathbf{T}$ implies $|\mathbf{S}| = |\mathbf{T}|$ and $\mathsf{left}\left(\mathbf{S}^{\downarrow x}\right) = \mathsf{left}\left(\mathbf{T}^{\downarrow x}\right)$.*

The intuition of Definitions 4.8 and 4.9 is that two PDS $\mathbf{S}$ and $\mathbf{T}$ are equivalent if and only if their observable behavior is identical. The following lemma makes this intuition formally precise.

Lemma 4.12. *For any $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{S}'$ we have*

$$\mathbf{S} \equiv \mathbf{S}' \quad \Longleftrightarrow \quad tr(\mathbf{S}, \mathbf{e}) = tr(\mathbf{S}', \mathbf{e}) \text{ for all } \mathcal{R}\text{-DDE } \mathbf{e},$$

and if $\mathcal{R}$ is a non-adaptive domain,

$$\mathbf{S} \equiv \mathbf{S}' \quad \Longleftrightarrow \quad tr(\mathbf{S}, \mathbf{e}) = tr(\mathbf{S}', \mathbf{e}) \text{ for all non-adaptive } \mathcal{R}\text{-DDE } \mathbf{e}.$$

Proof. We prove the first equivalence by induction over $\mathcal{R}$. If $\mathcal{R}$ is the empty domain the statement follows trivially. Otherwise let $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$. We prove the two directions of the equivalence separately.

- $\Longrightarrow$. Assume $\mathbf{S} \equiv \mathbf{S}'$ and let $\mathbf{e} \in \mathcal{E}_{\mathcal{R}}$ be arbitrary. If $\mathbf{e} = (\,)$, we have trivially $tr(\mathbf{S}, \mathbf{e}) = tr(\mathbf{S}', \mathbf{e})$ since $|\mathbf{S}| = |\mathbf{S}'|$ by Proposition 4.11. Thus, assume $\mathbf{e} = (x, \tau)$ for $x \in \mathcal{X}$. For any $y \in \mathcal{Y}(x)$ we have

  $$\begin{aligned} tr(\mathbf{S}, \mathbf{e})^{\cap[(x,y)\ldots]} &= (x, y), tr(\mathbf{S}^{\downarrow x \uparrow y}, \tau(y)) \\ &\overset{\text{(I.H.)}}{=} (x, y), tr(\mathbf{S}'^{\downarrow x \uparrow y}, \tau(y)) \\ &= tr(\mathbf{S}', \mathbf{e})^{\cap[(x,y)\ldots]}. \end{aligned}$$

  In the second step we have invoked the induction hypothesis, using the fact that $\mathbf{S}^{\downarrow x \uparrow y} \equiv \mathbf{S}'^{\downarrow x \uparrow y}$ by Definition 4.9. Hence, $tr(\mathbf{S}, \mathbf{e}) = tr(\mathbf{S}', \mathbf{e})$, concluding this first part of the proof.

- $\Longleftarrow$. Assume $tr(\mathbf{S}, \mathbf{e}) = tr(\mathbf{S}', \mathbf{e})$ for all $\mathcal{R}$-DDE $\mathbf{e}$. As $tr(\cdot, \mathbf{e})$ is a total function, we have $|\mathbf{S}| = |\mathbf{S}'|$. Let $x \in \mathcal{X}$, $y \in \mathcal{Y}(x)$ be arbitrary. For arbitrary $\mathbf{e}' \in \mathcal{E}_{\psi(x,y)}$, let the $\mathcal{R}$-DDE $\mathbf{e} = (x, \tau)$ be such that $\tau(y) = \mathbf{e}'$ (and $\tau(y') \in \mathcal{E}_{\psi(x,y')}$ arbitrary for $y' \neq y$). Then,

  $$\begin{aligned} (x, y), tr(\mathbf{S}^{\downarrow x \uparrow y}, \mathbf{e}') &= tr(\mathbf{S}, \mathbf{e})^{\cap[(x,y)\ldots]} \\ &= tr(\mathbf{S}', \mathbf{e})^{\cap[(x,y)\ldots]} = (x, y), tr(\mathbf{S}'^{\downarrow x \uparrow y}, \mathbf{e}'). \end{aligned}$$

  This implies that $tr(\mathbf{S}^{\downarrow x \uparrow y}, \mathbf{e}') = tr(\mathbf{S}'^{\downarrow x \uparrow y}, \mathbf{e}')$ for all $\psi(x, y)$-DDE $\mathbf{e}'$. By induction hypothesis we have thus $\mathbf{S}^{\downarrow x \uparrow y} \equiv \mathbf{S}'^{\downarrow x \uparrow y}$, concluding the second part of the proof.

Finally, observe that if we restrict the set of environments to those which are non-adaptive, essentially the same proof is valid. Only one thing changes: In the second part of the proof ($\Longleftarrow$), we let the $\mathcal{R}$-DDE $\mathbf{e} = (x, \tau)$ be such that $\tau(y') = \mathbf{e}'$ for *all* $y' \in \mathcal{Y}(x)$. This is a valid (non-adaptive) $\mathcal{R}$-DDE if the domain $\mathcal{R}$ is non-adaptive. □

Remark 4.13. It is well-known that there exist (pairs of) systems for which the optimal adaptive distinguishing advantage is strictly larger than the optimal non-adaptive one. Lemma 4.12 implies (assuming a non-adaptive domain) that if the optimal non-adaptive distinguishing advantage is zero, then the optimal adaptive distinguishing advantage is zero as well.

A sketch of a similar statement can be found in a footnote of [6]:

> ”Using complexity leveraging, we can transform any adaptive distinguisher into a non-adaptive one with an exponential loss in the distinguishing advantage. If the optimal non-adaptive distinguishing advantage is 0 as is the case for two identical distributions, then the optimal adaptive distinguishing advantage must also be 0.”

Note that an argument via Lemma 4.12 is more elementary and more concise.

4.3 Distinguishers and Distinguishing Advantage

Many cryptographic security definitions are based on the indistinguishablity of a (real) system and an ideal system, leading to the notion of a *distinguisher* and the *distinguishing advantage*. A distinguisher is, roughly speaking, an environment that can output an additional bit $B \in \{0, 1\}$ representing the distinguisher's guess. This section defines these concepts formally.

Definition 4.14. A *deterministic discrete $\mathcal{R}$-distinguisher* $\mathbf{d}$ (or $\mathcal{R}$-DDD) is an $\mathcal{R}$-DDE $\mathbf{e}$ together with a partial function $b : \mathcal{T}_{\mathcal{R}} \nrightarrow \{0, 1\}$, i.e., it is a pair

$$(\mathbf{e}, b) \in \mathcal{E}_{\mathcal{R}} \times \{0, 1\}^{\subseteq \mathcal{T}_{\mathcal{R}}}.$$

We let $\mathcal{D}_{\mathcal{R}}$ denote the set of deterministic discrete $\mathcal{R}$-distinguishers.

A *probabilistic discrete $\mathcal{R}$-distinguisher* $\mathbf{D} : \mathcal{E}_{\mathcal{R}} \times \{0, 1\}^{\subseteq \mathcal{T}_{\mathcal{R}}} \to \mathbb{R}_{\geq 0}$ (or $\mathcal{R}$ -PDD) is an $\mathbb{R}_{\geq 0}$-weighted distribution over $\mathcal{E}_{\mathcal{R}} \times \{0, 1\}^{\subseteq \mathcal{T}_{\mathcal{R}}}$.

Definition 4.15. Given two weight-1 $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{T}$ and a weight-1 $\mathcal{R}$-PDD $\mathbf{D}$, the *distinguishing advantage* $\Delta^{\mathbf{D}}(\mathbf{S}, \mathbf{T})$ is defined as

$$\Delta^{\mathbf{D}}(\mathbf{S}, \mathbf{T}) := \Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T}, \mathrm{E})) = 1) - \Pr^{\mathbf{DS}}(\mathrm{B}(tr(\mathrm{S}, \mathrm{E})) = 1),$$

where $\mathbf{DT}$ denotes the random experiment of choosing $\mathrm{D} = (\mathrm{E}, \mathrm{B})$ according to $\mathbf{D}$ and, independently, S according to $\mathbf{S}$. The experiment $\mathbf{DS}$ is defined analogously.

Moreover, we define the *optimal distinguishing advantage* $\Delta(\mathbf{S},\mathbf{T})$ by

$$\Delta(\mathbf{S},\mathbf{T}) := \sup_{\mathbf{D}} \Delta^{\mathbf{D}}(\mathbf{S},\mathbf{T}),$$

where the supremum is over all weight-1 $\mathcal{R}$-PDD.

The following lemma is well-known and often used in the cryptographic literature.

Lemma 4.16. *For any two weight-*1 *$\mathcal{R}$-PDS* $\mathbf{S}$ *and* $\mathbf{T}$ *we have*

(i) $\Delta(\mathbf{S},\mathbf{T}) = \sup_{\mathbf{d}\in\mathcal{D}_{\mathcal{R}}} \Delta^{\mathbf{d}}(\mathbf{S},\mathbf{T}) = \sup_{\mathbf{e}\in\mathcal{E}_{\mathcal{R}}} \delta(tr(\mathbf{S},\mathbf{e}), tr(\mathbf{T},\mathbf{e}))$.

(ii) $\Delta(\mathbf{S},\mathbf{T}) = 1 - \inf_{\mathbf{e}\in\mathcal{E}_{\mathcal{R}}} |tr(\mathbf{S},\mathbf{e}) \sqcap tr(\mathbf{T},\mathbf{e})|$.

Proof. *(i)* The direction $\Delta(\mathbf{S},\mathbf{T}) \geq \sup_{\mathbf{d}\in\mathcal{D}_{\mathcal{R}}} \Delta^{\mathbf{d}}(\mathbf{S},\mathbf{T})$ is obvious. The other direction follows since we have for any probabilistic discrete $\mathcal{R}$-distinguisher $\mathbf{D}$ due to independence

$$\Delta^{\mathbf{D}}(\mathbf{S},\mathbf{T}) = \sum_{\mathbf{d}\in\mathsf{supp}(\mathbf{D})} \mathbf{D}(\mathbf{d}) \cdot \Delta^{\mathbf{d}}(\mathbf{S},\mathbf{T}) \leq \sup_{\mathbf{d}\in\mathsf{supp}(\mathbf{D})} \Delta^{\mathbf{d}}(\mathbf{S},\mathbf{T}).$$

Moreover, it is straightforward to verify that

$$\begin{aligned}
&\sup_{\mathbf{d}=(\mathbf{e},b)} \Delta^{\mathbf{d}}(\mathbf{S},\mathbf{T})\\
&= \sup_{\mathbf{d}=(\mathbf{e},b)} \left(\Pr^{\mathbf{T}}(b(tr(\mathbf{T},\mathbf{e}))=1) - \Pr^{\mathbf{S}}(b(tr(\mathbf{S},\mathbf{e}))=1)\right)\\
&= \sup_{\mathbf{d}=(\mathbf{e},b)} \sum_{\substack{m\in\mathcal{T}_{\mathcal{R}}\\ b(m)=1}} \left(\Pr^{\mathbf{T}}(tr(\mathbf{T},\mathbf{e})=m) - \Pr^{\mathbf{S}}(tr(\mathbf{S},\mathbf{e})=m)\right)\\
&= \sup_{\mathbf{e}\in\mathcal{E}_{\mathcal{R}}} \sum_{m\in\mathcal{T}_{\mathcal{R}}} \max\left(0, \Pr^{\mathbf{T}}(tr(\mathbf{T},\mathbf{e})=m) - \Pr^{\mathbf{S}}(tr(\mathbf{S},\mathbf{e})=m)\right)\\
&= \sup_{\mathbf{e}\in\mathcal{E}_{\mathcal{R}}} \delta(tr(\mathbf{S},\mathbf{e}), tr(\mathbf{T},\mathbf{e})),
\end{aligned}$$

where the last step is due to Fact 2.8.

