UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE E TECNOLOGIE

CORSO DI LAUREA MAGISTRALE IN MATEMATICA

TESI DI LAUREA MAGISTRALE

Information Cohomology and Entropy

Relatore:
Professor Luca Barbieri Viale

Candidato:
Luca Castiglioni
Matricola:
967292

Anno Accademico 2022/2023

Contents

Introduction

Shannon entropy first appeared in Claude Shannon's article "A mathematical theory of communication", as a measure of the uncertainty associated with a probability distribution on a finite set. Although it was introduced to provide an answer to an information theoretic problem, namely to show that it is possible to send information through a channel at a positive rate with arbitrarily small probability of error, it has subsequently found applications in various mathematical disciplines, including statistical mechanics, probability theory, and portfolio theory [4]. More recently, D.Bennequin and P.Baudot [2] showed that entropy appears spontaneously as a cohomological class in *information cohomology*, which is an invariant associated with a finite statistical system. Finally, J.P. Vigneaux, in his doctoral dissertation, extended these ideas in several directions, in particular he introduced the notion of *information structure*, which is a category-theoretical formalization of the mentioned statistical systems. The aim of this thesis is to present this new point of view, mainly following Vigneaux's article [13].
As a preliminary step, we have included a first section containing an introduction to Shannon entropy. In this section we also define the functions S_α, called α-entropies,

where $\alpha \in (0, +\infty) \setminus \{1\}$. These functions can be considered a generalization of entropy, which we denote S_1 for a reason that will be made clear in section 1.
The second section is dedicated to introducing the category of information structures. We begin by providing a precise definition of objects (Def. 2.2.) and morphisms (Def. 2.6). After that, we prove the existence of products and coproducts.
Roughly speaking, information structures are pairs $(\mathbf{S}, \mathscr{E})$, where: $\mathbf{S}$ is a small category (actually is a partially ordered set) whose set of objects represents a set of observables, and arrows encode the refinement relations between them; $\mathscr{E}$ is a functor on $\mathbf{S}$, that associated with each observable the set of possible values it can assume. The definition of information structure is sufficiently general and flexible to treat in a unified manner several different cases, some of which are presented as examples. The fundamental example remains the one in which observables are random variables that share the same sample space. However, we will show that not all information structures are of this type, providing a characterization of this behavior that relates to the phenomena of contextuality treated in [1] through a sheaf-theoretical approach.
In the third section, we construct the framework necessary to define information cohomology as a derived functor. For this purpose, we define a presheaf of $\mathbb{R}$-algebras on the category $\mathbf{S}$, which can therefore be seen as a ringed site with the trivial topology. Within the abelian category of presheaves of $\mathscr{A}$-modules, we consider $\mathbb{R}_{\mathbf{S}}$ to be the constant presheaf with value $\mathbb{R}$, equipped with the trivial action of $\mathscr{A}$. Then, we construct the unnormalized bar resolution [8] of $\mathbb{R}_{\mathbf{S}}$, which turns out to be a projective resolution $\beta_{\bullet}(\mathbb{R}_{\mathbf{S}})$. Finally, we define the information cohomology associated with $\mathbf{S}$, with coefficients in a presheaf of $\mathscr{A}$-modules $\mathcal{F}$ as the cohomology of the complex $\mathrm{Hom}_{\mathscr{A}}(\beta_{\bullet}(\mathbb{R}_{\mathbf{S}}), \mathcal{F})$.
In the fourth and final section, we relate information cohomology and α-entropies. To this end, we introduce on each information structure $(\mathbf{S}, \mathscr{E})$ a class of covariant functors (*probability functors*) which associate each observable X with a subset of the probability laws on $\mathscr{E}_X$. Given a probability functor $\mathscr{Q}$, we consider the abelian presheaf which maps each observable X to the abelian group of the real-valued measurable functions on $\mathscr{Q}_X$. For each $\alpha > 0$ we define an $\mathscr{A}$-module structure, dependent on α, on this presheaf, obtaining in this way a familiy $\{\mathscr{F}_\alpha\}_{\alpha>0}$ of presheaves of $\mathscr{A}$-modules.
Once these definitions are provided, we show that $H^{\bullet}(\mathbf{S}, \mathscr{F}_\alpha)$, the information cohomology with coefficients in $\mathscr{F}_\alpha$, is "functorial" in $\mathbf{S}$ (Proposition 4.2.). It follows that information cohomology is an invariant for equivalent structures.
We then proceed to analyze the structure of the cocycles of the complex $\mathrm{Hom}_{\mathscr{A}}(\beta_*(\mathbb{R}_{\mathbf{S}}), \mathscr{F}_\alpha)$. In particular, it turns out that the 1-cocycles are uniquely determined by collections $(f[X])_{X \in Ob(\mathbf{S})}$, of elements in $\mathscr{F}_\alpha(X)$ i.e. functions of probabilities distributions on $\mathscr{E}_X$. These collections must satisfy the cocycle condition, expressed by $f[XY] = f[X] + X.f[Y]$, where $X.$ indicates the action of $\mathscr{A}_{XY}$. In this framework, for each $\alpha > 0$ the function S_α yields a 1-cocycle (Definition 4.5.) of the corresponding chain complex, since the cocycle condition coincides with the chain rule for α-entropies. This same fact enables us to

demonstrate the main result of this thesis, which we now present in detail.

Theorem 0.1. *Let* $(\mathbf{S}, \mathscr{E})$ *be a finite information structure equipped with an adapted probability functor* $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$ *whose restriction maps (internal marginalizations) are all surjective. Assume that for any object* $X \in Ob(\mathbf{S})$*, exists an arrow* $Z \to X$ *in* $\mathbf{S}$ *such that* Z *is non-trivially reducible. Then, for each* $\alpha > 0$*, there is an isomorphism of real vector spaces*

$$\begin{gathered}\chi_\alpha : \prod_{C \in \pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R} \longrightarrow Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \quad \textit{given by} \\ \chi((\lambda_C)_{C \in \pi_0(\mathbf{S}^*_{\mathscr{Q}})})[X] = \lambda_C S_\alpha[X] \quad \forall C \in \pi_0(\mathbf{S}^*), \forall\, X \in C\end{gathered} \tag{0.1}$$

Under this isomorphism, the subgroup of 1*-coboundaries is identified with the diagonal subspace* $\Delta = \mathbb{R} \cdot (1, \dots, 1, \dots)$ *if* $\alpha \neq 1$*, otherwise is the zero subspace. Therefore,*

$$H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) = \begin{cases} \prod_{\pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R} & \textit{if} \;\; \alpha = 1 \\ \left(\prod_{\pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R}\right) / \Delta & \textit{if} \;\; \alpha \neq 1 \end{cases} \tag{0.2}$$

Let us now explain notations and concepts we have not encountered so far. For instance, $\mathbf{S}^*_{\mathscr{Q}}$ denotes the subcategory of $\mathbf{S}$ consisting of objects X such that $\mathscr{Q}(X)$ contains a non-atomic probability. Moreover $\pi_0(\mathbf{S}^*_{\mathscr{Q}})$ is the set of the connected components of the category $\mathbf{S}^*_{\mathscr{Q}}$. Finally, an observable Z is non-trivially reducible if there exist two distinct observables X and Y such that $Z = XY$, and some conditions (Definition 4.7) on the functors $\mathscr{E}$ and $\mathscr{Q}$ are satisfied. These conditions allow us to use the functional equation (4.53) to show the existence of a real constant λ such that $f[X] = \lambda S_\alpha[X]$, $f[Y] = \lambda S_\alpha[Y]$ and $f[XY] = \lambda S_\alpha[XY]$.
This theorem proves that every 1-cocycle is locally a real multiple of the cocycle defined by entropy. This means that the conditions required for being a 1-cocycle, are enough to derive the form of the functions S_α. In this sense we can affirm that entropy appears spontaneously in this theory.

1 Entropy

Consider a random variable $X : (\Omega, \mathfrak{B}, p) \to \mathscr{E}_X$[1], where $\mathscr{E}_X$ is a finite set. Let P be the probability distribution of X, which means that

$$P(x) = p(X = x) := p(X^{-1}(\{x\})) \qquad \forall\, x \in \mathscr{E}_X. \tag{1.1}$$

[1] $(\Omega, \mathfrak{B}, p)$ is a probability space. Hence Ω is the sample space, which is merely a non-empty set; $\mathfrak{B} \subset \wp(\Omega)$ is the σ-algebra of events; $p : \mathfrak{B} \to [0, 1]$ is a probability measure

The *Shannon entropy* of the random variable X is a real number $S_1(X)$[2] that quantifies how uncertain we are about the outcome of X before it is revealed. Its value is computed by

$$S_1(X) := -\sum_{x\in\mathscr{E}_X} P(x)\log_2 P(x) \tag{1.2}$$

with the convention $0\log_2 0 := 0$, justified by the fact that $\lim_{t\to 0} t\log_2 t = 0$.
The following two observations are intended to provide some insights into the role of entropy as a measure of uncertainty. First notice that if the distribution of X is an atomic probability, i.e. there exists an element $\bar{x} \in \mathscr{E}_X$ such that $P(\bar{x}) = 1$, then the outcome of X is known in advance, there is no uncertainty about it. The entropy takes this into account, in fact from equation (1.2) we obtain $S_1(X) = 0$. On the other hand, is also true that if a random variable (taking values in a finite set) has zero entropy, then its distribution must be atomic, because all the terms in the sum that defines the entropy (1.2) have the same sign.
We then observe that, if the cardinality of the set of possible events is fixed, the maximum entropy is reached in correspondence with uniformly distributed random variables. Is possible to prove this fact looking for the maximum of the function

$$\begin{gathered}(p_1, p_2, \dots, p_n) \mapsto -\sum_{i=1}^{n} p_i \log_2 p_i \quad \text{constrained to the set} \\ \Delta^{n-1} = \left\{ (p_1, p_2, \dots, p_n) \in \mathbb{R}^n \middle| \sum_{i=1}^{n} c_i = 1 \text{ and } p_i \geq 0 \;\forall\, i : 1, \dots, n \right\} \subset \mathbb{R}^n,\end{gathered} \tag{1.3}$$

where $n \in \mathbb{N}$ is the number of possible events. To make sense of how this argument works, note that the points of Δ^{n-1} are in bijection with the probability distributions on a set of n elements. Anyway, the case of uniformly distributed random variables corresponds also to maximum uncertainty about what the outcome will be, since the probability distribution does not give us any clues about it.

Remark 1.1. The definition of entropy we gave above can be extended, with no modifications, to discrete variables with values in an infinite set. However, the sum involved could become a series that may not converge. An example of this behavior is given by a random variable X taking values in $\mathbb{N} \setminus \{1\}$, and distributed according to

$$p(X = n) := c\frac{1}{n\log^2 n} \qquad n \in \mathbb{N} \setminus \{1\}, \quad \frac{1}{c} = \sum_{n>1} \frac{1}{n\log^2 n}. \tag{1.4}$$

Is also possible to go further and define the entropy of continuous random variables as in [4]. However, in the present thesis, only finite set valued random variables are involved.

[2]This notation was chosen for the sake of continuity with the article [13]. The subscript "1" is justified by the existence of the α-entropies for any positive real number α, which will be introduced later.

For this reason, from now on, whenever we consider a random variable, it will be implied that it can assume only a finite number of values.

As can be seen from (1.2), the entropy of a random variable depends only on its distribution. Thus the sample space Ω is not involved in its computation. Moreover, composing a random variable X with a bijection yields another random variable with the same distribution. Hence, $S_1(X)$ does not depend on "the names" of elements in $\mathscr{E}_X$, but on their number. In light of these remarks, entropy can be viewed, perhaps more appropriately, as a function of probability laws defined on finite sets. In order to make this precise, we pick a probability law P on a finite set $\mathcal{X}$. Since the distribution of the random variable $id_{\mathcal{X}} : (\mathcal{X}, \wp(\mathcal{X}), P) \to \mathcal{X}$ is precisely P, we define

$$S_1(P) := S_1(id_{\mathcal{X}}) = -\sum_{z \in \mathcal{X}} P(z) \log_2 P(z). \tag{1.5}$$

We now introduce an important property of entropy, usually called *chain rule*, which will also turn out to be fundamental within this thesis, since it allows us to see entropy as a cocycle of a suitable chain complex.