(ii) The claim follows from *(i)* and Fact 2.8 (for weight-1 distributions $\mathbf{X}$ and $\mathbf{Y}$ we have $\delta(\mathbf{X},\mathbf{Y}) = 1 - |\mathbf{X} \sqcap \mathbf{Y}|$).

□

4.4 The Pseudo-Metric $\widehat{\Delta}$ and the Distance Lemma

In this section, we define a new pseudo-metric $\widehat{\Delta}$ on probabilistic discrete systems. Not only is this pseudo-metric natural and elementary, it also is more minimal than the classical distinguishing advantage Δ, since it is environmentless, i.e., it does not depend on a distinguisher or environment interacting with the systems.

We then prove the *Distance Lemma*, which states that for any two PDS $\mathbf{S}$ and $\mathbf{T}$ we have $\widehat{\Delta}(\mathbf{S},\mathbf{T}) = \Delta(\mathbf{S},\mathbf{T})$.

Definition 4.17. The distance between weight-1 $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{T}$ is defined by

$$\widehat{\Delta}(\mathbf{S},\mathbf{T}) := 1 - |[\mathbf{S}] \sqcap [\mathbf{T}]|\,.$$

Note that by Fact 2.8, the distance can be equivalently expressed as

$$\widehat{\Delta}(\mathbf{S},\mathbf{T}) = \inf_{(\mathbf{S}',\mathbf{T}') \in [\mathbf{S}] \times [\mathbf{T}]} \delta(\mathbf{S}',\mathbf{T}').$$

A fundamental consequence of the definition of $\widehat{\Delta}$ is that the probability-theoretic interpretation of the intersection weight discussed in Section 3.4 can now be used to reason naturally about the distance of systems: We have $\widehat{\Delta}(\mathbf{S},\mathbf{T}) \le \epsilon$ if and only if we can think of $\mathbf{S}$ and $\mathbf{T}$ being *equal* with probability at least $1-\epsilon$.

Observe that at this point, it is not obvious that $\widehat{\Delta}$ is a pseudo-metric as claimed. The Distance Lemma connects the new pseudo-metric $\widehat{\Delta}$ to the classical distinguishing advantage Δ, essentially stating that the two pseudo-metrics are equal.

Lemma 4.18 (Distance Lemma)**.** *For all weight-*1 *$\mathcal{R}$-PDS* $\mathbf{S}$ *and* $\mathbf{T}$ *we have*

$$\widehat{\Delta}(\mathbf{S},\mathbf{T}) = \Delta(\mathbf{S},\mathbf{T}),$$

and there exist $(\mathbf{S}',\mathbf{T}') \in [\mathbf{S}] \times [\mathbf{T}]$ *and such that* $\widehat{\Delta}(\mathbf{S},\mathbf{T}) = 1 - |\mathbf{S}' \sqcap \mathbf{T}'|$.

The direction $\widehat{\Delta}(\mathbf{S},\mathbf{T}) \ge \Delta(\mathbf{S},\mathbf{T})$ is not very surprising. It captures the intuition of the distinguishing advantage $\Delta(\mathbf{S},\mathbf{T})$ being upper-bounded by the probability that $\mathbf{S}$ and $\mathbf{T}$ are not equal. The opposite direction, however, is far from obvious.

The significance of Lemma 4.18 is at least twofold.

- First, it gives an intuitive and elementary justification of the classical distinguishing advantage Δ, as the natural probability-theoretic interpretation of $\widehat{\Delta}$ is inherited.
- Second, it allows to understand and prove the closeness of two systems with respect to $\widehat{\Delta}$ in an elegant and elementary manner, and then to obtain an indistinguishability result in the classical understanding without any additional effort (see also Chapter 5).

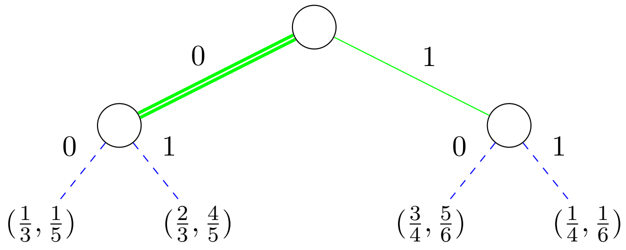
0
1
0
1
0
1
$(\frac{1}{3}, \frac{1}{5})$
$(\frac{2}{3}, \frac{4}{5})$
$(\frac{3}{4}, \frac{5}{6})$
$(\frac{1}{4}, \frac{1}{6})$

Figure 4.4: Behavior of the weight-1 PDS $\mathbf{S}$ (left) and $\mathbf{T}$ (right). On input 0, system $\mathbf{S}$ (system $\mathbf{T}$) outputs 1 with probability $\frac{2}{3}$ (with probability $\frac{4}{5}$).

4.4.1 Two Examples

Before proving the Distance Lemma, we present two examples.

Example 4.19. Figure 4.4 shows the behavior of two PDS $\mathbf{S}$ and $\mathbf{T}$ which answer a single query. The optimal distinguishing advantage is $\frac{2}{15}$, and an optimal distinguisher will always input 0. Consider the PDS $\mathbf{S}'$ defined by

$$\mathbf{S}' := \left\{ \left(\mathsf{const0}, \frac{13}{60}\right), \left(\mathsf{id}, \frac{7}{60}\right), \left(\mathsf{flip}, \frac{8}{15}\right), \left(\mathsf{const1}, \frac{2}{15}\right) \right\},$$

and the PDS $\mathbf{T}'$ defined by

$$\mathbf{T}' := \left\{ \left(\mathsf{const0}, \frac{1}{6}\right), \left(\mathsf{id}, \frac{1}{30}\right), \left(\mathsf{flip}, \frac{2}{3}\right), \left(\mathsf{const1}, \frac{2}{15}\right) \right\}.$$

It is easy to verify that $\mathbf{S}' \in [\mathbf{S}]$ and $\mathbf{T}' \in [\mathbf{T}]$. Moreover, we have $1 - |\mathbf{S}' \sqcap \mathbf{T}'| = \frac{2}{15} = \Delta(\mathbf{S}, \mathbf{T})$.

Example 4.20. Figure 4.5 shows the behavior of two PDS $\mathbf{S}$ and $\mathbf{T}$ which answer up to two queries. The optimal distinguishing advantage is $\frac{59}{99}$, and an optimal distinguisher will always input 0 as first query, and $x_2 = y_1$ as second query, where y_1 is the output from the first query. For the DDS $\mathbf{q_0}, \mathbf{q_1}, \mathbf{q_2}$ and $\mathbf{q_3}$ depicted in Figure 4.6, consider the PDS $\mathbf{S}'$ defined by

$$\mathbf{S}' := \left\{ \left(\mathbf{q_0}, \frac{2}{3}\right), \left(\mathbf{q_1}, \frac{1}{9}\right), \left(\mathbf{q_2}, \frac{1}{9}\right), \left(\mathbf{q_3}, \frac{1}{9}\right) \right\},$$

and the PDS $\mathbf{T}'$ defined by

$$\mathbf{T}' := \left\{ \left(\mathbf{q_0}, \frac{1}{11}\right), \left(\mathbf{q_1}, \frac{7}{11}\right), \left(\mathbf{q_2}, \frac{2}{11}\right), \left(\mathbf{q_3}, \frac{1}{11}\right) \right\}.$$

It is easy to verify that $\mathbf{S}' \in [\mathbf{S}]$ and $\mathbf{T}' \in [\mathbf{T}]$. Moreover, we have $1 - |\mathbf{S}' \sqcap \mathbf{T}'| = \frac{59}{99} = \Delta(\mathbf{S}, \mathbf{T})$.

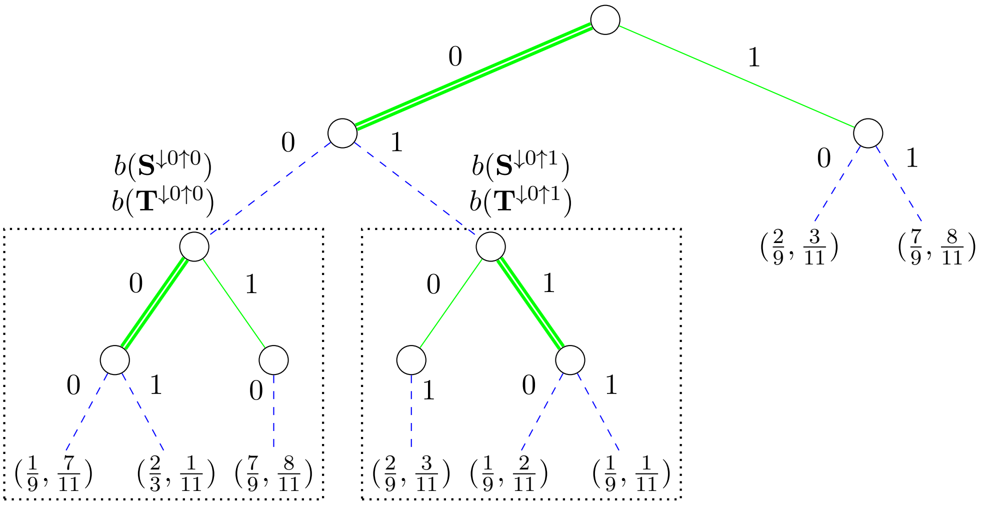
0
1
0
1
$b(\mathbf{S}^{\downarrow 0 \uparrow 0})$
$b(\mathbf{T}^{\downarrow 0 \uparrow 0})$
$b(\mathbf{S}^{\downarrow 0 \uparrow 1})$
$b(\mathbf{T}^{\downarrow 0 \uparrow 1})$
0
1
$(\frac{2}{9}, \frac{3}{11})$
$(\frac{7}{9}, \frac{8}{11})$
0
1
0
1
0
1
0
1
0
1
$(\frac{1}{9}, \frac{7}{11})$
$(\frac{2}{3}, \frac{1}{11})$
$(\frac{7}{9}, \frac{8}{11})$
$(\frac{2}{9}, \frac{3}{11})$
$(\frac{1}{9}, \frac{2}{11})$
$(\frac{1}{9}, \frac{1}{11})$

Figure 4.5: Behavior of the weight-1 PDS $\mathbf{S}$ (left) and $\mathbf{T}$ (right). On first input $x_1 = 0$, system $\mathbf{S}$ (system $\mathbf{T}$) outputs $y_1 = 1$ with probability $\frac{2}{9}$ (with probability $\frac{3}{11}$).

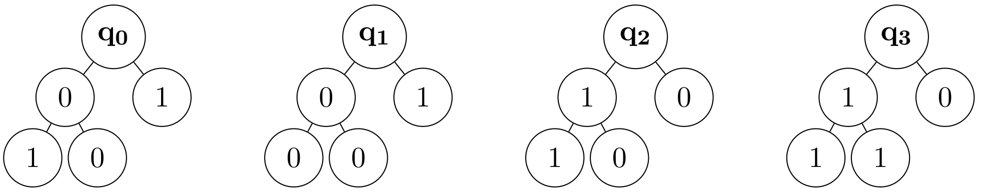

Figure 4.6: Four $\mathcal{R}$-DDS $\mathbf{q_0}, \mathbf{q_1}, \mathbf{q_2}$, and $\mathbf{q_3}$ for the recursive domain $\mathcal{R} = (\{0,1\}, x \mapsto \{0,1\}, (x,y) \mapsto \mathcal{R}_{x,y})$, where $\mathcal{R}_{0,y} = (\{0,1\}, x \mapsto \{0,1\}, (x,y) \mapsto \dashv)$ and $\mathcal{R}_{1,y} = \dashv$. On input $x_1 = 0$ and $x_2 = 1$ the DDS $\mathbf{q_3}$ outputs $y_1 = 1$ and $y_2 = 1$.