Let us fix a sample space $(\Omega, \mathfrak{B}, p)$ for a moment. Let $X : \Omega \to \mathscr{E}_X$ and $Y : \Omega \to \mathscr{E}_Y$ be random variables. We can always consider the random vector

$$\begin{aligned} (X, Y) : \Omega &\to \mathscr{E}_X \times \mathscr{E}_Y \\ \omega &\to (X(\omega), Y(\omega)) \,. \end{aligned} \tag{1.6}$$

Denoted with P the joint distibution of (X, Y), it is known that for any $y \in \mathscr{E}_Y$ such that $p(Y = y) \neq 0$, is defined on $\mathscr{E}_X$ the conditional law $P_{|Y=y}$ given by

$$P_{|Y=y}(x) := \frac{P(x, y)}{P(y)} = \frac{p(\{X = x\} \cap \{Y = y\})}{p(Y = y)} \qquad x \in \mathscr{E}_X. \tag{1.7}$$

This assertion clearly holds true even when the roles of X and Y are exchanged. The entropy of the random vector $S_1(X, Y)$ (computed with the joint distribution) is called *joint entropy*, while the quantity

$$S_1(X|Y) := \sum_{\substack{y \in \mathscr{E}_y \\ p(Y=y) \neq 0}} P(y) S_1(P_{|Y=y}) = \sum_{(x,y) \in \mathscr{E}_X \times \mathscr{E}_Y} P(x, y) \log_2 P_{Y=y}(x) \tag{1.8}$$

is known as *conditional entropy* [4]. This latter is a measure of how much uncertainty remains about the outcome of the variable X after the variable Y has been measured, averaged over the possible outcomes of Y. The joint entropy and the conditional entropy are related by the chain rule, which can be formulated as:

$$S_1(X) + S_1(Y|X) = S_1(X, Y) = S_1(Y) + S_1(X|Y). \tag{1.9}$$

This property sounds quite natural for a measure of uncertainty, since it essentially states that the uncertainty of a random vector (X, Y) is equal to the uncertainty of the outcome of Y plus the remaining uncertainty about the outcome of X once the value of Y has been revealed; and that the roles of X and Y can be exchanged. We will show later that the chain rule is a characterizing property of entropy, but first, let us reintroduce entropy in a more formal and structured way, even if it may be less evocative at first glance.
We already encountered the standard simplex Δ^{n-1} in (1.3) and remarked its relation with probabilities. Since we have restricted our focus to random variables with values in finite sets, we can identify entropy with the function

$$\begin{aligned} S_1 : \bigsqcup_{n\in\mathbb{N}} \Delta^n &\longrightarrow \mathbb{R}_{\geq 0} \\ P = (p_1, p_2, \dots, p_n) &\longmapsto S_1(P) := -\sum_{i=1}^{n} p_i \log_2 p_i. \end{aligned} \tag{1.10}$$

The chain rule can be restated as follows. For any $n \in \mathbb{N}$, for all $P \in \Delta^{n-1}$ and any partition $C = \{C_1, \dots, C_k\}$ of the set $\{1, 2, \dots, n\}$, we define the function

$$\begin{aligned} \pi_C : \{1, 2, \dots, n\} &\longrightarrow C \\ j &\longmapsto C_i \ s.t. \ j \in C_i. \end{aligned} \tag{1.11}$$

Then we can consider the marginalized law $P\pi_C^{-1}$ on $C \simeq \{1, 2, \dots, k\}$ and the conditional laws $P_{|C_i}$ for each $i : 1, 2, \dots, k$ such that $P(C_i) \neq 0$, which are given by

$$\begin{aligned} P\pi_C^{-1}(\{i'\}) &:= P(\pi_C^{-1}(\{i'\})) = P(C_{i'}) = \sum_{j\in C_{i'}} p_j \qquad \forall\, i' \in \{1, 2, \dots, k\}, \\ P_{|C_i}(\{j\}) &:= \frac{P(\{j\} \cap C_i)}{P(C_i)} \qquad \forall\, j \in \{1, 2, \dots, n\}. \end{aligned} \tag{1.12}$$

With these notations, it holds that

$$S_1(P) = S_1(P\pi_C^{-1}) + \sum_{\substack{C_i\in C \\ P(C_i)\neq 0}} P(C_i) S_1(P_{|C_i}).\ ^3 \tag{1.13}$$

This equation can be readily verified substituting the S_1's with their explicit expressions

[3] From now on, to lighten the notation, we will write this sum as $\sum_{i=1}^{k} P(C_i) S_1(P_{|C_i})$, while keeping in mind that the indices for which $P(C_i) = 0$ should not be considered.

and carrying out the calculations:

$$\begin{aligned}
&-\sum_{i=1}^{k} P(C_i)\log_2 P(C_i) - \sum_{i=1}^{k} P(C_i) \sum_{j=1}^{n} \frac{P(\{j\}\cap C_i)}{P(C_i)} \log_2\left(\frac{P(\{j\}\cap C_i)}{P(C_i)}\right) \\
&= -\sum_{i=1}^{k} P(C_i)\log_2 P(C_i) - \sum_{i=1}^{k}\sum_{j=1}^{n} P(\{j\}\cap C_i)\log_2\left(\frac{P(\{j\}\cap C_i)}{P(C_i)}\right) \\
&= -\sum_{i=1}^{k} P(C_i)\log_2 P(C_i) - \sum_{j=1}^{n} p_i \log_2 p_i + \sum_{i=1}^{k}\underbrace{\sum_{j=1}^{n} P(\{j\}\cap C_i)}_{P(C_i)}\log_2 P(C_i) = S_1(P).
\end{aligned} \tag{1.14}$$

The equation (1.13) states that the entropy associated with the choice of a random number between 1 and n is equal to the entropy associated with a first choice of a subset among the elements of a partition of $\{1, 2, \ldots, n\}$, summed with the entropy of choosing the number, knowing the subset in which it is located. From this we can see a strong analogy with the chain rule stated by means of random variables. We claim that actually the two formulations are equivalent. In fact, the equation (1.9) can be recovered by (1.13) letting $n := |\mathscr{E}_X \times \mathscr{E}_Y|$ and $C := \{\pi^{-1}_{\mathscr{E}_X}(x)\}_{x\in\mathscr{E}_X}$[4] or $C := \{\pi^{-1}_{\mathscr{E}_Y}(y)\}_{y\in\mathscr{E}_Y}$. Vice versa, choosing $(\Omega, \mathfrak{B}, p) := (\{1, 2, \ldots, n\}, \wp\{1, 2, \ldots, n\}, P)$, $X := Id_\Omega$ and $Y := \pi_C$ the equation (1.9) yields (1.13).

Let us now explain how entropy is characterized by the chain rule, as was previously mentioned. We remark that C.Shannon, in his foundational paper [10], derived the expression for the entropy of a probability distribution (1.10) by imposing on a function $H : \bigsqcup_{n\in\mathbb{N}} \Delta^n \to \mathbb{R}_{\geq 0}$ some constraints that encode properties considered natural for a measure of uncertainty. In particular, he required: that H be continuous in all its arguments; that the function $n \mapsto H(\frac{1}{n}, \ldots, \frac{1}{n})$ be monotonically increasing in $n \in \mathbb{N}$; and that H satisfies the chain rule. However, these properties are not necessary to characterize entropy, but it is enough to assume that H is measurable and satisfies the chain rule. As a matter of fact, the chain rule for H (1.13), when applied on the partitions $\{1, 2\} \sqcup \{3\} = \{1, 2, 3\} = \{1, 3\} \sqcup \{2\}$, implies[5] that the function $p \mapsto H(p, 1-p)$ satisfies the functional equation

$$u(1-x) + (1-x)u\left(\frac{y}{1-x}\right) = u(y) + (1-y)u\left(\frac{1-x-y}{1-y}\right) \tag{1.15}$$

for all $(x, y) \in [0, 1)^2$ such that $x + y \in [0, 1]$. In the subsection 4.3 we will prove,

[4] We have to choose a bijection $\{1, 2, \ldots, n\} \simeq \mathscr{E}_X \times \mathscr{E}_Y$ so that $\{\pi^{-1}_{\mathscr{E}_X}(x)\}_{x\in\mathscr{E}_X}$ can be viewed as a partition of $\{1, 2, \ldots, n\}$.

[5] this argument is presented in detail in Example 4.2 with the cocycle condition in place of the chain rule. However it will be clear that the two conditions are equivalent.

following [12], [7] and [3], that every measurable solution of this functional equation is a real multiple of the map $p \mapsto S_1(p, 1-p)$. This implies that H is equal to kS_1, for some real number k, at least on binary probability laws. But we can prove, by induction on the number of arguments of H, that the function H is completely determined by the values it assumes on binary probabilities. The inductive step is based on the equation

$$H(p_1, p_2, \dots, p_n) = H(p_1, 1-p_1) + (1-p_1)H(\frac{p_2}{1-p_1}, \dots, \frac{p_n}{1-p_1}) \qquad \forall (p_1, \dots, p_n) \in \Delta^{n-1} \tag{1.16}$$

which is obtained applying the chain rule on the partition $\{1, 2, \dots, n\} = \{1\} \sqcup \{2, \dots, n\}$. In this way, for any probability distribution P, we arrive to an expression of $H(P)$ involving only terms like $H(p, 1-p)$ for some $p \in [0, 1]$. Then, by substituting every occurrence of $H(p, 1-p)$ with $kS_1(p, 1-p)$ in such an expression, we get $H(P) = S_1(P)$.
Finally, we present a generalization of Shannon entropy. Let $\alpha \in (0, +\infty) \setminus \{1\}$, the function

$$\begin{aligned} S_\alpha : \bigsqcup_{n \in \mathbb{N}} \Delta^n &\longrightarrow \mathbb{R}_{\geq 0} \\ P = (p_1, p_2, \dots, p_n) &\longmapsto S_\alpha(P) := \frac{1}{(1-\alpha)\ln 2} \left(\sum_{i=1}^n p_i^\alpha - 1 \right). \end{aligned} \tag{1.17}$$

is called structural α-entropy, or Tsallis α-entropy, from the name of the physicist who first used it in the field of statistical mechanics. The family of functions $\{S_\alpha\}_\alpha$ generalizes the entropy S_1 in the sense that $\lim_{\alpha \to 1} S_\alpha = S_1$. Indeed, for all $n \in \mathbb{N}$, for all probability law $(p_1, \dots, p_n) \in \Delta^{n-1}$, we have

$$\begin{aligned} \lim_{\alpha \to 1} \frac{1}{(1-\alpha)\ln 2} \left(\sum_{i=1}^n p_i^\alpha - 1 \right) &= \lim_{\alpha \to 1} \frac{-1}{\ln 2} \sum_{i=1}^n \frac{p_i(p_i^{\alpha-1} - 1)}{\alpha - 1} \\ &= -\sum_{i=1}^n p_i \frac{\ln p_i}{\ln 2} = S_1(p_1, \dots, p_n). \end{aligned} \tag{1.18}$$

We chose to introduce these functions because they admit a cohomological interpretation too. Actually, the theory developed in Section 4 does not make much distinctions between the case $\alpha = 1$ and the case $\alpha \neq 1$. Just like S_1, also all the functions S_α vanish on the atomic probabilities, while they have a maximum, once the number of arguments is fixed, in correspondence with the uniform distributions. However, if $\alpha > 1$, the function

$$n \mapsto S_\alpha(\frac{1}{n}, \dots, \frac{1}{n}) = \frac{1}{1-\alpha} \left(\frac{n}{n^\alpha} - 1 \right) \qquad n \in \mathbb{N} \tag{1.19}$$

is not monotonically increasing, but it is indeed decreasing. Therefore, the functions S_α cannot properly be seen as measures of the grade of uncertainty of probability laws.

Moreover, for any $\alpha \in (0, +\infty)$, the function S_α does not satisfy the chain rule, but rather a modified version of it.

$$\begin{gathered}\forall\, n \in \mathbb{N},\, \forall\, P \in \Delta^{n-1},\ \forall\, C = \{C_1, \dots, C_k\} \text{ partition of } \{1, \dots, n\} \\ S_\alpha(P) = S_\alpha(P\pi_C^{-1}) + \sum_{i=1}^{k} P(C_i)^\alpha S_\alpha(P_{|C_i}),\end{gathered} \tag{1.20}$$

Similarly to what we have seen for the chain rule, we can express this property in an equivalent way using the language of random variables: following notations of (1.9), it holds that

$$\begin{aligned} S_\alpha(X, Y) &= S_\alpha(Y) + \sum_{y \in \mathscr{E}_Y} P(y)^\alpha S_\alpha(P_{|Y=y}) \\ &= S_\alpha(X) + \sum_{x \in \mathscr{E}_X} P(x)^\alpha S_\alpha(P_{|X=x}). \end{aligned} \tag{1.21}$$

The proof consists, even in this case, in a quick verification: let P be the joint probability distribution, then we have

$$\begin{gathered} \sum_{(x,y) \in \mathscr{E}_X \times \mathscr{E}_Y} P(x, y)^\alpha - 1 = \sum_{x \in \mathscr{E}_X} P(x)^\alpha \left(\sum_{y \in \mathscr{E}_Y} P_{|X=x}(y)^\alpha \right) - 1 = \\ \left(\sum_{x \in \mathscr{E}_X} P(x)^\alpha \left(\sum_{y \in \mathscr{E}_Y} P_{|X=x}(y)^\alpha - 1 \right) + P(x)^\alpha \right) - 1 = \\ \sum_{x \in \mathscr{E}_X} P(x)^\alpha \left(\sum_{y \in \mathscr{E}_Y} P_{|X=x}(y)^\alpha - 1 \right) + \sum_{x \in \mathscr{E}_X} P(x)^\alpha - 1. \end{gathered} \tag{1.22}$$

Hereafter, we will refer to this "deformed" version of the chain rule simply as *chain rule for α-entropies.*
We conclude observing that the argument which led to equation (1.15) can be repeated in the case $\alpha \neq 1$, applying the chain rule for α-entropies in place of the standard chain rule. This results in the functional equation

$$\begin{gathered} \forall\, (x, y) \in [0, 1)^2 \ \text{ such that } \ x + y \in [0, 1] \\ u(1 - x) + (1 - x)^\alpha u\left(\frac{y}{1 - x}\right) = u(y) + (1 - y)^\alpha u\left(\frac{1 - x - y}{1 - y}\right) \end{gathered} \tag{1.23}$$

In subsection 4.3 we prove, following [3], that every solution of this functional equation is a real multiple of the function $p \mapsto S_\alpha(p, 1 - p)$. Therefore, also the α-entropies are characterized by the chain rule.

2 The category of Information Structures

2.1 Definitions and examples

Definition 2.1. A *conditional meet semilattice* is a poset[6] $\mathbf{S}$ satisfying the additional property that for any two elements $x, y \in \mathbf{S}$, if there exists a third element $z \in \mathbf{S}$ such that $x \leq z$ and $y \leq z$, then the meet $x \wedge y$ exists in $\mathbf{S}$. We will call a conditional meet semilattice *unital*, if it admits a maximum $\top$.

Recall that any poset $\mathbf{S}$ can be equivalently viewed as a small category, called again $\mathbf{S}$, in this way:

- $Ob(\mathbf{S})$ is the set of elements of the poset $\mathbf{S}$;
- for all $x, y \in \mathbf{S}$, $\mathrm{Hom}_{\mathbf{S}}(x, y) := \begin{cases} \{*\} & if\ x \leq y \\ \emptyset & if\ x \nleq y. \end{cases}$

In this description of a poset, the maximum corresponds to the terminal object and the meet corresponds to the product.

Definition 2.2. An *information structure* is the datum of an unital conditional meet semilattice $\mathbf{S}$ equipped with a functor $\mathscr{M} : \mathbf{S} \longrightarrow \mathbf{Meas}$[7] (denote $\mathscr{M}(X) := \big(\mathscr{E}_X, \mathfrak{B}_X\big)$ for each $X \in Ob(\mathbf{S})$) that satisfies :

1. $\mathscr{E}_\top = \{*\}$, with the trivial σ -algebra;
2. $\forall X \in \mathbf{S}$, $\forall x \in \mathscr{E}_X$, the singleton $\{x\}$ is an element of the σ-algebra $\mathfrak{B}_X$;
3. for any arrow in $\mathbf{S}$, say f, the restriction morphism $\mathscr{M}(\mathrm{f})$ is surjective;
4. for every diagram $X \xleftarrow{\pi_X} X \times Y \xrightarrow{\pi_Y} Y$,the canonical measurable map

$$\begin{aligned} \iota : \mathscr{M}_{X\times Y} &\longrightarrow \mathscr{M}_X \times \mathscr{M}_Y \\ z &\longmapsto (\mathscr{M}(\pi_X)(Z), \mathscr{M}(\pi_Y)(z)) \end{aligned} \tag{2.1}$$

   is a monomorphism.

In the introduction, we have already indicated how the category $\mathbf{S}$ and the functor $\mathscr{M}$ should be interpreted. We only note that, since we have not imposed any restrictions on the cardinality of the sets of possible outcomes, it is necessary that they be equipped with a measurable space structure in order to introduce probability measures. We can then add that, given two observables X and Y, the product $XY := X \wedge Y$, if it exists,

[6] A partially ordered set.
[7] The category of measurable spaces and measurable functions between them.

represents the observable quantity obtained by measuring X and Y together. Therefore, the absence of certain products encodes the impossibility of performing some joint measurements. However, it is possible to conduct the joint measurement of X and Y if there exists an observable Z which determines both the outcomes of X and Y. In fact, Z also determines the outcome of the joint measurement. Finally, the terminal object represents the observable whose value is known with certainty before measuring it.

Definition 2.3. An information structure is said *finite* if $\forall X \in S$, $\mathscr{E}_X$ is a finite set, while it is *bounded* if the poset $\mathbf{S}$ has finite height.[8]

Observe that if $(\mathbf{S}, \mathscr{M})$ is a finite information structure, for each observable $X \in Ob(\mathbf{S})$, the σ-algebra $\mathfrak{B}_X$ coincide with $\wp(\mathscr{E}_X)$ because of condition 2. in Definition 2.2. Therefore, the structure of measurable space on $\mathscr{E}_X$ can be forgotten without losing any information. For these reason we will denote $(\mathbf{S}, \mathscr{M}) := (\mathbf{S}, \mathscr{E})$ in these cases, where $\mathscr{E} : \mathbf{S} \to \mathbf{Sets}$ is the functor obtained composing $\mathscr{M}$ with the forgetful functor $U : \mathbf{Meas} \to \mathbf{Sets}$.

Example 2.1. Rather than being a single example, this is a family of examples.
Let I be a finite set. Consider the abstract simplex $\Delta(I) = \wp(I)$. It can be seen as a poset assigning

$$\forall\, A, B \subset I \quad A \leq B \iff B \subset A. \tag{2.2}$$

Note that $A \wedge B = A \cup B$, hence $\Delta(I)$ admits all finite products. We call $\mathbf{K} \subseteq \Delta(I)$ a subsimplicial complex of $\Delta(I)$ if it is a full subcategory, and for any given element $C \in \mathbf{K}$ all its subset are elements of $\mathbf{K}$ too. Given a collection of measurable spaces $\{(\mathscr{E}_i, \mathfrak{B}_i)\}_{i \in I}$ consider the functor

$$\mathscr{M} : \Delta(I) \longrightarrow \mathbf{Meas}\,, \quad \begin{array}{ccc} A & \xrightarrow{\mathscr{M}} & \left(\prod_{i\in A} \mathscr{E}_i, \bigotimes_{i \in A} \mathfrak{B}_i\right) \\ \big\downarrow{\scriptstyle A \supseteq B} & & \big\downarrow{\scriptstyle projection} \\ B & \xrightarrow{\mathscr{M}} & \left(\prod_{i\in B} \mathscr{E}_i, \bigotimes_{i \in B} \mathfrak{B}_i\right) \end{array}$$

where $\bigotimes_{i\in A} \mathfrak{B}_i$ denotes the product σ-algebra, i.e. the smallest σ-algebra on $\prod_{i\in A} \mathscr{E}_i$ which makes all the canonical projections measurable. By restricting the functor $\mathscr{M}$ to $\mathbf{K}$, we obtain an information structure: indeed it is a poset admitting all meets and with a maximum, given by the empty set. Moreover the condition 1. of Definition 2.2. is clearly verified; for 2. we need to assume that for all $i \in \mathbf{K}$, $\mathfrak{B}_i$ contains all the singletons of $\mathscr{E}_i$; 3. is also clear because the canonical projections are all surjective; as for the condition 4. we just note that the map $\iota : \mathscr{E}_{A \cup B} \to \mathscr{E}_A \times \mathscr{E}_B$ is an isomorphism. The structures arising in this way are called *simplicial information structures.*

Example 2.2. Another fundamental family of examples is constituted by *concrete information structures.* Consider a set Ω, and denote $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ the set of all its finite

[8]The height of a poset is the cardinality of its longest chain.

partitions. Each random variable X defined on Ω with values in a finite set yields a finite partition of Ω, given by $\{X^{-1}(x_1), \dots, X^{-1}(x_m)\}$, where $\mathscr{E}_X = \{x_1, \dots, x_m\}$ is the target of X. Moreover, given any bijection $h : \mathscr{E}_X \xrightarrow{\sim} \mathcal{E}$, the random variable $h \circ X$ induces the same partition as X. Hence, the set of finite partition of Ω can be identified with the set of all random variables on Ω with values in a finite set, modulo the relation $\sim$ such that, for any two random variables $X : \Omega \to \mathscr{E}_X$ and $Y : \Omega \to \mathscr{E}_Y$, we have $X \sim Y$ iff there exists a bijection $h : \mathscr{E}_X \to \mathscr{E}_Y$ such that $Y = h \circ X$.
For this reason, we will use the same notation for a random variable and the induced finite partition. Consider X, Y as above, we say that X refines Y if is possible to write any element of Y as a disjoint union of elements of X (as partitions) or equivalently if $\sigma(Y) \subseteq \sigma(X)$[9] (as random variables). This defines a partial order relation on $\mathbf{Obs}_{\text{fin}}(\Omega)$, such that $X \leq Y$ iff X refines Y. In the resulting poset, the trivial partition $\{\Omega\}$ is the top element. The product between two partition is the coarsest partition that refines both. Explicitly, for any two partitions/random variables X, Y, their product is the finite partition $XY = \{x \cap y | x \in x, y \in Y\}$, which can be seen also as the partition induced by the random vector (X, Y). In order to construct an information structure, we consider the functor

$$\begin{aligned} \diamond : \mathbf{Obs}_{\text{fin}}(\Omega) &\longrightarrow \mathbf{Sets} \\ X &\longmapsto X \text{ seen as a finite set,} \end{aligned} \tag{2.3}$$

and whenever there is an arrow $X \xrightarrow{\text{f}} Y$, $\diamond(\text{f})$ is the function which maps any element of X to the unique element of Y which contains it. This functor clearly satisfies the conditions 1. and 3.; the condition 2. does not require any verification since the structure is finite; for 4. just note that

$$\begin{aligned} \iota : \diamond XY &\longrightarrow \diamond X \times \diamond Y \quad \text{is injective.} \\ x \cap y &\longmapsto (x, y) \end{aligned} \tag{2.4}$$

We remark that in Baudot and Bennequin's article [2], an information structure is defined to be a full subcategory $\mathbf{S} \subset \mathbf{Obs}_{\text{fin}}(\Omega)$ that is itself a unital conditional meet semilattice. This means that $\{\Omega\} = \top \in \mathbf{S}$ and that for any $X, Y \in \mathbf{S}$, if exists a partition Z that refines both X and Y, then XY belong to $\mathbf{S}$. Observe that restricting the functor $\diamond$ to $\mathbf{S}$ yields an information structure according to Definition 2.2.

Definition 2.4. Let Ω be a set. A concrete information structure on Ω is a pair $(\mathbf{S}, \diamond)$, where

- $\mathbf{S}$ is a full subcategory of $\mathbf{Obs}_{\text{fin}}(\Omega)$ such that $\{\Omega\} = \top \in \mathbf{S}$ and, for any $X, Y \in \mathbf{S}$, if exists a partition Z that refines both X and Y, then XY belong to $\mathbf{S}$

[9]For any random variable $X : (\Omega, \mathfrak{F}, P) \to (\mathscr{E}_X, \mathfrak{B}_X)$, $\sigma(X) \subset \mathfrak{F}$ is the sigma-algebra generated by the sets like $X^{-1}(B)$ for $B \in \mathfrak{B}_{\mathfrak{X}}$

- $\diamond$ is the restriction to $\mathbf{S}$ of the functor defined in (2.3).

Remark 2.1. Let Ω be a set. Consider $\Sigma := \{X_1, X_2, \dots, X_n\} \in \mathbf{Obs}_{\text{fin}}(\Omega)$.
There is a unique product preserving functor $\sigma : \Delta(n) \longrightarrow \mathbf{Obs}_{\text{fin}}(\Omega)$[10] such that $\sigma(i) = X_i$ for all $i : 1, \dots, n$. In fact we can write any $J \subseteq \{1, \dots, n\}$ as a union of singletons $J = \bigcup_{m \in J} \{m\}$. But since the union coincides the product in $\Delta(n)$, it must hold $\sigma(J) = \prod_{m \in J} X_m$. Moreover, the action of σ on arrows is determined by the universal property of the product. The image of σ is a full subcategory of $\mathbf{Obs}_{\text{fin}}(\Omega)$. It is also a conditional meet semilattice because it admits all the products. Hence it gives rise, together with the restriction of $\diamond$, to an information structure $(\mathbf{S}(\Sigma), \diamond)$.
If we try to repeat the same construction with a subcomplex $\mathbf{K} \subset \Delta(n)$ in place of $\Delta(n)$, some problems may occur. As an example, let $\Omega = \{0,1\} \times \{0,1\}$; let X_1 and X_2 be the partitions induced by the projections on the first and second component respectively ,while $X_3 = \{(0,0)\} \sqcup \{(0,1),(1,0),(1,1)\}$. Let $\mathbf{K}$ be the subcomplex of $\Delta(3)$ whose maximal faces are $\{1,2\}$ and $\{3\}$.

$$\begin{array}{ccccc} & & \{1\} & & \\ & \nearrow & & \searrow & \\ \{1,2\} & \longrightarrow & \{2\} & \longrightarrow & \emptyset \\ & & & \nearrow & \\ & & \{3\} & & \end{array}$$

Observe that $\sigma(\{1,2\}) = X_1 X_2$ is the atomic partition, which is initial and in particular refines X_3. Nevertheless the product $X_1 X_3$ does not belong to the image of σ, which therefore is not a conditional meet semilattice.