4.4.2 Proof of the Distance Lemma

Recall that the Distance Lemma states that for all weight-1 $\mathcal{R}$-PDS $\mathbf{S}$ and $\mathbf{T}$ we have

$$\widehat{\Delta}(\mathbf{S}, \mathbf{T}) = \Delta(\mathbf{S}, \mathbf{T}),$$

and there exist $(\mathbf{S}', \mathbf{T}') \in [\mathbf{S}] \times [\mathbf{T}]$ and such that $\widehat{\Delta}(\mathbf{S}, \mathbf{T}) = 1 - |\mathbf{S}' \sqcap \mathbf{T}'|$.

Proof of Lemma 4.18 (Distance Lemma). Due to the definition of $\widehat{\Delta}$ (Definition 4.17) and Lemma 4.16 *(ii)* we have:

$$\widehat{\Delta}(\mathbf{S}, \mathbf{T}) = \Delta(\mathbf{S}, \mathbf{T}) \iff |[\mathbf{S}] \sqcap [\mathbf{T}]| = \inf_{\mathbf{e} \in \mathcal{E}_{\mathcal{R}}} |tr(\mathbf{S}, \mathbf{e}) \sqcap tr(\mathbf{T}, \mathbf{e})| \,.$$

Let $\mathcal{F}_{\mathcal{R}} := \{tr(\cdot, \mathbf{e}) \mid \mathbf{e} \in \mathcal{E}_{\mathcal{R}}\}$. Since by Lemma 4.12 the equivalence relation $\equiv_{\mathcal{F}_{\mathcal{R}}}$ as defined in Chapter 3 (see Section 3.2) is equal to $\equiv$ from Definition 4.9, the lemma can thus be expressed equivalently as

$$\text{The pair } (\mathcal{S}_{\mathcal{R}}, \mathcal{F}_{\mathcal{R}}) \text{ is existentially observation-compatible.} \tag{4.1}$$

We prove (4.1) by induction over $\mathcal{R}$. If $\mathcal{R}$ is the empty domain, the statement is obvious. Thus, let $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$ be a non-empty domain. Recall Definitions 4.3 and 4.5, and in particular the definitions of $\mathcal{V}^x_{\mathcal{R}}$ and $\mathcal{W}^x_{\mathcal{R}}$:

$$\begin{aligned}\mathcal{V}^x_{\mathcal{R}} &= \left\{(y, \mathbf{s}') \mid y \in \mathcal{Y}(x), \mathbf{s}' \in \mathcal{S}_{\psi(x,y)}\right\}, \text{ and}\\ \mathcal{W}^x_{\mathcal{R}} &= \left\{\tau \mid \tau : \left(\cup_{y\in\mathcal{Y}(x)}\mathcal{E}_{\psi(x,y)}\right)^{\mathcal{Y}(x)}, \forall y \in \mathcal{Y}(x) : \tau(y) \in \mathcal{E}_{\psi(x,y)}\right\}.\end{aligned}$$

Consider the set $\mathcal{F}^x_{\mathcal{R}} \subseteq \mathcal{T}_{\mathcal{R}}{}^{\mathcal{V}^x_{\mathcal{R}}}$ defined by

$$\begin{aligned}\mathcal{F}^x_{\mathcal{R}} &:= \{f \mid \tau \in \mathcal{W}^x_{\mathcal{R}}, f(y, \mathbf{s}') := (x, y), tr(\mathbf{s}', \tau(y))\}\\ &= \{f \mid v : \left(\cup_{y\in\mathcal{Y}(x)}\mathcal{F}_{\psi(x,y)}\right)^{\mathcal{Y}(x)}, \forall y \in \mathcal{Y}(x) : v(y) \in \mathcal{F}_{\psi(x,y)},\\ &\qquad f(y, \mathbf{s}') := (x, y), v(y)(\mathbf{s}')\}.\end{aligned}$$

For every $x \in \mathcal{X}$ we instantiate Lemma 3.28 with $\mathcal{L}_x = \mathcal{Y}(x)$ and the tuple $\langle(\mathcal{S}_{\psi(x,y)}, \mathcal{F}_{\psi(x,y)})\rangle_{y\in\mathcal{L}_x}$. As $(\mathcal{S}_{\psi(x,y)}, \mathcal{F}_{\psi(x,y)})$ is existentially observation-compatible by induction hypothesis, the lemma implies that $(\mathcal{V}^x_{\mathcal{R}}, \mathcal{F}^x_{\mathcal{R}})$ is existentially observation-compatible for all $x \in \mathcal{X}$.

Invoking Lemma 3.24 on the tuple $\langle(\mathcal{V}^x_{\mathcal{R}}, \mathcal{F}^x_{\mathcal{R}})\rangle_{x\in\mathcal{X}}$ allows us to conclude that $(\mathcal{S}_{\mathcal{R}}, \mathcal{F}_{\mathcal{R}} - \{tr(\cdot, (\,))\})$ is existentially observation-compatible.

Finally, observe that $tr(\cdot, (\,))$ is a constant function, thus $(\mathcal{S}_{\mathcal{R}}, \mathcal{F}_{\mathcal{R}})$ is existentially observation-compatible by Lemma 3.23.

□

4.5 Winning Probability of Discrete Games

Finally, we consider a simplistic type of games which are discrete systems as discussed before, but where the final output is a bit $win \in \{0, 1\}$ which denotes whether the game was won or not.

Definition 4.21. A *finite recursive game domain* is a non-empty finite recursive domain $\mathcal{R} = (\mathcal{X}, \mathcal{Y}, \psi)$ such that either

- $\mathcal{X} = \{\diamond\}$, $\mathcal{Y}(\diamond) = \{0, 1\}$, and $\psi(x, y) = \dashv$ for all $x \in \mathcal{X}, y \in \mathcal{Y}(x)$, or
- $\diamond \notin \mathcal{X}$ and $\psi(x, y)$ is a finite recursive game domain for all $x \in \mathcal{X}, y \in \mathcal{Y}(x)$.

Recall that $\mathcal{T}_{\mathcal{R}}$ denotes the set of all $\mathcal{R}$-transcripts. We let $\mathcal{T}^w_{\mathcal{R}} \subseteq \mathcal{T}_{\mathcal{R}}$ denote the set of all *winning* $\mathcal{R}$-transcripts. Formally,

- $((\diamond, 1), (\,)) \in \mathcal{T}^w_{\mathcal{R}}$ if $\mathcal{X} = \{\diamond\}$, $\mathcal{Y}(\diamond) = \{0, 1\}$ and

- $((x,y),t) \in \mathcal{T}^w_{\mathcal{R}}$ if $t \in \mathcal{T}^w_{\psi(x,y)}$ for $x \in \mathcal{X}, y \in \mathcal{Y}(x)$.

Definition 4.22. For any finite recursive game domain $\mathcal{R}$, we call an $\mathcal{R}$-DDS $\mathbf{g}$ a *deterministic discrete $\mathcal{R}$-game* and an $\mathcal{R}$-PDS $\mathbf{G}$ a *probabilistic discrete $\mathcal{R}$-game.*

Definition 4.23. The optimal winning probability of a weight-1 $\mathcal{R}$-game $\mathbf{G}$ is defined as

$$\nu(\mathbf{G}) := \sup_{\mathbf{E}} \Pr^{\mathbf{EG}}(tr(\mathbf{G},\mathbf{E}) \in \mathcal{T}^w_{\mathcal{R}}),$$

where the supremum is over all weight-1 $\mathcal{R}$-PDE, and $\mathbf{EG}$ denotes the random experiment of choosing $\mathbf{E}$ according to $\mathbf{E}$ and $\mathbf{G}$ independently according to $\mathbf{G}$.

Definition 4.24. A deterministic discrete $\mathcal{R}$-game $\mathbf{g} \in \mathcal{S}_{\mathcal{R}}$ is *always-lose* if and only if its final output bit is always zero, i.e., $\mathbf{g}(\diamond) = (0, \langle\,\rangle)$ if $\diamond \in \mathcal{X}$ and otherwise $\mathbf{g}(x) = (y, \mathbf{s}')$ for an always-lose game $\mathbf{s}'$. If $\mathbf{g}$ is not always-lose it is *winnable.* This partitions $\mathcal{S}_{\mathcal{R}}$ into the set $\mathcal{S}^w_{\mathcal{R}}$ of winnable games and the set $\mathcal{S}^l_{\mathcal{R}}$ of always-lose games.

Definition 4.25. The (environment-less) *winnability* $\rho(\mathbf{G})$ of an $\mathcal{R}$-game $\mathbf{G}$ is defined by

$$\rho(\mathbf{G}) := \widehat{\mathbf{G}}(\mathcal{S}^w_{\mathcal{R}}).$$

Moreover, the (environment-less) *winnability* $\widehat{\nu}(\mathcal{G})$ of a set $\mathcal{G}$ of probabilistic $\mathcal{R}$-games is defined by

$$\widehat{\nu}(\mathcal{G}) := \inf_{\mathbf{G}\in\mathcal{G}} \rho(\mathbf{G}).$$

The following lemma states that the environment-less winnability $\widehat{\nu}([\mathbf{G}])$ of an equivalence class $[\mathbf{G}]$ coincides with the optimal winning probability $\mathbf{G}$ (for any probabilistic game $\mathbf{G}$).

Lemma 4.26. *For any probabilistic weight-*1 *$\mathcal{R}$-game $\mathbf{G}$ we have*

$$\widehat{\nu}([\mathbf{G}]) = \nu(\mathbf{G}),$$

and there exists a game $\mathbf{G}^* \in [\mathbf{G}]$ such that $\rho(\mathbf{G}^*) = \widehat{\nu}([\mathbf{G}])$.

Proof (sketch). Let $\mathbf{G}_l$ be the game which is derived from $\mathbf{G}$ by fixing all final outputs in the support of $\mathbf{G}$ to $win = 0$.

$$\begin{aligned}
\Delta(\mathbf{G}, \mathbf{G}_l) &= \sup_{\mathbf{e} \in \mathcal{E}_{\mathcal{R}}} \delta(tr(\mathbf{G}, \mathbf{e}), tr(\mathbf{G}_l, \mathbf{e})) \\
&= \sup_{\mathbf{e} \in \mathcal{E}_{\mathcal{R}}} \sum_{m \in \mathcal{T}_{\mathcal{R}}} \max\left(\Pr^{\mathbf{eG}}(tr(\mathbf{G}, \mathbf{e}) = m) - \Pr^{\mathbf{eG}_l}(tr(\mathbf{G}_l, \mathbf{e}) = m), 0\right) \\
&= \sup_{\mathbf{e} \in \mathcal{E}_{\mathcal{R}}} \sum_{m \in \mathcal{T}_{\mathcal{R}}^w} \max\left(\Pr^{\mathbf{eG}}(tr(\mathbf{G}, \mathbf{e}) = m) - \Pr^{\mathbf{eG}_l}(tr(\mathbf{G}_l, \mathbf{e}) = m), 0\right) \\
&= \sup_{\mathbf{e} \in \mathcal{E}_{\mathcal{R}}} \sum_{m \in \mathcal{T}_{\mathcal{R}}^w} \Pr^{\mathbf{eG}}(tr(\mathbf{G}, \mathbf{e}) = m) \\
&= \nu(\mathbf{G}).
\end{aligned}$$

In the third step, we have used that for any non-winning transcript $m \notin \mathcal{T}_{\mathcal{R}}^w$ we have $\Pr^{\mathbf{eG}}(tr(\mathbf{G}, \mathbf{e}) = m) \leq \Pr^{\mathbf{eG}_l}(tr(\mathbf{G}_l, \mathbf{e}) = m)$.