Definition 2.5. A morphism of conditional meet semilattices is a functor $\phi : \mathbf{S}_1 \to \mathbf{S}_2$ that preserves all finite products in $\mathbf{S}_1$ i.e.

$$\phi(X \times Y) = \phi(X) \times \phi(Y) \quad \forall X, Y \in Ob(\mathbf{S}_1) \tag{2.5}$$

Definition 2.6. A morphism of information structures is a pair $(\phi, \hat{\phi}) : (\mathbf{S}_1, \mathscr{M}_1) \to (\mathbf{S}_2, \mathscr{M}_2)$, where ϕ is a morphism of conditional meet semilattices, and $\hat{\phi} : \mathscr{M}_1 \Rightarrow \mathscr{M}_2 \circ \phi$ is a natural transformation.

$$\begin{array}{ccc} \mathbf{S}_1 & \xrightarrow{\ \phi\ } & \mathbf{S}_2 \\ & \searrow^{\mathscr{M}_1} \ \overset{\hat{\phi}}{\nearrow\!\!\!\nearrow} & \downarrow \mathscr{M}_2 \\ & & \mathbf{Meas} \end{array} \tag{2.6}$$

Given $(\psi, \hat{\psi}) : (\mathbf{S}_2, \mathscr{M}_2) \longrightarrow (\mathbf{S}_3, \mathscr{M}_3)$, the composition $(\psi, \hat{\psi}) \circ (\phi, \hat{\phi})$ is defined as the pair $\left(\psi \circ \phi, (\text{id}_\phi \times \hat{\psi}) \circ \hat{\phi}\right)$ [11]. More explicitly, we are considering the natural transformation

[10] $\Delta(n) := \wp(\{1, \dots, n\}) = \Delta(\{1, \dots, n\})$

[11] the notation $\text{id}_\phi \times \hat{\psi}$ indicates the horizontal composition of natural transformations

whose X component is given by the composition

$$\mathscr{M}_1(X) \xrightarrow{\hat{\phi}_X} \mathscr{M}_2(\phi(X)) \xrightarrow{\hat{\psi}_{\phi(X)}} \mathscr{M}_3\left(\psi\left(\phi(X)\right)\right) \tag{2.7}$$

Finally, for every information structure, the identity map is given by $(\mathrm{id}_{\mathbf{S}}, \mathrm{id}_{\mathscr{M}})$.

The fact that $(\mathrm{id}_{\mathbf{S}}, \mathrm{id}_{\mathscr{M}})$ is actually a neutral element for the composition defined above is clear. That composition is associative is a consequence of the properties of horizontal composition, but can be seen also as follows. Recall that we have a functor

$$\begin{array}{ccc} (\_\circ\phi) : Fun(\mathbf{S}_2,\mathbf{Meas}) & \to & Fun(\mathbf{S}_1,\mathbf{Meas}), \\ \mathscr{M} & \longmapsto & \mathscr{M}\circ\phi \\ \downarrow \eta & & \downarrow (\_\circ\phi)(\eta) \\ \mathscr{N} & \longmapsto & \mathscr{N}\circ\phi \end{array} \tag{2.8}$$

also denoted with ϕ^*.[12] The natural transformation $\phi^*(\eta)$ is defined on components by $\phi^*(\eta)_X := \eta_{\phi(X)}$ for X in $\mathbf{S}_1$, so it coincides with $\mathrm{id}_\phi \times \hat{\psi}$. Now for any sequence of morphisms

$$\mathbf{S}_1 \xrightarrow{(\phi,\hat{\phi})} \mathbf{S}_2 \xrightarrow{(\psi,\hat{\psi})} \mathbf{S}_3 \xrightarrow{(\xi,\hat{\xi})} \mathbf{S}_4$$

The equality

$$\underbrace{(\psi\phi)^*(\hat{\xi})}_{\phi^*\psi^*(\hat{\xi})} \circ (\phi^*(\hat{\psi})) \circ \hat{\phi} = \phi^*(\psi^*(\hat{\xi}) \circ \hat{\psi}) \circ \hat{\phi}$$

holds true, and associativity follows. This implies that information structures form a category, which we denote **InfoStr**.

Once the category is built, we proceed studying its limits and colimits. It turns out that

Proposition 2.1. InfoStr *admits a zero object, the countable products and all coproducts.*

Proof. The zero object is the pair made up of: the category $\mathbf{1}$, which has only one object $\top$ and only one arrow, or equivalently the poset that contains only the terminal object; the functor $\mathscr{M}_0$ that sends $\top$ to $(\{*\}, \{\emptyset, \{*\}\})$ which is terminal in **Meas**. This pair satisfies trivially all properties required to be an information structure. Moreover, for any information structure $(\mathbf{S}', \mathscr{M}')$, there is a unique morphism of information structures $(\mathbf{S}', \mathscr{M}) \to \mathbf{1}$, which sends all objects of $\mathbf{S}'$ to $\top$. On the other hand, a morphism of information structures $\mathbf{1} \to (\mathbf{S}', \mathscr{M})$ in forced to map $\top \to \top$, and then in unique.

Coproducts: let $\mathcal{I}$ be a set and $\{(\mathbf{S}_i, \mathscr{M}_i)\}_{i\in I}$ be a family of information structures. Consider the category $\mathbf{S}$ such that:

$$Ob(\mathbf{S}) = \bigsqcup_{i\in\mathcal{I}} Ob(\mathbf{S}_i)/\sim \tag{2.9}$$

[12]this functor takes the name ϕ_p between categories of presheaves

where $\sim$ is the equivalence relation that identifies all terminal objects. Denote $\top$ the equivalence class of them.

$$\forall\, X, Y \in Ob(\mathbf{S}) \quad \mathrm{Hom}_{\mathbf{S}}(X,Y) := \begin{cases} \mathrm{Hom}_{\mathbf{S}_i}(X,Y) & if\ X, Y \in \mathbf{S}_i \ or\ Y = \top \\ \emptyset & if\ X \in \mathbf{S}_i,\ Y \in \mathbf{S}_j,\ i \neq j \end{cases} \tag{2.10}$$

This category is a conditional meet semilattice because all $\mathbf{S}_i$ are. Then we define a functor $\mathscr{M} : \mathbf{S} \to \mathbf{Meas}$ by

$$\mathscr{M}(X) = \begin{cases} \mathscr{M}_i(X) & \text{if } X \in \mathbf{S}_i \\ \big(\{*\}, \{\emptyset, \{*\}\}\big) & \text{if } X = T \end{cases} \qquad X \in Ob(\mathbf{S}). \tag{2.11}$$

The pair $(\mathbf{S}, \mathscr{M})$ is an information structure because the axioms are verified locally on each $\mathbf{S}_i$. Furthermore for any index i, we have a morphism of posets $j_i : \mathbf{S}_i \to \mathbf{S}$ which maps any elements to itself, but seen in the disjoint union. In order to obtain morphisms in **InfoStr**, we define natural transformations $\hat{j}_i$ by :

$$\hat{j}_i(X_i) := \mathrm{id}_{\mathscr{M}_i(X_i)} \quad \forall\, i \in \mathcal{I}, \quad \forall X_i \in \mathbf{S}_i \tag{2.12}$$

Now we are left to prove that $(\mathbf{S}, \mathscr{M})$ has the universal property of the coproduct. To achieve this, consider a cocone $\{(\psi_i, \hat{\psi}_i) : (\mathbf{S}_i, \mathscr{M}_i) \to (\mathbf{R}, \mathscr{N})\}_{i \in \mathcal{I}}$. A morphism $(\psi, \hat{\psi})$ making all the triangles

$$\begin{array}{ccc} (\mathbf{S}, \mathscr{M}) & \xrightarrow{\ \psi\ } & (\mathbf{R}, \mathscr{N}) \\ {\scriptstyle j_i}\uparrow & \nearrow_{\psi_i} & \\ (\mathbf{S}_i, \mathscr{M}_i) & & \end{array} \tag{2.13}$$

commutative, is completely determined. Indeed, for any $i \in \mathcal{I}$ and any $Y_i \in Ob(\mathbf{S}_i)$, $\psi(Y_i)$ must be equal to $\psi_i(Y_i)$. The same holds for the maps $\hat{\psi}_{Y_i}$ and $\hat{\psi}_{iY_i}$ since $\hat{j}_i$ is the identity map.

Countable products: let $\mathcal{I}$ be a countable set and $\{(\mathbf{S}_i, \mathscr{M}_i)\}_{i \in I}$ a family of information structures. Define $\mathbf{S}$ to be the product in **Cat**[13] of the categories $\mathbf{S}_i$. This means that $Ob(\mathbf{S}) := \prod_{i \in I} Ob(\mathbf{S}_i)$ and

$$\forall\, (X_i)_{i\in I}, (Y_i)_{i \in I} \in Ob(\mathbf{S}) \quad \mathrm{Hom}_{\mathbf{S}}\left((X_i)_{i \in I}, (Y_i)_{i \in I}\right) := \prod_{i \in I} \mathrm{Hom}_{\mathbf{S}_i}(X_i, Y_i). \tag{2.14}$$

[13] The category of all small categories.

We endow $\mathbf{S}$ with the functor

$$\begin{aligned} \mathscr{M} : \mathbf{S} &\longrightarrow \mathbf{Meas} \\ (X_i)_{i\in I} &\longmapsto \prod_{i\in I} \mathscr{M}_i(X_i) = (\prod_{i\in I} \mathscr{E}_i(X_i), \bigotimes_{i\in I} \mathfrak{B}_{i,X_i}) \end{aligned} \tag{2.15}$$

such that given $(\mathrm{f}_i)_{i\in I} : (X_i)_{i\in I} \to (Y_i)_{i\in I}$, the measurable function $\mathscr{M}((\mathrm{f}_i)_{i\in I}) := \prod_{i\in I} \mathrm{f}_i$ is induced by the universal property of the product and is clearly surjective. The pair $(\mathbf{S}, \mathscr{M})$ just defined satisfies the conditions 1. ... 4. of Definition 2.2:
1. Follows once noting that the terminal object in $\mathbf{S}$ is the collection of terminal objects of the $\mathbf{S}_i$.
2. Since I is countable, all the singletons $\{(x_i)_{i\in I}\} \in \mathscr{E}((X_i)_{i\in I})$ can be written as a countable intersection of elements in the product σ-algebra. Denoted $\pi_{\mathscr{E}(X_j)} : \prod_{i\in I} \mathscr{E}(X_i) \to \mathscr{E}(X_j)$ the canonical projection,

$$\{(x_i)_{i\in I}\} = \bigcap_{i\in I} \pi^{-1}_{\mathscr{E}(X_i)}(\{x_i\}). \tag{2.16}$$

Thus it is itself an element of the product σ-algebra.
3. Already verified
4. For any two objects $(X_i)_{i\in I}$,$(Y_i)_{i\in I}$ in $\mathbf{S}$, the measurable map

$$\iota : \prod_{i\in I} \mathscr{M}_i(X_i \times Y_i) \longrightarrow \prod_{i\in I} \mathscr{M}_i(X_i) \times \prod_{i\in I} \mathscr{M}_i(Y_i) \tag{2.17}$$

is the product of the injective maps $\iota_i : \mathscr{M}(X_i \times Y_i) \to \mathscr{M}(X_i) \times \mathscr{M}(Y_i)$ and hence it is injective too.
For any $i \in I$, the projection $p_i : \mathbf{S} \to \mathbf{S}_i$ (from the product in **Cat**) becomes a morphism of information structures $(p_i, \hat{p}_i)$ defining

$$\hat{p}_{i(X_i)_i} : (\prod_{i\in I} \mathscr{E}_i(X_i), \bigotimes_{i\in I} \mathfrak{B}_{i,X_i}) \longrightarrow (\mathscr{E}_i, \mathfrak{B}_{i,X_i}) \qquad (X_i)_{i\in I} \in \mathbf{S} \tag{2.18}$$

as the projection from the product in **Meas**. We claim that $(\mathbf{S}, \mathscr{M})$ together with the morphisms $(p_i, \hat{p}_i)$, has the universal property of the product in **InfoStr**. To prove this, consider an I-cone $\{(t_i, \hat{t}_i : (\mathbf{R}, \mathscr{N}) \to (\mathbf{S}_i, \mathscr{M}_i)\}$ on the discrete diagram $\{(\mathbf{S}_i, \mathscr{M}_i)\}_{i\in I}$. Since $\mathbf{S}$ is the product in **Cat** of the $\mathbf{S}_i$'s, there is a unique functor

$$\langle t_i \rangle_{i\in I} : \mathbf{R} \to \mathbf{S} \;\text{ such that }\; p_i \circ \langle t_i \rangle_{i\in I} = t_i \quad \forall\, i \in I. \tag{2.19}$$

Furthermore, for any object U in $\mathbf{R}$, there exists only one measurable map making the diagrams

$$
\begin{array}{ccc}
\mathscr{N}(U) & \overset{\exists!\,\langle \hat{t}_i(U)\rangle}{\dashrightarrow} & \mathscr{M}\left((t_i(U))_{i\in I}\right) = \prod_{i\in I} \mathscr{M}_i(t_i(U)) \\
& \searrow^{\hat{t}_i(U)} & \big\downarrow \hat{p}_i(U) \\
& & \mathscr{M}_i(t_i(U))
\end{array}
\tag{2.20}
$$

commutative for all $i \in I$. Thus, $\langle \hat{t_i} \rangle_{i\in I}$ is also uniquely determined to be the natural transformation whose component in U is $\langle \hat{t}_i(U) \rangle$. □

2.2 Representations

Throughout this subsection we consider only finite information structures. Removing the reference to a fixed sample space lent a more natural definition of morphisms between information structures. Nevertheless, there are information structures whose objects can be seen as partitions of some fixed set, although they are not necessarily concrete in the sense of Definition 2.4. This is made precise by the following

Definition 2.7. Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure. A classical representation of $(\mathbf{S}, \mathscr{E})$ in a set Ω is a morphism of information structures $(\rho, \hat{\rho}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{Obs}_{\text{fin}}(\Omega), \diamond)$ such that $\hat{\rho}_X$ is a bijection for any $X \in \mathbf{S}$.