By the Distance Lemma (Lemma 4.18) there exists $(\mathbf{G}^*, \mathbf{G}_l^*) \in [\mathbf{G}] \times [\mathbf{G}_l]$ such that $1 - |\mathbf{G}^* \sqcap \mathbf{G}_l^*| = \Delta(\mathbf{G}, \mathbf{G}_l) = \nu(\mathbf{G})$. As the winnability of $\mathbf{G}_l$ is zero, the winnability of $\mathbf{G}_l^*$ must be zero as well. Thus, the winnability of $\mathbf{G}^*$ is at most $\nu(\mathbf{G})$, i.e.,

$$\rho(\mathbf{G}^*) \leq \nu(\mathbf{G}).$$

Since the game can actually be won with probability $\nu(\mathbf{G})$, we must have for *any* $\mathbf{G}' \in [\mathbf{G}]$

$$\nu(\mathbf{G}) \leq \rho(\mathbf{G}').$$

Hence, by antisymmetry $\rho(\mathbf{G}^*) = \nu(\mathbf{G})$ and by transitivity $\rho(\mathbf{G}^*) \leq \rho(\mathbf{G}')$ for all $\mathbf{G}' \in [\mathbf{G}]$, implying $\rho(\mathbf{G}^*) = \inf_{\mathbf{G}' \in [\mathbf{G}]} \rho(\mathbf{G}')$. This concludes the proof. □

We note that Lemma 4.26 can be used to significantly simplify the proof of the following Lemma of [13].

Lemma 4.27 (informal, Lemma 6 [13])**.** *For weight-*1 *$\mathcal{R}_1$-game $\mathbf{G}_1$ and $\mathcal{R}_2$-game $\mathbf{G}_2$, let $[\mathbf{G}_1, \mathbf{G}_2]^\wedge$ be the independent parallel conjunction-composition of the two games (i.e., the game which is won exactly if both parallel subgames are won). Then*

$$\nu([\mathbf{G}_1, \mathbf{G}_2]^\wedge) = \nu(\mathbf{G}_1)\, \nu(\mathbf{G}_2).$$

4.6 Conclusions

We have given a new representation of Maurer's theory of discrete systems. Based on this new representation, gave an environment-less perspective of the distance of discrete systems as well as the winning probability of discrete games. Invoking the results on observation compatibility from Chapter 3, we proved that the new environment-less notions are actually *equivalent* to their classical (environment-based) counterparts.

We conclude by discussing some open questions and future work.

- The discrete systems as defined in this chapter are based on the usual fully-adaptive single-execution semantics. The results shown in this chapter do not immediately carry over to different semantics. It is an open question how the definitions can be generalized in a clean manner to partially-adaptive multi-execution semantics such that the Distance Lemma still holds.
- The discussed notion of games has a winning condition which is encoded into the output of the system. This is overly-specific and more general games do not have this property (see [12]). Since the environment being able to observe the winning condition has no influence on the winning probability (in the considered setting), it is easy to see that Lemma 4.26 also holds for a more general type of games. Ideally, we would also want a direct proof that does not depend on the Distance Lemma.

Chapter 5

Generalizing Indistinguishability Amplification

The goal of indistinguishability amplification is to construct an object which is ϵ-close to its ideal from objects which are only ϵ'-close to their ideal for ϵ much smaller than ϵ'. The most basic type of this construction is to XOR two independent bits $\mathbf{B}_1$ and $\mathbf{B}_2$. It is easy to verify that if $\mathbf{B}_1$ and $\mathbf{B}_2$ are ϵ_1- and ϵ_2-close to the uniform bit $\mathbf{U}$, respectively, $\mathbf{B}_1 \oplus \mathbf{B}_2$ will be $2\epsilon_1\epsilon_2$-close to the uniform bit. The crucial property of the XOR construction is the following: If at least *one* of the bits $\mathbf{B}_1$ or $\mathbf{B}_2$ is perfectly uniform, then their XOR is perfectly uniform as well. Interestingly, it suffices to assume only such a *neutralizing* property of a construction to prove an indistinguishability amplification result. The notion of *neutralizing constructions* was introduced in [13] to prove a generalization of the above XOR construction for discrete systems.

In this chapter, we further generalize the notion of *neutralizing constructions* from [13] and use the pseudo-metric $\widehat{\Delta}$ presented in Chapter 4 to prove a more general indistinguishability amplification theorem than the Product Theorem of [13].

5.1 $\mathcal{A}$-neutralizing Constructions

We first define the abstract notion of a *construction*. Intuitively, one can think of a construction as a discrete system which can be connected to multiple subsystems via an inside interface, and which communicates with environments via an outside interface. Implemented as an actual discrete system, such an object is described formally in [12] and called a *converter*. As the concrete interaction and the interfaces are immaterial to the statements made in this chapter, we introduce a more abstract object which is simply a function mapping systems to another system.

Definition 5.1. For finite recursive domains $\mathcal{R}_1,\dots,\mathcal{R}_k$ and $\mathcal{R}$, a *deterministic* $(\mathcal{R}_1,\dots,\mathcal{R}_k \rightsquigarrow \mathcal{R})$*-construction* is a function $\mathbf{c} : \mathcal{S}_{\mathcal{R}_1} \times \cdots \times \mathcal{S}_{\mathcal{R}_k} \to \mathcal{S}_{\mathcal{R}}$ for which there exists a function $f : \mathcal{E}_{\mathcal{R}} \to \mathcal{E}_{\mathcal{R}_1} \times \cdots \times \mathcal{E}_{\mathcal{R}_k}$ and a function $g : \mathcal{T}_{\mathcal{R}_1} \times \cdots \times \mathcal{T}_{\mathcal{R}_k} \to \mathcal{T}_{\mathcal{R}}$ such that for every $(\mathbf{s}_1,\dots,\mathbf{s}_k) \in \mathcal{S}_{\mathcal{R}_1} \times \cdots \times \mathcal{S}_{\mathcal{R}_k}$ and $\mathbf{e} \in \mathcal{E}_{\mathcal{R}}$

$$tr(\mathbf{c}(\mathbf{s}_1,\dots,\mathbf{s}_k),\mathbf{e}) = g(tr(\mathbf{s}_1,\mathbf{e}_1),\dots,tr(\mathbf{s}_1,\mathbf{e}_k)) \quad \text{for } (\mathbf{e}_1,\dots,\mathbf{e}_k) = f(\mathbf{e}).$$

Moreover, a *probabilistic* $(\mathcal{R}_1,\dots,\mathcal{R}_k \rightsquigarrow \mathcal{R})$*-construction* $\mathbf{C}$ is a distribution over deterministic $(\mathcal{R}_1,\dots,\mathcal{R}_k \rightsquigarrow \mathcal{R})$-constructions based on the monoid $\mathbb{R}_{\geq 0}$.

Remark 5.2. The above definition of a construction can be easily generalized for abstract objects $\mathcal{A}_1,\dots,\mathcal{A}_k$ and $\mathcal{A}$ that are observable by functions $\mathcal{F}_1,\dots,\mathcal{F}_k$ and $\mathcal{F}$, respectively.

In the following, expressions involving multiple weight-1 distributions denote the *independent composition* of the distributions. For example, for a weight-1 construction $\mathbf{C}$ and weight-1 $\mathcal{R}_i$-PDS $\mathbf{S}_i$, the $\mathcal{R}$-PDS $\mathbf{C}(\mathbf{S}_1,\dots,\mathbf{S}_k)$ is defined by

$$\mathbf{C}(\mathbf{S}_1,\dots,\mathbf{S}_k)(\mathbf{s}) := \sum_{\substack{\mathbf{c}\in\mathsf{supp}(\mathbf{C}),\mathbf{s}_i\in\mathcal{S}_{\mathcal{R}_i}\\ \mathbf{c}(\mathbf{s}_1,\dots,\mathbf{s}_k)=\mathbf{s}}} \mathbf{C}(\mathbf{c}) \cdot \prod_{i\in[k]} \mathbf{S}_i(\mathbf{s}_i).$$

The proof of the following lemma is straightforward via Lemma 4.12 and thus omitted.

Lemma 5.3. *For a probabilistic* $(\mathcal{R}_1,\dots,\mathcal{R}_k \rightsquigarrow \mathcal{R})$*-construction* $\mathbf{C}$ *and* k *pairs* $(\mathbf{S}_i,\mathbf{S}'_i)$ *such that* $\mathbf{S}_i$ *and* $\mathbf{S}'_i$ *are weight-*1 $\mathcal{R}_i$*-PDS with* $\mathbf{S}_i \equiv \mathbf{S}'_i$ *we have*

$$\mathbf{C}(\mathbf{S}_1,\dots,\mathbf{S}_k) \equiv \mathbf{C}(\mathbf{S}'_1,\dots,\mathbf{S}'_k).$$

Notation 5.4. For k pairs $(x_1,y_1),\dots,(x_k,y_k)$ with $x_i,y_i \in \mathcal{Z}_i$ and $b \in \{0,1\}^k$ we define

$$\langle x_1/y_1,\dots,x_k/y_k\rangle_b := (z_1,\dots,z_k),$$

where $z_i = x_i$ if $b_i = 0$ and $z_i = y_i$ otherwise ($b_i = 1$).

Definition 5.5. A probabilistic k-ary construction $\mathbf{C}$ is $\mathcal{A}$*-neutralizing* for $(\mathbf{F}_1,\mathbf{I}_1),\dots,(\mathbf{F}_k,\mathbf{I}_k)$ and a monotone[1] set $\mathcal{A} \subseteq \{0,1\}^k$ if for any choice of bits $b \in \mathcal{A}$ we have

$$\mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1,\dots,\mathbf{F}_k/\mathbf{I}_k\rangle_b) \equiv \mathbf{C}(\mathbf{I}_1,\dots,\mathbf{I}_k).$$

Definition 5.6. A probabilistic k-ary construction $\mathbf{C}$ is q*-neutralizing* for $(\mathbf{F}_1,\mathbf{I}_1),\dots,(\mathbf{F}_k,\mathbf{I}_k)$ if it is $\mathcal{A}$-neutralizing for[2] $\mathcal{A} = \{b \mid b \in \{0,1\}^k, \mathsf{hw}(b) \geq q\}$.

[1] A set $\mathcal{A} \subseteq \{0,1\}^k$ is monotone if for every $a \in \mathcal{A}$ we have $a' \in \mathcal{A}$ for every $a' \in \{0,1\}^k$ with $a'_i \geq a_i$.

[2] Recall that $\mathsf{hw}(b)$ denotes the hamming weight of b.

5.1.1 A General $\mathcal{A}$-neutralizing Construction

We briefly present a simple $\mathcal{A}$-neutralizing construction for general $\mathcal{A}$.

For a finite field $\mathbb{F}$, let $A \in \mathbb{F}^{q \times k}$ be a $(q \times k)$-matrix for $q \leq k$, and let $\mathcal{A} \subseteq \{0,1\}^k$ be the (monotone) set containing all $v \in \{0,1\}^k$ with $v_{i_1} = \cdots = v_{i_q} = 1$ for q distinct indices, such that the columns $i_1, \ldots, i_q$ of A are linearly independent.

Consider the deterministic construction $\mathbf{c} : \mathbb{F}^k \to \mathbb{F}^q$ defined by[3]

$$\mathbf{c}(x_1, \ldots, x_k) := A \cdot (x_1, \ldots, x_k)^{\mathsf{T}}.$$

It is easy to see that $\mathbf{c}$ is $\mathcal{A}$-neutralizing for $(\mathbf{X}_1, \mathbf{U}), \ldots, (\mathbf{X}_k, \mathbf{U})$, where $\mathbf{X}_i$ are arbitrary weight-1 distributions over $\mathbb{F}$ and $\mathbf{U}$ is the uniform distribution over $\mathbb{F}$.

Observe moreover that $\mathbf{c}$ is q-neutralizing if A is a hyper-invertible matrix as introduced in [2]. Assuming the field $\mathbb{F}$ has sufficiently many elements ($|\mathbb{F}| \geq q + k$) such a matrix always exists (see [2] for a concrete polynomial-based construction).

More generally, consider the construction $\mathbf{c}'$ which combines k functions $f_i : \mathcal{X} \to \mathbb{F}$ for some finite set $\mathcal{X}$ to q functions $f'_j : \mathcal{X} \to \mathbb{F}$ by

$$\mathbf{c}'(f_1, \ldots, f_k) := \left(f'_1, \ldots, f'_q\right),$$

where $f'_i := x \mapsto A_i \cdot (f_1(x), \ldots, f_k(x))^{\mathsf{T}}$ and A_i is the i-th row of A. As above, $\mathbf{c}'$ is $\mathcal{A}$-neutralizing for $(\mathbf{F}_1, \mathbf{R}), \ldots, (\mathbf{F}_k, \mathbf{R})$, where $\mathbf{F}_i$ are arbitrary weight-1 distributions over $\mathbb{F}^{\mathcal{X}}$ and $\mathbf{R}$ is the uniform distribution over $\mathbb{F}^{\mathcal{X}}$.