Using the bijectivity of $\hat{\rho}$, we can associate to any $X \in Ob(\mathbf{S})$ a simple random variable $\tilde{X} : \Omega \to \mathscr{E}_X$ defined by $\tilde{X}^{-1}(x) := \hat{\rho}_X(x)$ for all $x \in \mathscr{E}_X$. It is possible to characterize the structures that admit a representation in some space by means of the properties of the functor $\mathscr{E}$. This is actually what we are going to do now.
The functor $\mathscr{E} : \mathbf{S} \to \mathbf{Set}$ is a diagram in **Sets**, which is a complete category. Hence the limit exist, and is explicitly described by:

$$
\lim_{\mathbf{S}} \mathscr{E} = \left\{ (t_X)_{X\in Ob(\mathbf{S})} \in \prod_{X\in Ob(\mathbf{S})} \mathscr{E}_X \mid \text{for any arrow } \varphi : X \to Y,\ \mathscr{E}(\varphi)(t_X) = (t_Y) \right\} \tag{2.21}
$$

The functions that constitute the universal cone are the restrictions to $\lim_{\mathbf{S}} \mathscr{E}$ of the projections from the product $\pi_{\mathscr{E}_Z} : \prod_{X\in Ob(\mathbf{S})} \mathscr{E}_X \to \mathscr{E}_Z$, where $Z \in Ob(\mathbf{S})$. In fact, for any other cone $(D, \{d_X\}_{X\in Ob(\mathbf{S})})$ over the diagram $\mathscr{E}$, we have a unique function $\langle d_X \rangle_{X\in Ob(\mathbf{S})}$ induced by the universal property of the product. But since for any arrow $\mathrm{f} : X \to Y$ in $\mathbf{S}$ we have $d_Y = \mathrm{f} \circ d_X$ (being D a cone), the image of the function $\langle d_X \rangle_{X\in Ob(\mathbf{S})}$ is contained in $\lim_{\mathbf{S}} \mathscr{E} \subset \prod_{X\in Ob(\mathbf{S})} \mathscr{E}_X$. Thus, there exists a unique map $d : D \to \lim_{\mathbf{S}} \mathscr{E}$ such that $d_Z = \pi_{\mathscr{E}_Z} \circ d$ for all $Z \in Ob(\mathbf{S})$.
We remind that the elements of $\lim_{\mathbf{S}} \mathscr{E}$ are usually called compatible families, or global sections of $\mathscr{E}$.

Definition 2.8. An information structure $(\mathbf{S}, \mathscr{E})$ is called noncontextual if for all $X \in Ob(\mathbf{S})$ and for all $x \in \mathscr{E}_X$, there exists $\bar{s} \in \lim_{\mathbf{S}} \mathscr{E}$ such that $\pi_{\mathscr{E}_X}(\bar{s}) = x$.

We can reformulate this definition saying that for all $X \in Ob(\mathbf{S})$, every element in $\mathscr{E}_X$ belongs to a compatible family.

Theorem 2.1. *Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure. It holds that*

$$(\mathbf{S}, \mathscr{E}) \textit{ has a classical representation} \quad \Longleftrightarrow \quad (\mathbf{S}, \mathscr{E}) \textit{ is noncontextual.} \tag{2.22}$$

Proof. $\Rightarrow$) Suppose that $(\rho, \hat{\rho})$ is a classical representation of $(\mathbf{S}, \mathscr{E})$ in some set Ω. For any map $\mathrm{f} : X \to Y$ in $\mathbf{S}$ we have the commutative diagram

$$\begin{array}{ccccc} \Omega & \xrightarrow{\gamma_X} & \diamond\rho(X) & \underset{\hat{\rho}_X^{-1}}{\xrightarrow{\simeq}} & \mathscr{E}_X \\ & \searrow^{\gamma_Y} & \downarrow{\scriptstyle \diamond\rho(f)} & & \downarrow{\scriptstyle \mathscr{E}(\mathrm{f})} \\ & & \diamond\rho(Y) & \underset{\hat{\rho}_Y^{-1}}{\xrightarrow{\simeq}} & \mathscr{E}_Y \end{array}$$

where γ_X maps $\omega \in \Omega$ to the element of $\diamond\rho(X)$ to which it belongs. We can see that Ω is the vertex of a cone over the diagram $\mathscr{E}$. Therefore exists a unique map $\Psi : \Omega \to \lim_{\mathbf{S}} \mathscr{E}$ such that $\forall\, X \in \mathbf{S},\ \pi_{\mathscr{E}_X} \circ \Psi = \hat{\rho}_X^{-1} \circ \gamma_X$. Consider an element $y \in \mathscr{E}_Y$, for some $Y \in Ob(\mathbf{S})$. Is always possible to take an element ω_y in the preimage of y via $\hat{\rho}_Y^{-1} \circ \gamma_Y$, because this map is surjective. We claim that $\Psi(\omega_y)$ is the compatible family we were looking for. In fact $\pi_{\mathscr{E}_Y}(\Psi(\omega_y)) = \hat{\rho}_Y^{-1} \circ \gamma_Y(\omega_y) = y$.
$\Leftarrow$) We define a classical representation of $\mathbf{S}$ in $\lim_{\mathbf{S}} \mathscr{E}$. Let $\rho : \mathbf{S} \to \mathbf{Obs}_{\mathrm{fin}}(\lim_{\mathbf{S}} \mathscr{E})$ be the functor that associates to and object X the partition induced by $\pi_{\mathscr{E}_X}$, which is $\{\pi_{\mathscr{E}_X}^{-1}(x) | x \in \mathscr{E}_X\}$. Given an arrow $\mathrm{f} : X \to Y$, we have the comutative triangle

$$\begin{array}{ccc} \lim_{\mathbf{S}} \mathscr{E} & & \\ \downarrow{\scriptstyle \pi_{\mathscr{E}_X}} & \searrow^{\pi_{\mathscr{E}_Y}} & \\ \mathscr{E}_X & \underset{\mathscr{E}(\mathrm{f})}{\longrightarrow} & \mathscr{E}_Y . \end{array} \tag{2.23}$$

From this we deduce

$$\pi_{\mathscr{E}_Y}^{-1}(y) = \bigcup_{x \in \mathscr{E}(\mathrm{f})^{-1}(\{y\})} \pi_{\mathscr{E}_X}^{-1}(x) \tag{2.24}$$

which means that $\rho(X)$ refines $\rho(Y)$. We define $\rho(\mathrm{f})$ as this refinement. To obtain a morphism in **InfoStr**, consider the natural transformation $\hat{\rho} : \mathscr{E} \to \diamond \circ \rho$ given by

$$\begin{array}{ccc} \mathscr{E}_X & \xrightarrow{\hat{\rho}_X} & \diamond\rho(X) \\ \downarrow{\scriptstyle \mathscr{E}(\mathrm{f})} & & \downarrow{\scriptstyle \diamond\rho(\mathrm{f})} \\ \mathscr{E}_Y & \xrightarrow{\hat{\rho}_Y} & \diamond\rho(Y) \end{array} \qquad \begin{array}{ccc} x & \longrightarrow & \pi_{\mathscr{E}_X}^{-1}(x) \\ \downarrow & & \downarrow \\ \mathscr{E}\mathrm{f}(x) & \longrightarrow & \pi_{\mathscr{E}_Y}^{-1}(\mathscr{E}\mathrm{f}(x)) \end{array} \tag{2.25}$$

Observe that $\hat{\rho}_X$ is surjective by definition of $\rho(X)$, and is also injective. In fact for

$x, z \in \mathscr{E}_X$ such that $\pi^{-1}_{\mathscr{E}_X}(x) = \pi^{-1}_{\mathscr{E}_X}(z)$, applying $\pi_{\mathscr{E}_X}$, we get $x = z$. Thus $(\rho, \hat{\rho})$ is a classical representation of $(\mathbf{S}, \mathscr{E})$. □

Remark 2.2. A *measurement scenario* is defined in [1] as a triple $(\mathcal{X}, \mathcal{M}, O)$, where:

- $\mathcal{X}$ is a finite set of variables.
- $\mathcal{M} := \{C_i\}_{i \in I}$ is a family of subsets of $\mathcal{X}$, that represent the maximal possible measurement contexts.
- O is a finite set which represents the set of possible values or outcomes that each of the variables in $\mathcal{X}$ can assume.

The power set $\wp(\mathcal{X})$ can be ordered by inclusion, obtaining in this way a poset which in turn can be seen as a category. It is worth noting that this category is the opposite of the category $\Delta(\mathcal{X})$ defined in Example 2.1.
On the category $\wp(\mathcal{X})$ is defined the presheaf $\mathcal{E} : U \mapsto O^U = \prod_{x \in U} O$, with restriction morphisms $\mathcal{E}(U \subset U')$ given by the projections $\prod_{x \in U'} O \to \prod_{x \in U} O$. This presheaf is called *the sheaf of events*: it is indeed a sheaf if we see $\mathcal{X}$ as a discrete topological space. Furthermore, for any $C \in \mathcal{M}$, suppose given a probability law p_C on the set $\mathcal{E}(C)$, such that the resulting family $(p_C)_{C \in \mathcal{M}}$ is compatible, in the sense that for any $C, C' \in \mathcal{M}$ the marginalizations of p_C and $p_{C'}$ to $\mathcal{E}(C \cap C')$ are equal. This defines a subpresheaf $\mathcal{S}$ of $\mathscr{E}$ that can be interpreted as the presheaf of events which are actually possible according to the family $(p_C)_{C \in \mathcal{M}}$. Specifically

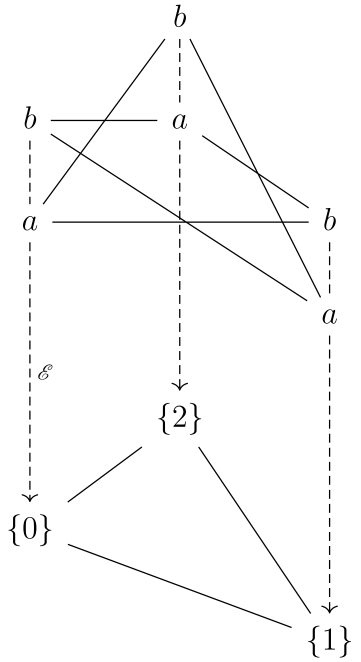
$$\mathcal{S}(U) := \left\{ s \in O^U \,|\, \forall\, C \in \mathcal{M}, s|_{U \cap C} \in supp(p_C|_{U \cap C}) \right\}, \tag{2.26}$$

where $s|_{U \cap C}$ denotes the projection of s in $O^{U \cap C}$ and $p_C|_{U \cap C}$ denotes the marginalization. Let $\mathbf{K}$ be the subcomplex of $\Delta(\mathcal{X})$ whose maximal faces are the measurement context $C \in \mathcal{M}$. It is possible to prove that $(\mathbf{K}, \mathcal{S})$ is an information structure according to Definition 2.2. We then follow the notation of [1] saying that an finite information structure $(\mathbf{S}, \mathscr{E})$ is:

- *logically contextual* at a value $x \in \mathscr{E}_X$ if x belongs to no compatible families in $\lim_{\mathbf{S}} \mathscr{E}$
- *strongly contextual* if $\mathscr{E}$ does not admit any global section, i.e. $\lim_{\mathbf{S}} \mathscr{E} = \emptyset$.

We conclude the discussion with some examples.

Example 2.3. Let $\mathbf{K}$ be the subcomplex of $\Delta(2)$ whose minimal faces are: $\{0, 1\}, \{0, 2\}, \{1, 2\}$. We want to build a simplicial information structure on $\mathbf{K}$. For this purpose, we introduce a functor $\mathscr{E} : \mathbf{K} \to \mathbf{Sets}$ defining it directly on all objects and arrows in $\mathbf{K}$. We start setting $\mathscr{E}(\{i\}) := \{a, b\}$ for $i = 0, 1, 2$. Note that a and b are just "names" for the elements of a two-elements set. Usually these two elements are denoted 0 and 1 but we choose to not adopt this notations here to avoid confusion with the elements $0, 1 \in \Delta(2)$. Since the functor $\mathscr{E}$ must satisfy the condition 4. in the definition of information structure, the set $\mathscr{E}(\{i,j\})$ can be identified with a subset of $\mathscr{E}(\{i\}) \times \mathscr{E}(\{j\})$, thus it makes sense to define $\mathscr{E}(\{i,j\}) := \{(a,b),(b,a)\}$ for any $\{i,j\} \in \mathbf{K}$. Moreover, $\mathscr{E}$ transforms arrows in $\mathbf{K}$ (reversed inclusions) into restrictions of the canonical projections from the product in **Sets**.

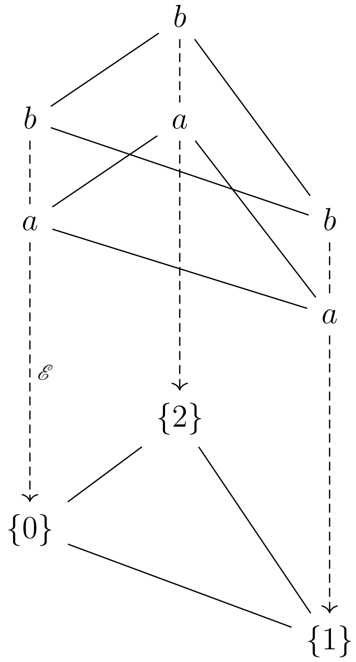


(2.27)

The triangle at the ground floor is a picture of $\mathbf{K}$, the side with vertices $\{i\}$ and $\{j\}$ represents the object $\{i,j\}$. Above each vertex $\{i\}$ of this triangle lies $\mathscr{E}(\{i\})$, and above each side $\{i,j\}$ there is $\mathscr{E}(\{i,j\})$: every segment that connects an element of $\mathscr{E}(\{i\})$ with an element of $\mathscr{E}(\{j\})$ specifies a pair in $\mathscr{E}(\{i\}) \times \mathscr{E}(\{j\})$. We can now observe that an element $s \in \prod_{X \in \mathbf{K}} \mathscr{E}_X$ corresponds to a family of points an dashes indexed by objects of $\mathbf{K}$. This family forms a closed path if and only if $s \in \lim_{\mathbf{S}} \mathscr{E}$. Looking at the picture we deduce that in this case $\lim_{\mathbf{S}} \mathscr{E} = \emptyset$, so the structure is *strongly contextual.*
Alternatively, one could notice that the shape of $\mathscr{E}$ models the constraint that for any measurement involving two observables together, they yield different outcomes. Thus, no joint measurement of the three observables can satisfy the constraint, in fact, it's impossible for three elements taken from $\{a, b\}$ to be pairwise different. Observing that an element in the limit specifies a possible outcome of such a joint measurement, we conclude that $\lim_{\mathbf{S}} \mathscr{E} = \emptyset$.

Example 2.4. Here is a variation of the previous example. We only modify the value $\mathscr{E}$ on the two-element objects.
Specifically, $\mathscr{E}(\{i,j\}) := \{(a,a),(b,b)\}$ for any $\{i,j\} \subset \{0,1,2\}$. In this setting the picture

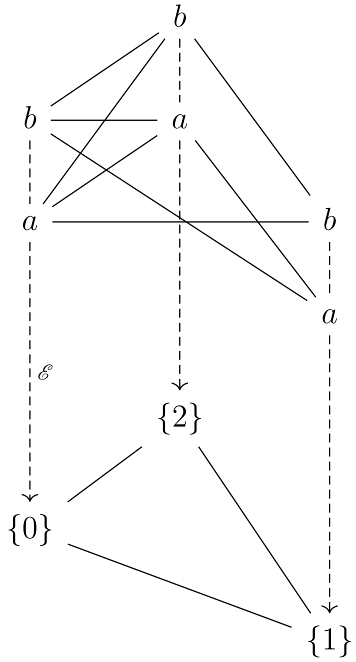
becomes



(2.28)

We see that any point or dash is part of a closed path. Hence the structure is noncontextual. Moreover $\lim_{\mathbf{S}} \mathscr{E} \simeq \{a, b\}$ since any component of an element of the limit determines all others. In this case the three observables are constrained to be identical.

Example 2.5. We present another variation. Again we modify only the definition of $\mathscr{E}$ on objects, according to: $\mathscr{E}(\{0,2\}) := \{a,b\} \times \{a,b\}$; $\mathscr{E}(\{0,1\}) := \{(a,b),(b,a)\}$; $\mathscr{E}(\{1,2\}) := \{(a,a),(b,b)\}$.



(2.29)

In this case $\lim_{\mathbf{S}} \mathscr{E} \neq \emptyset$, indeed $\big(a, b, b, (a, b), (a, b), (b, b)\big) \in \lim_s \mathscr{E}$. However there is no closed path that includes the segment (b, b) over $\{0, 2\}$.
Thus the structure is *logically contextual at* $(b, b) \in \mathscr{E}(\{0, 2\})$ but not *strongly contextual.*

Example 2.6. Given a concrete structure $(\mathbf{S}, \diamond)$ on a set Ω, it may happen that $\lim_{\mathbf{S}} \diamond \neq \Omega$. Recall that there exists a map $\psi : \Omega \to \lim_{\mathbf{S}} \diamond$ induced by universal property of the limit. Hence any element in Ω specifies an element in $\lim_{\mathbf{S}} \diamond$.
Let $\Omega := \{1, 2, 3, 4\}$ and consider the finite partitions partitions $X_i = \big\{\{i\}, \{i\}^c\big\}$ for $i = 1, 2, 3, 4$. Let $\mathbf{S}$ be the concrete structure whose underlying poset is represented by

$$\begin{array}{ccccccc} & X_1X_2 & & & & X_2X_3 & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ X_1 & & & X_2 & & & X_3 \\ & \searrow & & \downarrow & & \swarrow & \\ & & & T & & & \end{array} \tag{2.30}$$

Applying the functor $\diamond$ we get

$$\begin{array}{ccccccc} & \big\{\{1\}, \{2\}, \{3, 4\}\big\} & & & & \big\{\{2\}, \{3\}, \{1, 4\}\big\} & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ \big\{\{1\}, \{1\}^c\big\} & & & \big\{\{2\}, \{2\}^c\big\} & & & \big\{\{3\}, \{3\}^c\big\} \\ & \searrow & & \downarrow & & \swarrow & \\ & & & \big\{\{1, 2, 3, 4\}\big\} & & & \end{array} \tag{2.31}$$

Observe that $(\{1\}, \{3\}, \{1\}, \{2\}^c, \{3\}, \{1, 2, 3, 4\}) \in \lim_{\mathbf{S}} \diamond$ but this compatible family can not be the image of any $\omega \in \Omega$, simply because such an ω would be both 1 and 3. Thus $\lim_{\mathbf{S}} \diamond \neq \Omega$. Notice that by including X_1X_3 in $\mathbf{S}$, the observables X_1 and X_3 can be measured together, and the outcom of such a measurement belongs to the set $\big\{\{1\}, \{3\}, \{2, 4\}\big\}$. Thus the presence of elements like the one above is no longer possible.
With this modification we have $\lim_{\mathbf{S}} \diamond \simeq \Omega$.

3 Information cohomology

In this section we introduce the information cohomology following [13]. The reader is supposed to be familiar with abelian categories and derived functors. We refer to [14] for notations and definitions related with these topics.
Let $\mathbf{S}$ be a unital conditional meet semilattice with terminal object $\top$. For any object $X \in Ob(\mathbf{S})$ the categorical product in $\mathbf{S}$ induces a monoid structure on the set $\mathscr{S}_X :=$

$\{Y \in Ob(\mathbf{S}) | \exists \mathrm{f} : X \to Y\}$ of the objects which are coarser than X. In fact, for any $Y_1, Y_2 \in \mathscr{S}_X$, the object X provides a common lower bound, thus the product $Y_1 Y_2$ exists. This defines a binary operation on the set $\mathscr{S}$, which is shown to be associative, commutative, and with neutral element given by $\top$. Whenever there is an arrow $\mathrm{f} : Z \to X$ in $\mathbf{S}$, the inclusion $\mathscr{S}_X \subset \mathscr{S}_Z$ is an homomorphism of monoids. Therefore we obtain a presheaf of monoids $\mathscr{S} : \mathbf{S}^{\mathrm{op}} \to \mathbf{Mon}$ such that $\mathscr{S}(\mathrm{f})$ is the inclusion $\mathscr{S}_X \subset \mathscr{S}_Z$.

Definition 3.1. Given a monoid $\mathbf{M}$, the monoid $\mathbb{R}$-algebra associated to $\mathbf{M}$, denoted $\mathbb{R}[\mathbf{M}]$, is the free $\mathbb{R}$-module on the underlying set of $\mathbf{M}$ endowed with the product

$$\left(\sum_{i=1}^{n} r_i[m_i]\right) \cdot \left(\sum_{j=1}^{k} r'_j[m'_j]\right) = \left(\sum_{i=1}^{n} \sum_{j=1}^{k} r_i r'_j [m_i \cdot_M m'_j]\right) \quad \text{where } r_i, r'_j \in \mathbb{R} \text{ and } m_i, m'_j \in \mathbf{M}$$

which is associative and admits $[1_M]$ as a neutral element.

Definition 3.2. Let $\mathbf{S}$ be a unital conditional meet semilattice.
$\mathscr{A} : \mathbf{S}^{\mathrm{op}} \to \mathbf{Rings}$ is the presheaf of rings described by

$$\begin{array}{ccc} X & \longmapsto & \mathbb{R}[\mathscr{S}_X] \\ \downarrow \varphi & & \mathscr{A}(\varphi) \uparrow \text{extension of } \mathscr{S}(\varphi) \text{ by } \mathbb{R}\text{−linarity} \\ Y & \longmapsto & \mathbb{R}[\mathscr{S}_Y] \end{array} \tag{3.1}$$

Definition 3.3. Let $\mathbf{S}$ as above. We denote with $\mathbb{R}_{\mathbf{S}} : \mathbf{S}^{\mathrm{op}} \to \mathbf{Ab}$ the constant presheaf of $\mathscr{A}$-modules which maps all $X \in Ob(\mathbf{S})$ to $(\mathbb{R}, +, 0)$, equipped with the trivial action of $\mathscr{A}_X$. ($Y \cdot r = r \ \forall \ Y \in \mathscr{A}_X$, $r \in \mathbb{R}$)

The functor $\mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \_) : \mathbf{PMod}(\mathscr{A}) \to \mathbf{Ab}$ is left exact. Moreover, the category $\mathbf{PMod}(\mathscr{A})$ is abelian and has enough injective objects. Hence the right derived functors $R^n \mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \_) = \mathrm{Ext}^n(\mathbb{R}_{\mathbf{S}}, \_)$ are well defined and form a universal δ-functor [14].

Definition 3.4. The *information cohomology* of $\mathbf{S}$ with coefficients in the $\mathscr{A}$-module $\mathcal{F}$ is

$$H^\bullet(\mathbf{S}, \mathcal{F}) := R^\bullet \mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \_)(\mathcal{F}) = \mathrm{Ext}^\bullet(\mathbb{R}_{\mathbf{S}}, \mathcal{F}) \tag{3.2}$$

Recall that properties of $\mathbf{PMod}(\mathscr{A})$ listed above, ensure that $\mathcal{F}$ has an injective resolution

$$0 \to \mathcal{F} \xrightarrow{\varsigma} \mathcal{J}^0 \xrightarrow{d^0_{\mathcal{J}}} \mathcal{J}^1 \xrightarrow{d^1_{\mathcal{J}}} \mathcal{J}^2 \dots,$$

and the group (actually a real vector space) $\mathrm{Ext}^n(\mathbb{R}_{\mathbf{S}}, \mathcal{F})$ is the n-th cohomology group of the complex obtained applying the functor $\mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \_)$ to the complex $\mathcal{J}^\bullet$.
It is, however, more convenient to fix a projective resolution $\mathcal{P}_\bullet \xrightarrow{\varepsilon} \mathbb{R}_{\mathbf{S}} \to 0$ and use it to compute the information cohomology with every presheaf of coefficients. This is possible because of the following

Proposition 3.1.