5.2 Indistinguishability Amplification for q-neutralizing Constructions

Lemma 5.7 (cf. Lemma 3 of [13]). *For any weight-1 $\mathcal{R}$-PDS $\mathbf{S}, \mathbf{T}$ and any weight-1 $\mathbf{B} \in \Gamma_{\{0,1\}\mathbb{R}_{\geq 0}}$*

$$\widehat{\Delta}(\langle \mathbf{S}/\mathbf{T} \rangle_{\mathbf{B}}, \mathbf{T}) = \mathbf{B}(0) \cdot \widehat{\Delta}(\mathbf{S}, \mathbf{T}).$$

Proof. We give a succinct proof via the classical distinguishing advantage and the Distance Lemma. A proof without Distance Lemma is in principle possible

[3] One can think of an element of $\mathbb{F}^n$ as a single-query DDS with unary input alphabet $\{\diamond\}$ and output alphabet $\mathbb{F}^n$.

though more involved. Observe that

$$\begin{aligned}
&\Delta^{\mathbf{D}}(\langle \mathbf{S}/\mathbf{T}\rangle_{\mathbf{B}}, \mathbf{T}) \\
&= \Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T},\mathrm{E})) = 1) - \Pr^{\mathbf{D}\langle \mathbf{S}/\mathbf{T}\rangle_{\mathbf{B}}}(\mathrm{B}(tr(\langle \mathrm{S}/\mathrm{T}\rangle_{\mathrm{B}},\mathrm{E})) = 1) \\
&= \mathbf{B}(0) \cdot \Big(\Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T},\mathrm{E})) = 1) - \Pr^{\mathbf{DS}}(\mathrm{B}(tr(\mathrm{S},\mathrm{E})) = 1)\Big) \\
&\quad + \mathbf{B}(1) \cdot \Big(\Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T},\mathrm{E})) = 1) - \Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T},\mathrm{E})) = 1)\Big) \\
&= \mathbf{B}(0) \cdot \Big(\Pr^{\mathbf{DT}}(\mathrm{B}(tr(\mathrm{T},\mathrm{E})) = 1) - \Pr^{\mathbf{DS}}(\mathrm{B}(tr(\mathrm{S},\mathrm{E})) = 1)\Big) \\
&= \mathbf{B}(0) \cdot \Delta^{\mathbf{D}}(\mathbf{S}, \mathbf{T}).
\end{aligned}$$

The claim follows from the Distance Lemma (Lemma 4.18). □

The following lemma describes a general proof technique and can be used as a tool to prove indistinguishability amplification results for any $\mathcal{A}$-neutralizing construction.

Lemma 5.8. *Let the probabilistic k-ary construction $\mathbf{C}$ be $\mathcal{A}$-neutralizing for $(\mathbf{F}_1, \mathbf{I}_1), \ldots, (\mathbf{F}_k, \mathbf{I}_k)$ and let $\mathbf{B}, \mathbf{B}' \in \Gamma_{(\mathcal{A}\cup\{0^k\})\mathbb{R}_{\geq 0}}$ be weight-1 distributions such that $\mathbf{B}(0^k) > 0$ and $\mathbf{B}'(0^k) = 0$. Then,*

$$\begin{aligned}
&\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \\
&\qquad \leq \mathbf{B}(0^k)^{-1} \cdot \sum_{e \in \{0,1\}^k} \delta(mask(\mathbf{B}, e), mask(\mathbf{B}', e)) \cdot \mathbf{E}(e),
\end{aligned}$$

where $mask(x, m)$ is the tuple derived from x by removing all elements at the indices at which $m_i = 0$, and $\mathbf{E}(e_1, \ldots, e_k) = \prod_{i\in[k]} \mathrm{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

Proof. By Lemma 5.7 we have for weight-1 $\mathbf{B}'' \in \Gamma_{\{0,1\}\mathbb{R}_{\geq 0}}$ with $\mathbf{B}''(0) = \mathbf{B}(0^k)$

$$\begin{aligned}
&\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \\
&= \mathbf{B}(0^k)^{-1} \cdot \widehat{\Delta}(\langle \mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k)/\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)\rangle_{\mathbf{B}''}, \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k))
\end{aligned}$$

Observe that $\langle \mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k)/\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)\rangle_{\mathbf{B}''} \equiv \mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \ldots, \mathbf{F}_k/\mathbf{I}_k\rangle_{\mathbf{B}})$ and $\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k) \equiv \mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \ldots, \mathbf{F}_k/\mathbf{I}_k\rangle_{\mathbf{B}'})$ since $\mathbf{C}$ is $\mathcal{A}$-neutralizing. Thus,

$$\begin{aligned}
&\widehat{\Delta}(\langle \mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k)/\mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)\rangle_{\mathbf{B}''}, \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \\
&\quad = \widehat{\Delta}(\mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \ldots, \mathbf{F}_k/\mathbf{I}_k\rangle_{\mathbf{B}}), \mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \ldots, \mathbf{F}_k/\mathbf{I}_k\rangle_{\mathbf{B}'})).
\end{aligned}$$

According to the Distance Lemma[4] (Lemma 4.18) there exist $(\mathbf{F}'_i, \mathbf{I}'_i) \in [\mathbf{F}_i] \times [\mathbf{I}_i]$ for every $i \in [k]$ such that $|\mathbf{F}'_i \sqcap \mathbf{I}'_i| = |[\mathbf{F}_i] \sqcap [\mathbf{I}_i]|$. Thus,

$$\begin{aligned}
&\widehat{\Delta}(\mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \dots, \mathbf{F}_k/\mathbf{I}_k \rangle_{\mathbf{B}}), \mathbf{C}(\langle \mathbf{F}_1/\mathbf{I}_1, \dots, \mathbf{F}_k/\mathbf{I}_k \rangle_{\mathbf{B}'})) \\
&= \widehat{\Delta}(\mathbf{C}(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}}), \mathbf{C}(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}'})) \\
&\le \delta(\mathbf{C}(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}}), \mathbf{C}(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}'})) \\
&\le \delta(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}}, \langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}'}),
\end{aligned}$$

where the last step is due to Lemma 3.13.

We exhibit a random experiment $\mathcal{E}$ with random variables $\mathrm{F}'_i \sim \mathbf{F}'_i, \mathrm{I}'_i \sim \mathbf{I}'_i$, $\mathrm{B} \sim \mathbf{B}$, and $\mathrm{B}' \sim \mathbf{B}'$, such that $\mathrm{L} := \langle \mathrm{F}'_1/\mathrm{I}'_1, \dots, \mathrm{F}'_k/\mathrm{I}'_k \rangle_{\mathrm{B}} \sim \langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}}$ and $\mathbf{R} := \langle \mathrm{F}'_1/\mathrm{I}'_1, \dots, \mathrm{F}'_k/\mathrm{I}'_k \rangle_{\mathrm{B}'} \sim \langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}'}$. Define $\mathrm{E}_i := [\mathrm{F}'_i \neq \mathrm{I}'_i]$ and $\mathrm{E} := (\mathrm{E}_1, \dots, \mathrm{E}_k)$.

Observe that the joint distribution of F'_i and I'_i as well as B and B' can be chosen arbitrary (as long as the marginal distributions are respected). Recall the definition of $\mathcal{C}_\delta(\cdot, \cdot)$ from Lemma 3.33. Let the joint distribution of F'_i and I'_i be $\mathcal{C}_\delta(\mathbf{F}'_i, \mathbf{I}'_i)$. Moreover, the joint distribution of B and B' depends on E such that[5]

$$\begin{aligned}
&\mathrm{Pr}^{\mathcal{E}}(\mathit{mask}(\mathrm{B}, e) = b, \mathit{mask}(\mathrm{B}', e) = b', \mathrm{E} = e) \\
&\qquad = \mathcal{C}_\delta(\mathit{mask}(\mathbf{B}, e), \mathit{mask}(\mathbf{B}', e))(b, b') \cdot \mathbf{E}(e).
\end{aligned}$$

Thus we have by Lemma 3.33

$$\begin{aligned}
&\delta(\langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}}, \langle \mathbf{F}'_1/\mathbf{I}'_1, \dots, \mathbf{F}'_k/\mathbf{I}'_k \rangle_{\mathbf{B}'}) \\
&\le \mathrm{Pr}^{\mathcal{E}}(\mathrm{L} \neq \mathrm{R}) \\
&= \sum_{e \in \{0,1\}^k} \mathrm{Pr}^{\mathcal{E}}(\mathrm{L} \neq \mathrm{R}, \mathrm{E} = e) \\
&= \sum_{e \in \{0,1\}^k} \mathrm{Pr}^{\mathcal{E}}(\mathit{mask}(\mathrm{B}, e) \neq \mathit{mask}(\mathrm{B}', e), \mathrm{E} = e) \\
&= \sum_{e \in \{0,1\}^k} \delta(\mathit{mask}(\mathbf{B}, e), \mathit{mask}(\mathbf{B}', e)) \cdot \mathbf{E}(e),
\end{aligned}$$

which concludes the proof. □

Using Lemma 5.8 we show the following indistinguishability amplification theorem for all q-neutralizing constructions.

[4] The Distance Lemma is invoked merely for the *existence* of $(\mathbf{F}'_i, \mathbf{I}'_i) \in [\mathbf{F}_i] \times [\mathbf{I}_i]$ with $|\mathbf{F}'_i \sqcap \mathbf{I}'_i| = |[\mathbf{F}_i] \sqcap [\mathbf{I}_i]|$.

[5] Note that even though the joint distribution of B and B' depends on E, the random variable B is still independent of $((\mathrm{F}'_1, \mathrm{I}'_1), \dots, (\mathrm{F}'_k, \mathrm{I}'_k))$.

Theorem 5.9. *If the probabilistic k-ary construction $\mathbf{C}$ is q-neutralizing for $(\mathbf{F}_1, \mathbf{I}_1), \ldots, (\mathbf{F}_k, \mathbf{I}_k)$, then*

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \leq \sum_{i=k-q+1}^{k} f_{i-(k-q),i} \cdot \mathsf{hw}(\mathbf{E})(i),$$

where

$$f_{n,m} := \frac{1}{2} \cdot \left(1 + \sum_{j=n}^{m} \binom{m}{j} \cdot \binom{j-1}{n-1}\right),$$

and $\mathbf{E}(e_1, \ldots, e_k) = \prod_{i \in [k]} \mathrm{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

Proof. For $q \geq 1$ and $k \geq q$ we represent distributions $\mathbf{B}_{q,k}, \mathbf{B}'_{q,k}$ using multisets $A_{q,k}, A'_{q,k}$ over $\mathcal{A} \cup \{0^k\}$, with the natural understanding that $\mathbf{B}_{q,k}$ ($\mathbf{B}'_{q,k}$) is uniformly distributed over $A_{q,k}$ ($A'_{q,k}$), i.e., $\mathbf{B}_{q,k}(a) = A_{q,k}(a)/|A_{q,k}|$.