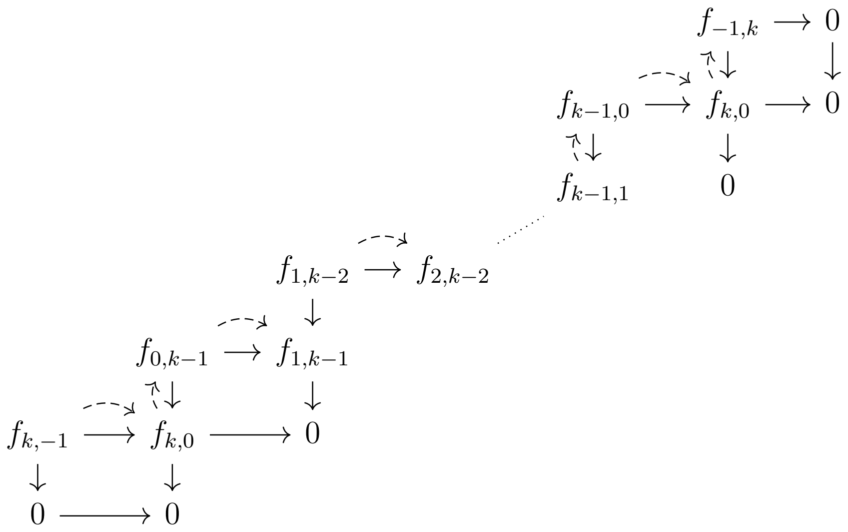
$$H^n(\mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \mathcal{J}^{\bullet})) \simeq H^n(\mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_{\bullet}, \mathcal{F})) \quad \forall\, n \in \mathbb{N}_0 \tag{3.3}$$

Proof. $\mathrm{Hom}_{\mathscr{A}}(\_, \_) : \mathbf{PMod}(\mathscr{A})^{\mathrm{op}} \times \mathbf{PMod}(\mathscr{A}) \to \mathbf{Ab}$ is a bifunctor, such that if $\mathcal{J}$ is injective, $\mathrm{Hom}_{\mathscr{A}}(\_, \mathcal{J})$ is exact, and if $\mathcal{P}$ is projective (iff it is injective in $\mathbf{PMod}(\mathscr{A})^{\mathrm{op}}$), the functor $\mathrm{Hom}_{\mathscr{A}}(\mathcal{P}, \_)$ is exact.

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & 0 & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_0, \mathcal{F}) & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_1, \mathcal{F}) & \longrightarrow & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \mathcal{J}^0) & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_0, \mathcal{J}^0) & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_1, \mathcal{J}^0) & \longrightarrow & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \mathcal{J}^1) & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_0, \mathcal{J}^1) & \longrightarrow & \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_1, \mathcal{J}^1) & \longrightarrow & \\
 & & \downarrow & & \downarrow & & \downarrow & &
\end{array} \tag{3.4}$$

In this bicomplex, the vertical differential is given by $d_v^n := d_{\mathcal{J}}^n \circ \_$ for $n \geq 0$ and $d_v^{-1} := \varsigma \circ \_$. Similarly, the horizontal differential is given by $d_h^{\bullet} := \_ \circ d_{\bullet}^{\mathcal{P}}$ for $n \geq 0$ and $d_h^{-1} = \_ \circ \varepsilon$. Thus, based on the definitions of injective and projective objects, we ca infer that all rows and columns are exact except for the first row and the first column.
This implies that the cohomology of the first row is isomorphic to the cohomology of the first column, which is (3.3). We only quote the construction of the isomorphism, omitting all the necessary verifications.
Let $f_{k,-1} \in \mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \mathcal{J}^k)$ for some $k \in \mathbb{N}$, which is a cocycle for the vertical differential, i.e. $d_{\mathcal{J}}^k \circ f_{k,-1} = 0$. Consider $f_{k,0} := d_h^{-1}(f_{k,-1})$: it is a cocycle for d_v because $d_v^k(f_{k,0}) = d_h \circ d_v(f_{k,-1}) = 0$. Therefore, since the first column is exact, there exists an element $f_{k-1,0} \in \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_0, \mathcal{J}^{k-1})$ such that $d_v(f_{k-1,0}) = f_{k,0}$. We then move again to the right along d_h: $f_{k-1,1} := d_h^0(f_{k-1,0})$ is again a cocycle w.r.t. the vertical differential, because $d_v(d_h^0(f_{k-1,0})) = d_h(f_{k,0}) = d_h(d_h(f_{k,-1})) = 0$. This allows us to find $f_{k-2,1}$. Iterating this argument results in a "ladder" path on the bicomplex (3.4), starting at the spot $(k, -1)$

and arriving at $(-1,k)$, as displayed in the following diagram



(3.5)

The element $f_{-1,k} \in \mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_k, \mathcal{F})$ is a cocycle w.r.t the horizontal differential, since

$$\varsigma \circ d_h(f_{-1,k}) = d_h(f_{0,k}) = d_f(d_h(f_{0,k-1})) = 0. \tag{3.6}$$

It is worth noting that $f_{-1,k}$ is not determined by $f_{k,-1}$. Indeed, every time the row index decreases, a choice has to be made. Therefore, $f_{k,-1} \mapsto f_{-1,k}$ is not even a function. However, it can be verified that

$$\begin{aligned} H^k(\mathrm{Hom}_{\mathscr{A}}(\mathbb{R}_{\mathbf{S}}, \mathcal{J}^\bullet)) &\xrightarrow{\sim} H^k(\mathrm{Hom}_{\mathscr{A}}(\mathcal{P}_\bullet, \mathcal{F})) \\ [f_{k,-1}] &\mapsto [f_{-1,k}] \end{aligned} \tag{3.7}$$

is well defined and is an isomorphism of $\mathbb{R}$-vector spaces. □

It remains to find a projective resolution of $\mathbb{R}_{\mathbf{S}}$ that allows for explicit calculations with cocycles. For this purpose we will use the *unnormalized bar resolution*, which is presented in Appendix B.

Let $\mathbb{R}_p : \mathbf{S} \to \mathbf{Rings}$ be the presheaf which is constant at the ring $(\mathbb{R}, +, \cdot, 0, 1)$. For any $X \in Ob(\mathbf{S})$, we have a ring homomorphism $\mathbb{R} \to \mathscr{A}_X, c \mapsto c[\top]$ that induces a structure of $\mathbb{R}$-algebra on $\mathscr{A}_X$. Moreover, $\mathscr{A}$ maps the morphisms in $\mathbf{S}$ to morphisms of $\mathbb{R}$-algebras, which means that the just defined ring homomorphisms join into a natural transformation $\iota : \mathbb{R}_p \to \mathscr{A}$.

We now apply the construction in Example B.2 to the natural trasformation $\iota : \mathbb{R}_p \to \mathscr{A}$, obtaining a relatively free allowable resolution $\beta_\bullet(\mathbb{R}_{\mathbf{S}}) \twoheadrightarrow \mathbb{R}_{\mathbf{S}}$. In particular, for any $n \geq 0$

and any $X \in Ob(\mathbf{S})$, we have

$$\mathscr{B}_n(X) := \beta_n(\mathbb{R}_{\mathbf{S}})(X) = \mathscr{A}_X \otimes_{\mathbb{R}} (\cdots \otimes_{\mathbb{R}} (\mathscr{A}_X \otimes_{\mathbb{R}} \mathbb{R}) \ldots) \simeq \underbrace{\mathscr{A}_X \otimes_{\mathbb{R}} \cdots \otimes_{\mathbb{R}} \mathscr{A}_X \otimes_{\mathbb{R}} \cdots \otimes_{\mathbb{R}} \mathscr{A}_X}_{n+1 \ times} . \tag{3.8}$$

And

$$\begin{aligned} \partial_n(X) : \mathscr{B}_n(X) &\longrightarrow \mathscr{B}_{n-1}(X) \\ Y_0 \otimes \ldots Y_n &\longmapsto \sum_{i=0}^{n} (-1)^i Y_0 \otimes Y_1 \otimes \cdots \otimes Y_i Y_{i+1} \otimes \cdots \otimes Y_n. \end{aligned} \tag{3.9}$$

Moreover, letting $\mathrm{f} : X \to Y$ to be an arrow in $\mathbf{S}$, we have $\mathscr{B}_n(\mathrm{f}) = (\mathscr{A} \otimes_{\mathbb{R}} \_ \circ \Box)^{n+1}(\mathbb{R}_{\mathbf{S}})(\mathrm{f}) = \mathscr{A}(\mathrm{f}) \otimes_{\mathbb{R}} \cdots \otimes_{\mathbb{R}} \mathscr{A}(\mathrm{f})$. The differentials are natural in X, as can be seen through a simple check.

We will show that the functors $\mathscr{B}_n$ are free $\mathscr{A}$-modules. We first recall the definition:

Definition 3.5. Let $\mathcal{C}$ be a small category. $\mathcal{O} : \mathcal{C}^{\mathrm{op}} \to \mathbf{Rings}$ is a presheaf of rings, and $\mathcal{G}$ is a presheaf of sets with the same source. The free $\mathcal{O}$-module on $\mathcal{G}$ is the functor

$$\begin{aligned} \mathcal{O}[\mathcal{G}] : &\mathcal{C}^{\mathrm{op}} \longrightarrow \mathbf{Ab} \\ &X \longmapsto \mathcal{O}_X[\mathcal{G}(X)] := \bigoplus_{x \in \mathcal{G}(X)} \mathcal{O}_X \cdot x \end{aligned} \tag{3.10}$$

Each $\mathcal{O}_X[\mathcal{G}(X)]$ is clearly an $\mathcal{O}_X$-module. Given an arrow $\mathrm{f} : X \to Y$, $\mathcal{O}[\mathcal{G}](\mathrm{f})$ is defined as a path from the top left to the bottom right in the commutative square

$$\begin{array}{ccc} \mathcal{O}_Y[\mathcal{G}(Y)] & \xrightarrow{\mathcal{O}_Y} & \mathcal{O}_Y[\mathcal{G}(X)] \\ \downarrow & & \downarrow \\ \mathcal{O}_X[\mathcal{G}(X)] & \longrightarrow & \mathcal{O}_X[\mathcal{G}(X)] \end{array}$$

Where the horizontal arrows are specified by the function $\mathcal{G}(f)$ on the bases, while the vertical ones are induced canonically by the universal property of the coproduct in $\mathbf{Ab}$. The abelian presheaf just defined makes the diagram A.1 commutative, so it is actually a presheaf of $\mathcal{O}$-modules. This construction yelds a functor $\mathcal{O}[\ ] : \mathbf{Psh}(\mathcal{C}, \mathbf{Sets}) \to \mathbf{PMod}(\mathcal{O})$, because, given a natural transformation $\zeta : \mathcal{G} \to \mathcal{G}'$, for any object X, there is a unique $\mathcal{O}_X$-linear map $\mathcal{O}[\zeta]_X : \mathcal{O}_X[\mathcal{G}(X)] \to \mathcal{O}_X[\mathcal{G}'(X)]$, such that $\mathcal{O}[\zeta]_X([y]) = [\zeta_X(y)] \quad \forall\, y \in \mathcal{G}(X)$. The naturality of $\mathcal{O}[\zeta]$ follows from the naturality of ζ.

For each $n \in \mathbb{N}_0$ there is a natural isomorphism $\mathscr{B}_n \simeq \mathscr{A}[\mathscr{S}^n]$[14], whose X component is

$$\begin{aligned} \mathscr{B}_n(X) \simeq \overbrace{\mathscr{A}_X \otimes_{\mathbb{R}} \cdots \otimes_{\mathbb{R}} \mathscr{A}_X}^{n+1 \text{ times}} \overset{\star}{\simeq} \bigoplus_{Y_1,\ldots,Y_n \in \mathscr{S}_X} \mathscr{A}_X[Y_1|Y_2|\ldots|Y_n] \simeq \mathscr{A}_X[\mathscr{S}^n_X] \\ Y \otimes Y_1 \otimes \cdots \otimes Y_n \longmapsto Y[Y_1|\ldots|Y_n] \end{aligned} \tag{3.11}$$

Remark 3.1. If $n = 0$, the functor $\mathscr{S}^n$ becomes constant at $\{*\}$ and the isomorphism $\star$ still holds, in fact it turns into the identity map on $\mathscr{A}_X$. In this case we will replace the notation $\mathscr{A}_X[\{*\}]$ with $\mathscr{A}_X[\ ]$. Actually, what should be included in square brackets is an element of $\mathbb{R}$, since the starting $\mathscr{A}$-module was $\mathbb{R}_\mathbf{S}$, but because of $\mathbb{R}$-linearity we have omitted it.

By means of $\star$, the differental $\partial_n(X) : \mathscr{B}_n(X) \to \mathscr{B}_{n-1}(X)$, for $n > 1$, takes the form

$$\partial_n(X)([Y_1|Y_2|\ldots|Y_n]) = Y_1[Y_2|\ldots|Y_n]+\sum_{k=1}^{n-1}(-)^k[Y_1|\ldots|Y_kY_{k+1}|\ldots|Y_n]+(-)^n[Y_1|\ldots|Y_{n-1}] \tag{3.12}$$

If $n = 1$ we have $\partial_1(X)([Y]) = Y[\ ] - [\ ]$, and for $n = 0$, the augmentation morphism $\varepsilon_X : \mathscr{A}_X[\ ] \to \mathbb{R}$ is determined an the only $\mathscr{A}_X$-linear map such that $\epsilon_X([\ ]) = 1$.
Our aim now is to prove that $\mathscr{B}_\bullet$ is indeed a projective resolution of $\mathbb{R}_\mathbf{S}$.