Let

$$A'_{q,k} := \bigcup_{j \in \{q, q+2, \ldots, k\}} \left\{ \left(b, \binom{j-1}{q-1}\right) \mid b \in \{0,1\}^k, \mathsf{hw}(b) = j \right\} \quad \text{and}$$

$$A_{q,k} := \{(0^k, 1)\} \cup \bigcup_{j \in \{q+1, q+3, \ldots, k\}} \left\{ \left(b, \binom{j-1}{q-1}\right) \mid b \in \{0,1\}^k, \mathsf{hw}(b) = j \right\}.$$

For a multiset M over $\{0,1\}^n$, let $\mathit{blind}_m(M)$ be the multiset over $\{0,1\}^{n-m}$ derived from M by removing the bits at m fixed positions, say the first m bits, for every element. We only consider multisets for which $\mathit{blind}_m(M)$ is well-defined, i.e., it does not matter at which m positions the bits are removed. We prove the following statement

$$\begin{aligned}\forall q \geq 1,\, \forall k \geq q: \quad & |A_{q,k}| = |A'_{q,k}| = f_{q,k} \\ & \wedge \forall j \geq q : \mathit{blind}_j(A_{q,k}) = \mathit{blind}_j(A'_{q,k}) \\ & \wedge \forall j < q : |\mathit{blind}_j(A_{q,k}) \triangle \mathit{blind}_j(A'_{q,k})| = 2 f_{q-j,k-j}.\end{aligned} \tag{5.1}$$

This implies the claim via Lemma 5.8, since we have

$$\begin{aligned}&\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \\ &\quad \leq \mathbf{B}_{q,k}(0^k)^{-1} \cdot \sum_{e \in \{0,1\}^k} \delta(\mathit{mask}(\mathbf{B}_{q,k}, e), \mathit{mask}(\mathbf{B}'_{q,k}, e)) \cdot \mathbf{E}(e) \\ &\quad = |A_{q,k}| \cdot \sum_{i=0}^{k} \frac{|\mathit{blind}_{k-i}(A_{q,k}) \triangle \mathit{blind}_{k-i}(A'_{q,k})|}{2|A_{q,k}|} \cdot \mathsf{hw}(\mathbf{E})(i) \\ &\quad = \sum_{i=k-q+1}^{k} f_{i-(k-q),i} \cdot \mathsf{hw}(\mathbf{E})(i).\end{aligned}$$

In the second step we have used that for any two multisets M, M' representing uniform weight-1 distributions $\mathbf{M}, \mathbf{M}'$ we have $\delta(\mathbf{M}, \mathbf{M}') = |M \triangle M'|/(2|M|)$ if $|M| = |M'|$.

We prove (5.1) by induction over k. Observe that

$$\begin{aligned} \mathit{blind}_1(A'_{q,k}) = & \bigcup_{j \in \{q, q+2, \ldots, k-1\}} \left\{ \left(b, \binom{j-1}{q-1} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\} \\ & \cup \bigcup_{j \in \{q-1, q+1, \ldots, k-1\}} \left\{ \left(b, \binom{j}{q-1} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\}. \end{aligned}$$

Similarly, we see that

$$\begin{aligned} \mathit{blind}_1(A_{q,k}) = & \{(0^{k-1}, 1)\} \\ & \cup \bigcup_{j \in \{q+1, q+3, \ldots, k-1\}} \left\{ \left(b, \binom{j-1}{q-1} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\} \\ & \cup \bigcup_{j \in \{q, q+2, \ldots, k-1\}} \left\{ \left(b, \binom{j}{q-1} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\}. \end{aligned}$$

If $q = 1$, it is easy to see that $|A_{q,k}| = |A'_{q,k}| = f_{q,k}$, as well as $\mathit{blind}_1(A'_{q,k}) = \mathit{blind}_1(A_{q,k})$ and $|\mathit{blind}_0(A_{q,k}) \triangle \mathit{blind}_0(A'_{q,k})| = 2f_{q,k}$ (since $A_{q,k}$ and $A'_{q,k}$ are disjoint).

Otherwise ($q \geq 2$), we use the well-known recurrence $\binom{j}{q-1} - \binom{j-1}{q-1} = \binom{j-1}{q-2}$ to obtain

$$\begin{aligned} & \mathit{blind}_1(A'_{q,k}) - \mathit{blind}_1(A_{q,k}) \cap \mathit{blind}_1(A'_{q,k}) \\ & = \bigcup_{j \in \{q-1, q+1, \ldots, k-1\}} \left\{ \left(b, \binom{j-1}{q-2} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\} \\ & = A'_{q-1,k-1}. \end{aligned}$$

Analogously, we see that

$$\begin{aligned} & \mathit{blind}_1(A_{q,k}) - \mathit{blind}_1(A_{q,k}) \cap \mathit{blind}_1(A'_{q,k}) \\ & = \{(0^{k-1}, 1)\} \cup \bigcup_{j \in \{q, q+2, \ldots, k-1\}} \left\{ \left(b, \binom{j-1}{q-2} \right) \mid b \in \{0,1\}^{k-1}, \mathsf{hw}(b) = j \right\} \\ & = A_{q-1,k-1}. \end{aligned}$$

As by induction hypothesis $\mathit{blind}_{q-1}(A_{q-1,k-1}) = \mathit{blind}_{q-1}(A'_{q-1,k-1})$, we have $\mathit{blind}_q(A_{q,k}) = \mathit{blind}_q(A'_{q,k})$. Since blinding does not change the cardinality of a multiset, it follows $|A_{q,k}| = |A'_{q,k}| = f_{q,k}$. Moreover, as $A_{q,k}$ and $A'_{q,k}$ are

$f_{n,m}$	m=1	2	3	4	5	6	7	8	9	10	11	12
n=1	1	2	4	8	16	**32**	64	128	256	512	1024	2048
2	–	1	3	9	25	65	**161**	385	897	2049	4609	10241
3	–	–	1	4	16	56	176	**512**	1408	3712	9472	23552
4	–	–	–	1	5	25	105	385	1281	3969	11649	32769
5	–	–	–	–	1	6	36	176	736	2752	9472	30592
6	–	–	–	–	–	1	7	49	273	1281	5313	20097
7	–	–	–	–	–	–	1	8	64	400	2080	9472
8	–	–	–	–	–	–	–	1	9	81	561	3201
9	–	–	–	–	–	–	–	–	1	10	100	760
10	–	–	–	–	–	–	–	–	–	1	11	121
11	–	–	–	–	–	–	–	–	–	–	1	12
12	–	–	–	–	–	–	–	–	–	–	–	1

Table 5.1: Values of $f_{n,m}$ for $m \leq 12$. For a k-ary q-neutralizing construction, the bound of Theorem 5.9 depends on the values $f_{1,k-q+1}, f_{2,k-q+2}, \ldots, f_{q,k}$. For $q = 3$ and $k = 8$, the corresponding values are marked in bold.

disjoint we have $|blind_0(A_{q,k}) \triangle blind_0(A'_{q,k})| = 2f_{q,k}$. Finally, for $j \geq 1$ and $j < q$ we have

$$\begin{aligned}|blind_j(A_{q,k}) \triangle blind_j(A'_{q,k})| &= |blind_{j-1}(A_{q-1,k-1}) \triangle blind_{j-1}(A'_{q-1,k-1})| \\ &\overset{\text{(I.H.)}}{=} 2f_{(q-1)-(j-1),(k-1)-(j-1)} = 2f_{q-j,k-j},\end{aligned}$$

which concludes the proof.

□

Example 5.10. If the probabilistic 8-ary construction $\mathbf{C}$ is 3-neutralizing for $(\mathbf{F}_1, \mathbf{I}_1), \ldots, (\mathbf{F}_8, \mathbf{I}_8)$ we can obtain the required values of $f_{n,m}$ from Table 5.1 to conclude via Theorem 5.9 that

$$\begin{aligned}&\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_8), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_8)) \\ &\qquad \leq 32 \cdot \mathsf{hw}(\mathbf{E})(6) + 161 \cdot \mathsf{hw}(\mathbf{E})(7) + 512 \cdot \mathsf{hw}(\mathbf{E})(8),\end{aligned}$$

where $\mathbf{E}(e_1, \ldots, e_8) = \prod_{i \in [8]} \text{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

Corollary 5.11 (Theorem 1 [13])**.** *If the probabilistic k-ary construction $\mathbf{C}$ is 1-neutralizing for $(\mathbf{F}_1, \mathbf{I}_1), \ldots, (\mathbf{F}_k, \mathbf{I}_k)$, then*

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \leq 2^{k-1} \cdot \mathsf{hw}(\mathbf{E})(k) = 2^{k-1} \cdot \prod_{i \in [k]} \widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i),$$

where $\mathbf{E}(e_1, \ldots, e_k) = \prod_{i \in [k]} \text{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

Proof. We invoke Theorem 5.9 with $q = 1$ to obtain

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \leq f_{1,k} \cdot \mathsf{hw}(\mathbf{E})(k)$$

It remains only to see that

$$f_{1,k} = \frac{1}{2} \cdot \left(1 + \sum_{j=1}^{k} \binom{k}{j} \cdot \binom{j-1}{0} \right) = \frac{1}{2} \cdot \sum_{j=0}^{k} \binom{k}{j} = 2^{k-1}.$$

□

Corollary 5.12 ($q = k - 1$)**.** *If the probabilistic k-ary construction* $\mathbf{C}$ *is* $(k-1)$*-neutralizing for* $(\mathbf{F}_1, \mathbf{I}_1), \dots, (\mathbf{F}_k, \mathbf{I}_k)$*, then*

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \leq \sum_{i=2}^{k} i \cdot \mathsf{hw}(\mathbf{E})(i),$$

where $\mathbf{E}(e_1, \dots, e_k) = \prod_{i \in [k]} \mathrm{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$*.*

Proof. We invoke Theorem 5.9 with $q = k - 1$ to obtain

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \leq \sum_{i=2}^{k} f_{i-1,i} \cdot \mathsf{hw}(\mathbf{E})(i).$$

It remains only to see that

$$f_{i-1,i} = \frac{1}{2} \cdot \left(1 + \sum_{j=i-1}^{i} \binom{i}{j} \cdot \binom{j-1}{i-2} \right) = \frac{1}{2} \cdot (1 + (i + (i-1))) = i.$$

□

5.3 Understanding the Bound

In this section, we discuss different aspects of the bound shown in Theorem 5.9 and finally present a simplified (non-tight) variant.

5.3.1 Parametrized Setting

Consider a parametrized k-ary q-neutralizing construction $\mathbf{C}_{k,q}$ (with parameters $q \geq 1$ and $k \geq q$) for an infinite sequence of pairs $(\mathbf{F}_1, \mathbf{I}_1), (\mathbf{F}_2, \mathbf{I}_2), \dots$. Assume $\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i) \leq \epsilon$ for all $i \in \mathbb{N}$ and a fixed $\epsilon < \frac{1}{2}$. Let $k^*_{q,\epsilon,\delta}$ be the smallest k such that

$$\widehat{\Delta}(\mathbf{C}_{k,q}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}_{k,q}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \leq \delta.$$

Table 5.2 lists the upper bound $k_{q,\epsilon,\delta}$ of $k^*_{q,\epsilon,\delta}$ for $\delta = 2^{-100}$ which is implied by Theorem 5.9.

$k_{q,\epsilon,\delta}$	$\epsilon{=}10^{-5}$	10^{-3}	0.01	0.05	0.1	0.2	0.3	0.4	0.49
$q{=}1$	7	12	18	30	43	75	135	308	3397
2	8	13	20	33	47	81	146	335	3806
3	9	14	21	35	50	86	155	360	4190
4	10	16	23	37	52	91	164	383	4563
5	11	17	24	39	55	96	173	405	4927
10	17	23	31	49	67	117	213	511	6697
20	27	34	44	65	89	155	286	708	10143
50	58	66	80	111	149	258	490	1275	20349
10^2	109	119	137	183	242	422	818	2202	37297
10^3	1012	1035	1105	1381	1815	3263	6625	18776	–

Table 5.2: Upper bound $k_{q,\epsilon,\delta}$ of $k^*_{q,\epsilon,\delta}$ for $\delta = 2^{-100}$ and different q, ϵ.

Example 5.13. Let $\mathbf{c}_{k,q}$ denote the deterministic q-neutralizing construction from Section 5.1.1 which transforms k independent random functions from $\mathcal{X}$ to $\mathbb{F}$ to q random functions from $\mathcal{X}$ to $\mathbb{F}$ for a finite field $\mathbb{F}$. Observe that $\mathbf{c}_{k,q}$ is exactly of the parametrized form described above.

Consider the following scenario: We would like to have (an object very close to) $(\mathbf{R}_1, \dots, \mathbf{R}_q)$, which denotes q independent uniform random functions from $\mathcal{X}$ to $\mathbb{F}$. However, our resources consist only of independent random functions $\mathbf{F}_i$ with $\widehat{\Delta}(\mathbf{F}_i, \mathbf{R}) = \epsilon$ for a too large ϵ. For fixed q, $\epsilon < \frac{1}{2}$, and $\delta \in [0, 1]$, we are thus interested in the smallest number k of random functions we need to combine to obtain

$$\widehat{\Delta}(\mathbf{c}_{k,q}(\mathbf{F}_1, \dots, \mathbf{F}_k), (\mathbf{R}_1, \dots, \mathbf{R}_q)) \leq \delta.$$

For example, if our random functions $\mathbf{F}_i$ are all 0.1-close to the uniform random function, we can construct $q = 5$ random functions that are (altogether!) 2^{-100}-close to 5 independent uniform functions by combining 55 random functions (see Table 5.2) using $\mathbf{c}_{55,5}$.

5.3.2 Amplification Threshold

For any q and k with $q \leq k$, we define the *amplification threshold* $\epsilon^*_{q,k}$ as the largest real number for which $\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i) < \epsilon^*_{q,k}$ for all $i \in [k]$ implies

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) < \epsilon^*_{q,k},$$

assuming $\mathbf{C}$ is q-neutralizing for the pairs $(\mathbf{F}_i, \mathbf{I}_i)$. Table 5.3 contains the a lower bound $\epsilon_{q,k} \leq \epsilon^*_{q,k}$ of the amplification threshold that is implied by Theorem 5.9. Note that $\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i) < \epsilon^*_{q,k}$ does not imply *strong* amplification.

$\epsilon_{q,k}$	k=2	3	4	5	6	7	8	9
q=1	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
2	–	0.1835	0.2659	0.3044	0.3285	0.3458	0.3590	0.3697
3	–	–	0.0914	0.1706	0.2140	0.2430	0.2645	0.2816
4	–	–	–	0.0542	0.1210	0.1624	0.1913	0.2135
5	–	–	–	–	0.0358	0.0915	0.1292	0.1567
6	–	–	–	–	–	0.0253	0.0722	0.1064
7	–	–	–	–	–	–	0.0188	0.0588
8	–	–	–	–	–	–	–	0.0146

Table 5.3: Lower bound $\epsilon_{q,k}$ of the amplification threshold $\epsilon^*_{q,k}$ for $k \leq 12$, rounded down. For $q = 1$ it is known that $\epsilon_{q,k} = \epsilon^*_{q,k}$.

$\tau_{q,k}$	k=2	3	4	5	6	7	8	9
q=1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	–	0.5000	0.7676	0.8689	0.9164	0.9422	0.9577	0.9677
3	–	–	0.2325	0.5000	0.6529	0.7442	0.8026	0.8424
4	–	–	–	0.1312	0.3472	0.5000	0.6045	0.6781
5	–	–	–	–	0.0837	0.2559	0.3956	0.5000
6	–	–	–	–	–	0.0579	0.1975	0.3220
7	–	–	–	–	–	–	0.0424	0.1577
8	–	–	–	–	–	–	–	0.0324

Table 5.4: Conjectured upper bound $\tau_{q,k}$ of the amplification threshold $\epsilon^*_{q,k}$ for $k \leq 12$, rounded up.

Table 5.4 shows the value $\tau_{q,k}$ which we conjecture to be a (very loose) upper bound of of $\epsilon^*_{q,k}$. It is derived under the (sole) assumption that there exists a q-neutralizing k-ary construction for every q and $k \geq q$ such that

$$\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \ldots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \ldots, \mathbf{I}_k)) \geq \sum_{i=k-q+1}^{k} \mathsf{hw}(\mathbf{E})(i),$$

where $\mathbf{E}(e_1, \ldots, e_k) = \prod_{i \in [k]} \text{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

5.3.3 A Simpler Bound

The following lemma gives an alternative representation of $f_{n,m}$ as well as an almost-tight closed-form bound.

Lemma 5.14. *For any $n \geq 1$ and $m \geq n$ we have* *(i)* $t_{n,m} = 2t_{n,m-1} + t_{n-1,m-1}$, *where* $t_{n,m} := \sum_{j=n}^{m} \binom{m}{j} \cdot \binom{j-1}{n-1}$. *This implies*

$$f_{n,m} = 2f_{n,m-1} + f_{n-1,m-1} - 1.$$

(ii)

$$2^{m-n} \cdot \binom{m-1}{n-1} \in \left[f_{n,m},\ 2f_{n,m} - \frac{1}{3} \right].$$

Proof (sketch). *(i)* is straightforward to show by induction over m. *(ii)* follows from *(i)*, also by induction. □

We present a simpler variant of the bound of Theorem 5.9 which follows from Lemma 5.14. Beware that this simple bound is strictly larger than the original bound, and thus cannot be tight.

Corollary 5.15. *If k-ary $\mathbf{C}(\cdot, \dots, \cdot)$ is q-neutralizing for $(\mathbf{F}_1, \mathbf{I}_1), \dots, (\mathbf{F}_k, \mathbf{I}_k)$, then*

(i)

$$\begin{aligned} &\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \\ &\leq 2^{k-q} \sum_{i=k-q+1}^{k} \binom{i-1}{i-1-(k-q)} \cdot \mathsf{hw}(\mathbf{E})(i) \\ &= 2^{k-q} \sum_{j=0}^{q-1} \binom{j+(k-q)}{j} \cdot \mathsf{hw}(\mathbf{E})(k-q+1+j), \end{aligned}$$

where $\mathbf{E}(e_1, \dots, e_k) = \prod_{i \in [k]} \mathrm{Bernoulli}(\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i))(e_i)$.

(ii) if $\widehat{\Delta}(\mathbf{F}_i, \mathbf{I}_i) \leq \epsilon$ *for all* $i \in [k]$ *we have*

$$\begin{aligned} &\widehat{\Delta}(\mathbf{C}(\mathbf{F}_1, \dots, \mathbf{F}_k), \mathbf{C}(\mathbf{I}_1, \dots, \mathbf{I}_k)) \\ &\leq 2^{k-q} \epsilon^{k-q+1} \sum_{j=0}^{q-1} \binom{j+k-q}{j} \binom{k}{q-1-j} \epsilon^{j} \\ &\leq 2^{k-q+1} \frac{k!}{(k-q+1)!} \cdot \epsilon^{k-q+1}. \end{aligned}$$

Proof (sketch). *(i)* follows directly from Theorem 5.9 and Lemma 5.14 *(ii)*. *(ii)* follows by further bounding the expression obtained from *(i)*. □

5.4 On the Bound's Tightness in Special Cases

The bound shown in Corollary 5.11 for 1-neutralizing constructions is known to be tight (see [13]). In this section, we briefly discuss the tightness of the bound of Corollary 5.12 for 2-neutralizing 3-ary constructions.

For any quasigroup $(\mathbb{G}, \star)$, let $\mathbf{c}_{3,2} : \mathbb{G}^3 \to \mathbb{G}^2$ denote the 3-ary construction[6]

$$\mathbf{c}_{3,2}(x, y, z) := (x \star y, y \star z).$$

It is straightforward to verify that $\mathbf{c}_{3,2}$ is 2-neutralizing for the pairs $(\mathbf{X}, \mathbf{U})$, $(\mathbf{Y}, \mathbf{U})$, and $(\mathbf{Z}, \mathbf{U})$, where $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ are arbitrary weight-1 distributions over $\mathbb{G}$ and $\mathbf{U}$ is the uniform weight-1 distribution over $\mathbb{G}$. Thus, we have by Corollary 5.12

$$\begin{aligned}&\delta(\mathbf{c}_{3,2}(\mathbf{X}, \mathbf{Y}, \mathbf{Z}),\ \mathbf{U}^2)\\ &\qquad \le 2 \cdot (\delta(\mathbf{X})\delta(\mathbf{Y}) + \delta(\mathbf{X})\delta(\mathbf{Z}) + \delta(\mathbf{Y})\delta(\mathbf{Z})) - 3 \cdot \delta(\mathbf{X})\delta(\mathbf{Y})\delta(\mathbf{Z}),\end{aligned}$$

where $\mathbf{U}^2$ is the uniform distribution over $\mathbb{G}^2$ and $\delta(\cdot)$ is the statistical distance from the uniform distribution, i.e., $\delta(\cdot) := \delta(\cdot, \mathbf{U})$.

For the special case $\delta(\mathbf{X}) = \delta(\mathbf{Y}) = \delta(\mathbf{Z}) = \epsilon \in [0, 1]$ we obtain thus

$$\delta(\mathbf{c}_{3,2}(\mathbf{X}, \mathbf{Y}, \mathbf{Z}),\ \mathbf{U}^2) \le 6\epsilon^2 - 3\epsilon^3.$$

Proposition 5.16. *For any $c \in \mathbb{R}_{>0}$ and any $x, y, z \in [0, \frac{1}{c}]$*

(i) $|x - y| \le x + y - 2cxy$.

(ii) $|xy + yz - xz| \le xy + yz + xz - 2cxyz$.

The bounds are tight, as is seen by setting $x = y = z = \frac{1}{c}$.

Proof. *(i)* Observe that $|x - y| = x + y - 2\min(x, y)$ and as $cx \le 1$ and $cy \le 1$, we obtain $\min(x, y) \ge \min(cyx, cxy) = cxy$.

(ii) We have $|xy + yz - xz| = |xy + (y - x)z| \le xy + |y - x|z$ by the triangle inequality. Finally, *(i)* implies that $xy + |y - x|z \le xy + yz + xz - 2cxyz$.

□

Fact 5.17. *Assume $(\mathbb{G}, \star) = (\mathbb{Z}_n, \oplus_n)$ for even $n \ge 2$, and let for $\epsilon_{\mathbf{X}} \in [0, \frac{1}{n}]$*

$$\mathbf{X}(x) = \begin{cases} \frac{1}{n} + \epsilon_{\mathbf{X}} & \text{if } x = 0 \\ \frac{1}{n} - \epsilon_{\mathbf{X}} & \text{if } x = \frac{n}{2} \\ \frac{1}{n} & \text{otherwise} \end{cases}$$

[6] As in Section 5.1.1, one can think of an element of $\mathbb{G}^n$ as a single-query DDS with unary input alphabet $\{\diamond\}$ and output alphabet $\mathbb{G}^n$. It is easy to see that for any distribution $\mathbf{X}$ over such systems we have $[\mathbf{X}] = \{\mathbf{X}\}$ and thus for any pair $(\mathbf{X}, \mathbf{Y})$ of such weight-1 distributions $\widehat{\Delta}(\mathbf{X}, \mathbf{Y}) = \delta(\mathbf{X}, \mathbf{Y})$.

Moreover, let $\mathbf{Y}$ *and* $\mathbf{Z}$ *be defined analogously for* $\epsilon_{\mathbf{Y}}, \epsilon_{\mathbf{Z}} \in [0, \frac{1}{n}]$. *It is straightforward (though somewhat tedious) to show that*

$$\begin{aligned}\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathrm{U}^2) = \left(2-\frac{3}{n}\right)&(\delta(\mathbf{X})\delta(\mathbf{Y})+\delta(\mathbf{X})\delta(\mathbf{Z})+\delta(\mathbf{Y})\delta(\mathbf{Z}))\\ +\frac{1}{n}\cdot(&|\delta(\mathbf{X})\delta(\mathbf{Y})+\delta(\mathbf{X})\delta(\mathbf{Z})-\delta(\mathbf{Y})\delta(\mathbf{Z})|\\ &+|\delta(\mathbf{X})\delta(\mathbf{Y})+\delta(\mathbf{Y})\delta(\mathbf{Z})-\delta(\mathbf{X})\delta(\mathbf{Z})|\\ &+|\delta(\mathbf{X})\delta(\mathbf{Z})+\delta(\mathbf{Y})\delta(\mathbf{Z})-\delta(\mathbf{X})\delta(\mathbf{Y})|).\end{aligned}$$

If $\delta(\mathbf{X}) = \delta(\mathbf{Y}) = \delta(\mathbf{Z}) = \epsilon \in [0, \frac{1}{n}]$ *we thus obtain*

$$\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathbf{U}^2) = \left(1-\frac{1}{n}\right)6\epsilon^2,$$

and for $\epsilon = \frac{1}{n}$

$$\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathbf{U}^2) = 6\epsilon^2 - 6\epsilon^3.$$

The following bound is tight by Proposition 5.16 since $\{\delta(\mathbf{X}), \delta(\mathbf{Y}), \delta(\mathbf{Z})\} \subseteq [0, \frac{1}{n}]$.

$$\begin{aligned}&\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathbf{U}^2)\\ &\quad\le 2\cdot(\delta(\mathbf{X})\delta(\mathbf{Y})+\delta(\mathbf{X})\delta(\mathbf{Z})+\delta(\mathbf{Y})\delta(\mathbf{Z})) - 6\cdot\delta(\mathbf{X})\delta(\mathbf{Y})\delta(\mathbf{Z}).\end{aligned}$$

Fact 5.18. *Let* $(\mathbb{G}, \star)$ *be an arbitrary quasigroup with* $|\mathbb{G}| = n$. *If the support of* $\mathbf{X}$, $\mathbf{Y}$, *and* $\mathbf{Z}$ *is a singleton set, respectively, we have* $\delta(\mathbf{X}) \to 1$ *as* $n \to \infty$ *(and analogously for* $\mathbf{Y}$ *and* $\mathbf{Z}$*). As the support of* $\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z})$ *is a singleton set as well we have*

$$\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathbf{U}^2) \to 1 \text{ as } n \to \infty.$$

This shows that for any bound of the form

$$\begin{aligned}&\delta(\mathbf{c}_{3,2}(\mathbf{X},\mathbf{Y},\mathbf{Z}),\ \mathbf{U}^2) \le\\ &\quad d_1\cdot\delta(\mathbf{X})\delta(\mathbf{Y}) + d_2\cdot\delta(\mathbf{X})\delta(\mathbf{Z}) + d_3\cdot\delta(\mathbf{Y})\delta(\mathbf{Z}) + d_4\cdot\delta(\mathbf{X})\delta(\mathbf{Y})\delta(\mathbf{Z})\end{aligned}$$

for constants $d_i \in \mathbb{R}$ *we must have* $\sum_i d_i \ge 1$.

It is easy to see that the above examples imply the proved general bound

$$2\cdot(\delta(\mathbf{X})\delta(\mathbf{Y})+\delta(\mathbf{X})\delta(\mathbf{Z})+\delta(\mathbf{Y})\delta(\mathbf{Z})) - 3\cdot\delta(\mathbf{X})\delta(\mathbf{Y})\delta(\mathbf{Z})$$

being *almost* tight. There is, however, a tiny gap in the constant in front of the term $\delta(\mathbf{X})\delta(\mathbf{Y})\delta(\mathbf{Z})$ between the shown constructions and the proven bound. It is not obvious whether this gap can be closed by a more clever 2-neutralizing construction or by a better bound.

5.5 Conclusions

We have introduced $\mathcal{A}$-neutralizing constructions and proved a general indistinguishability amplification theorem (Theorem 5.9) that holds for any q-neutralizing construction. This generalizes the previously known Product Theorem of [13], which is essentially the same statement for the special case $q = 1$.

We conclude by discussing some open questions and future work.

- Is the bound proved in Theorem 5.9 tight for all q-neutralizing constructions or can it be further improved? Even if the bound is indeed tight, it is likely that a better bound can be shown under stronger assumptions. An interesting question is what natural assumptions can be made in order to achieve this.
- There are at least three orthogonal dimensions in which Theorem 5.9 can be further generalized:
  - The most straightforward generalization is to go from q-neutralizing constructions to arbitrary $\mathcal{A}$-neutralizing constructions. Note that the presented proof of Theorem 5.9 for q-neutralizing constructions relies heavily on the symmetry of such constructions. While the general technique presented in Lemma 5.8 is directly applicable to any $\mathcal{A}$-neutralizing construction[7], instantiating it in an optimal way will likely require an additional clever idea.
  - The only crucial assumption needed to achieve indistinguishability amplification as in Theorem 5.9 seems to be the neutralizing property of the construction. Thus, it is plausible that the same bound can be shown for a much more general type of objects (not only discrete systems).
  - Can the argument be extended to a computational setting? We expect the hardness amplification results shown in [11] to be useful to achieve this. Presumably, this will yield a generalization of the bounds shown in [15].

[7] For a k-ary $\mathcal{A}$-neutralizing construction with fixed *small* k it not difficult to find distributions $\mathbf{B}$ and $\mathbf{B}'$ such that a good bound is obtained.

Chapter 6

Conclusions

The main focus of this work was on finding the *right* view on systems and their properties which naturally appear in computer science and cryptography. This is motivated by the observation that with the right perspective, many statements of interest suddenly become very easy to understand and to prove. In contrast, a wrong perspective often results in overly complicated and virtually unformalizable arguments even for seemingly simple statements.

In Chapter 3, we have introduced abstract finite distributions as well as their intersection. We defined the (existential) observation compatibility of a pair $(\mathcal{A}, \mathcal{F})$, which, loosely speaking, states the following for all pairs $\mathbf{X}$ and $\mathbf{Y}$ of distributions over $\mathcal{A}$:

> The minimal overlap of $f(\mathbf{X})$ and $f(\mathbf{Y})$ over all functions $f \in \mathcal{F}$
> is equal to the maximal overlap of any two distributions
> $\mathbf{X}'$ and $\mathbf{Y}'$ that are equivalent to $\mathbf{X}$ and $\mathbf{Y}$, respectively.

Moreover, we have presented two natural ways to lift observation compatibility, i.e., to construct new observation-compatible pairs $(\mathcal{A}, \mathcal{F})$ from given observation-compatible pairs $(\mathcal{A}_i, \mathcal{F}_i)$. Said lifting methods describe a broad class of observation-compatible pairs.

Then, in Chapter 4, we have introduced an inductive representation of Maurer's theory of discrete systems. Within this theory, we gave new definitions of the distance of probabilistic discrete systems as well as of the maximum winning probability of certain types of probabilistic discrete games. The distinctive property of these new definitions is that they are *environment-less*, i.e., expressed as intrinsic properties of the objects themselves, free of an environment interacting with the objects. We then have shown that the environment-less definitions are actually *equivalent* to the corresponding classical ones, applying the results on observation compatibility.

In Chapter 5, we used the new environment-less distance of probabilistic discrete systems to prove a general indistinguishability amplification theorem for q-neutralizing constructions. The proof is more elementary than the one of the Product Theorem of [13], and at the same time we show a more general statement (for arbitrary $q \in \mathbb{N}_{\geq 1}$, as opposed to only $q = 1$).

Overall, the results provide great confidence in the environment-less paradigm. Yet it is in the nature of things that finding the right view is a continuous process. We therefore conclude by discussing ideas for future work.

- Even though the results on lifting observation compatibility already describe a fairly broad class of pairs with the property, we conjecture that a much stronger lifting lemma can be shown. Roughly speaking, we conjecture that lifting succeeds if whenever multiple components of an object can be observed at once, arbitrary adaptivity between the projections on the components is allowed.
- We discussed two properties in the environment-less paradigm: the distance of probabilistic discrete systems as well as the winning probability of probabilistic discrete games. While these are key properties in most cryptographic statements, a fully environment-less (cryptographic) systems theory needs many different variants of (similar) properties that have not been covered yet. For example, an interesting problem is to interpret indistinguishability under partially-adaptive or multi-execution distinguishers within this new paradigm.
- The treatment in this work is purely information-theoretic. Moreover, it does not take into account quantum systems. An interesting question is whether the results shown on observation compatibility as well as the environment-less properties can be extended to quantum systems or to computational models. We expect that our results do not carry over in their entirety. However, it is likely that our statements can either be properly generalized or extended by additional arguments to achieve this goal.
- As discussed in Section 4.6, the presented definitions of discrete systems capture the standard fully-adaptive single-execution semantics. While it is certainly desirable to generalize these semantics, we propose to go even further than that. Ideally, one would distill the essential properties of all such discrete systems into a concise set of *axioms*[1]. For example, a property similar to observation compatibility might simply be stated as an axiom about a set of abstract systems $\mathcal{S}$. This would allow to reason about such systems without fixing a specific representation of the objects. Especially in settings with intricate semantics, which usually

[1] We use the term *axiom* for non-logical axioms (as in ”ring axioms” and not as in ”axiom of choice”).

demand a more involved representation of the corresponding objects, such an abstract systems theory may lead to an even greater simplification than for the systems notion discussed in this work. Presumably, this kind of theory would represent a new level of abstraction that is positioned above the discrete system level and below the most general system level which only captures the composition of systems (see [14, 8]).

Bibliography

[1] D. Aldous. Random walks on finite groups and rapidly mixing markov chains. In J. Azéma and M. Yor, editors, *Séminaire de Probabilités XVII 1981/82*, pages 243–297, Berlin, Heidelberg, 1983. Springer Berlin Heidelberg.

[2] Z. Beerliová-Trubíniová and M. Hirt. Perfectly-secure mpc with linear communication complexity. In R. Canetti, editor, *Theory of Cryptography*, pages 213–230, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[3] R. Chuaqui. Cardinal algebras and measures invariant under equivalence relations. *Transactions of the American Mathematical Society*, 142:61–79, 1969.

[4] H. Dobbertin. On Vaught's criterion for isomorphisms of countable boolean algebras. *algebra universalis*, 15(1):95–114, Dec 1982.

[5] P. A. Grillet. Interpolation properties and tensor product of semigroups. *Semigroup Forum*, 1(1):162–168, Dec 1970.

[6] E. Kiltz and H. Wee. Quasi-adaptive NIZK for linear subspaces revisited. In E. Oswald and M. Fischlin, editors, *Advances in Cryptology - EUROCRYPT 2015*, pages 101–128, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[7] D. Maharam. The representation of abstract measure functions. *Transactions of the American Mathematical Society*, 65(2):279–330, 1949.

[8] C. Matt, U. Maurer, C. Portmann, R. Renner, and B. Tackmann. Toward an algebraic theory of systems, Nov. 2018.

[9] U. Maurer. Indistinguishability of random systems. In L. R. Knudsen, editor, *Advances in Cryptology — EUROCRYPT 2002*, pages 110–132, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[10] U. Maurer. Constructive cryptography – a new paradigm for security definitions and proofs. In S. Moedersheim and C. Palamidessi, editors, *Theory of Security and Applications (TOSCA 2011)*, volume 6993 of *Lecture Notes in Computer Science*, pages 33–56. Springer-Verlag, Apr. 2011.

[11] U. Maurer. An information-theoretic approach to hardness amplification. In *2017 IEEE International Symposium on Information Theory (ISIT)*, June 2017.

[12] U. Maurer. Lecture notes cryptography foundations, February 2019.

[13] U. Maurer, K. Pietrzak, and R. Renner. Indistinguishability amplification. In A. Menezes, editor, *Advances in Cryptology - CRYPTO 2007*, pages 130–149, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

[14] U. Maurer and R. Renner. Abstract cryptography. In B. Chazelle, editor, *The Second Symposium on Innovations in Computer Science, ICS 2011*, pages 1–21. Tsinghua University Press, Jan. 2011.

[15] U. Maurer and S. Tessaro. Computational indistinguishability amplification: Tight product theorems for system composition. In S. Halevi, editor, *Advances in Cryptology — CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 350–368. Springer-Verlag, Aug. 2009.

[16] A. Tarski and B. Jónsson. *Cardinal Algebras: With an appendix: Cardinal products of isomorphism types.* Oxford University Press, 1949.