Proposition 3.2. *Keep notations from the Definition 3.5. Let $U : \mathbf{PMod}(\mathcal{O}) \to \mathbf{Psh}(\mathcal{C}, \mathbf{Sets})$ be the forgetful functor. There is a free-forgetful adjunction $\mathcal{O}[\ ] \dashv U$.*

Proof. The adjunction in this proposition is built upon the free-forgetful adjunction between the categories of sets and of modules over a ring. Details are omitted . □

Proposition 3.3. *For each $n \geq 0$, the presheaf $\mathscr{B}_n$ is a projective object in $\mathbf{PMod}(\mathscr{A})$.*

Proof. We have to show that for any diagram in $\mathbf{PMod}(\mathscr{A})$ like

$$\begin{array}{ccc} & & \mathscr{A}[\mathscr{S}^n] \\ & & \big\downarrow \alpha \\ \mathcal{N} & \xrightarrow{\ \sigma\ } & \mathcal{M} \end{array} \tag{3.13}$$

there exists a morphism $\alpha' : \mathscr{A}[\mathscr{S}^n] \to \mathcal{N}$ such that $\alpha = \sigma \circ \alpha'$. Because of Proposition 3.2 it suffices to prove the existence of a map of presheaves $\zeta : \mathscr{S}^n \to U\mathcal{N}$ such that $\bar{\alpha} = U\sigma \circ \zeta$. Indeed, provided such a ζ, we would have $\sigma \circ \bar{\zeta} = \overline{U\sigma \circ \zeta} = \bar{\bar{\alpha}} = \alpha$ and setting

[14]Here, with $\mathscr{S}^n$ we mean the presheaf of sets defined by the rule $X \mapsto \mathscr{S}^n_X$, and, on arrows, by $\mathscr{S}^n(\mathrm{f}) = (f, f, \ldots, f)$

$\alpha' := \bar{\zeta}$ the proof would be concluded.

$$\begin{array}{ccc} & & \mathscr{A}[\mathscr{S}^n] \\ & \swarrow^{\bar{\zeta}} & \downarrow \alpha \\ \mathcal{N} & \xrightarrow{\sigma} & \mathcal{M} \end{array} \tag{3.14}$$

Since for all $X \in Ob(\mathbf{S})$ the function σ_X is surjective, we can construct functions ζ_X such that $\sigma_X \circ \zeta_X = \alpha_X$, but they don't need to join in a natural transformation. But using the additional properties of the category $\mathbf{S}$ as a conditional meet semilattice we can choose functions that are compatible with restrictions. If $n = 0$, choose $\zeta_T(*) \in \sigma_T^{-1}(\alpha_T(*))$. For any object X in $\mathbf{S}$, exists a unique arrow $\tau_X : X \to T$. Defining $\zeta_X(*) := \mathcal{N}(\tau_X)(\zeta_T(*))$ we obtain a compatible family of $U\mathcal{N}$ i.e. an element of $\mathrm{Hom}_{Psh}(\underline{\{*\}}^{15}, U\mathcal{N})$.
For $n > 0$, as above we define first ζ_T choosing an element $\zeta_T((T, \ldots, T)) \in \sigma_T^{-1}(\alpha_T(T, \ldots, T))$. Next, for any $X \in Ob(\mathbf{S})$, we define ζ_X recursively:

$$\zeta_X((Y_1, Y_2, \ldots, Y_n)) := \zeta_{\prod_{i=1}^n Y_i}((Y_1, Y_2, \ldots, Y_n))$$

if the product is not isomorphic to X, otherwise we choose an element in $\sigma_X^{-1}(\alpha_X(Y_1, Y_2, \ldots, Y_n))$. □

Let $\mathcal{F}$ be an $\mathscr{A}$-module, the Proposition 3.3 implies that can we compute the groups $\mathrm{Ext}^\bullet(\mathbb{R}_\mathbf{S}, \mathcal{F})$ as the n-th cohomology groups of $(\mathrm{Hom}_\mathscr{A}(\mathscr{B}_\bullet, \mathcal{F}), \delta^\bullet)$. The differential of this complex is given by precomposition with ∂, the differential of $\mathscr{B}$. Explicitly, given $f \in \mathrm{Hom}_\mathscr{A}(\mathscr{B}_n, \mathcal{F})$ and X in $\mathbf{S}$

$$\begin{aligned} (\delta^n f)_X :&[Y_1|\ldots|Y_{n+1}] \mapsto f_X(\partial_n(X)([Y_1|\ldots|Y_{n+1}])) = \\ Y_1 \cdot f_X([Y_2|\ldots|Y_{n+1}]) &+ \sum_{k=1}^{n} (-)^k f_X([Y_1|\ldots|Y_k Y_{k+1}|Y_{n+1}]) + (-)^{n+1} f_X([Y_1|\ldots|Y_n]). \end{aligned} \tag{3.15}$$

Just to simplify notations, we will write $f_X[Y_1|\ldots|Y_n]$ in place of $f_X([Y_1|\ldots|Y_n])$. The natural transformations in $C^n(\mathbf{S}, \mathcal{F}) := \mathrm{Hom}_\mathscr{A}(\mathscr{B}_n, \mathcal{F})$ are called n-cochains; the n-cocycles are the elements of $Ker(\delta^n)$, denoted $Z^n(\mathbf{S}, \mathcal{F})$; the group of n-coboundaries is the image of δ^{n-1} which is a subgroup of $Z^n(\mathbf{S}, \mathcal{F})$ since $\delta^n \delta^{n-1} = 0$. With these notations, we find the familiar relation $H^n(\mathbf{S}, \mathcal{F}) = Z^n(\mathbf{S}, \mathcal{F}) / \delta^{n-1}(C^n(\mathbf{S}, \mathcal{F}))$. The $\mathscr{A}$-module $\mathcal{F}$ in question, is called presheaf of coefficients.
Let us conclude this section on information cohomology by presenting a property of funtoriality in the category $\mathbf{S}$ of observables. This could draw an analogy between information cohomology and other theories involving cohomology, such as the singular cohomology of

[15] if A is a set, we denote $\underline{\mathrm{A}}$ the presheaf of sets constant at A

a topological space. Moreover, the functoriality will allow us to state that probabilistic information cohomology (Sec. 4) is invariant for equivalent information structures.

Proposition 3.4. *Let $\phi : \mathbf{S} \to \mathbf{S}'$ be a morphism of unital conditional meet semilattices. Then is defined a morphism of presheaves of rings $\phi^\# : \mathscr{A} \to \phi^p\mathscr{A}'$, which induces a functor $\phi_* : \mathbf{PMod}(\mathscr{A}') \to \mathbf{PMod}(\mathscr{A})$. Moreover, $\phi^\#$ yields a chain map*

$$\Psi_\bullet : \mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}'_\bullet, \mathcal{F}') \to \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_\bullet, \phi_*\mathcal{F}') \tag{3.16}$$

and hence a map of graded vector spaces $H^\bullet(\Psi) : H^\bullet(\mathbf{S}', \mathcal{F}') \to H^\bullet(\mathbf{S}, \phi_*\mathcal{F}')$.

Proof. Since ϕ preserves finite products, the functions $\mathscr{S}_X \to \mathscr{S}'_{\phi(X)}, Y \mapsto \phi(Y)$ are homomorphisms of monoids for all $X \in Ob(\mathbf{S})$. They clearly define a natural transformation $\phi^\# : \mathscr{S} \to \phi^p\mathscr{S}'$, which uniquely determines a morphism of presheaves of rings $\mathscr{A} \to \phi^p\mathscr{A}'$, that we rename $\phi^\#$. Now, given $(\mathcal{F}', \mu')$ an $\mathscr{A}'$-module, we consider the $\mathscr{A}$-module structure given by

$$\begin{array}{ccc} \phi^p\mathscr{A}' \times \phi^p\mathcal{F}' & \xrightarrow{\phi^p(\mu')} & \phi^p\mathcal{F}' \\ {\scriptstyle \phi^\#\times \mathrm{id}}\uparrow & & \\ \mathscr{A} \times \phi^p\mathcal{F}' & & \end{array} \tag{3.17}$$

This yields a functor $\phi_* : \mathbf{PMod}(\mathscr{A}') \to \mathbf{PMod}(\mathscr{A})$, because if $\alpha : (\mathcal{F}', \mu') \to (\mathcal{G}', \nu')$ is $\mathscr{A}'$-linear, then it must be also $\mathscr{A}$-linear.

$$\begin{array}{ccccc} \mathscr{A} \times \phi^p\mathcal{F}' & \xrightarrow{\phi^\#\times\mathrm{id}} & \phi^p\mathscr{A}' \times \phi^p\mathcal{F}' & \xrightarrow{\phi^p(\mu')} & \phi^p\mathcal{F}' \\ \downarrow{\scriptstyle \mathrm{id}\times\phi^p(\alpha)} & & \downarrow{\scriptstyle \mathrm{id}\times\phi^p(\alpha)} & & \downarrow{\scriptstyle \phi^p(\alpha)} \\ \mathscr{A} \times \phi^p\mathcal{G}' & \xrightarrow{\phi^\#\times\mathrm{id}} & \phi^p\mathscr{A}' \times \phi^p\mathcal{G}' & \xrightarrow{\phi^p(\nu')} & \phi^p\mathcal{G}' \end{array} \tag{3.18}$$

Thus, applying ϕ_* to the complex of $\mathscr{A}'$-modules $\mathscr{B}'_\bullet$, we get a complex in $\mathbf{PMod}(\mathscr{A})$. Let $n \geq 0$. Having the morphisms $\mathscr{A}_X \to \mathscr{A}'_{\phi(X)}$ and $\mathscr{S}_X \to \mathscr{S}'_{\phi(X)}$ it is natural to define $\Phi^n_X : \mathscr{B}_n(X) \to \mathscr{B}'_n(\phi(X))$ as the unique $\mathscr{A}_X$-linear map such that

$$\begin{aligned} \mathscr{A}_X[\mathscr{S}^n_X] &\longrightarrow \mathscr{A}'_X[\mathscr{S}'^n_{\phi(X)}] \\ [Y_1|\dots|Y_n] &\longrightarrow [\phi(Y_1)|\dots|\phi(Y_n)]. \end{aligned} \tag{3.19}$$

These maps are clearly compatible with restriction morphisms. Thus we can consider, for any $n \in \mathbb{N}$, the natural transformation $\Phi^n : \mathscr{B}_n \to \phi_*(\mathscr{B}'_n)$, which turns out to be also $\mathscr{A}$-linear. We now claim that the just defined $\Phi^\bullet$ is a chain map. We check this on a generic X, let $n > 1$,

$$\begin{aligned}\Phi_X^{n-1} \circ \partial_n(X)([Y_1|\dots|Y_n]) &= \Phi_X^{n-1}(Y_1[Y_2|\dots|Y_n]+ \\ &+ \sum_{k=1}^{n-1}(-)^k[Y_1|\dots|Y_kY_{k+1}|\dots|Y_n] + (-)^n[Y_1|\dots|Y_{n-1}]) \\ &= \phi(Y_1)[\phi(Y_2|\dots|\phi(Y_n)]+ \\ &+ \sum_{k=1}^{n-1}(-)^k[\phi(Y_1)|\dots|\phi(Y_kY_{k+1})|\dots|\phi(Y_n)] + (-)^n[\phi(Y_1)|\dots|\phi(Y_{n-1})] \\ &= \phi_*\partial'_n(X)(\Phi_X^n([Y_1|\dots|Y_n])).\end{aligned} \tag{3.20}$$

We now define for each $n \geq 0$

$$\begin{aligned}\Psi_n : \mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}'_n, \mathcal{F}') &\longrightarrow \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_n, \phi_*\mathcal{F}') \\ f' &\longmapsto \phi_*(f') \circ \Phi^n.\end{aligned} \tag{3.21}$$

These homomorphisms of abelian groups commute with the differential $\delta^\bullet$, as can be seen from

$$\begin{aligned}\Psi_{n+1}(\delta'^n(f')) = \phi_*(\delta'^n(f')) \circ \Phi^{n+1} &= \phi_*(f') \circ \phi_*(\partial'_{n+1}) \circ \Phi^{n+1} \\ = \phi_*(f') \circ \Phi^n \circ (\partial_{n+1}) &= \delta^n(\Psi_n(f')).\end{aligned} \tag{3.22}$$

It is known that any cochain map induces a morphism of graded vector spaces in cohomology, therefore we conclude the proof considering the one induced by $\Psi_\bullet$ □

4 Probabilistic Information Cohomology

In this section we construct a family of presheaves indexed by a positive parameter α, and then we focus on information cohomology with coefficients in these presheaves. Within this framework the α-entropies arise as the only possible 1-cocycles.
We begin introducing the probabilities as a covariant functor on a finite information structure. By the way, all information structures throughout this section are supposed to be finite.

Definition 4.1. Let $(\mathbf{S}, \mathscr{E})$ be an information structure. $\mathscr{P} : \mathbf{S} \to \mathbf{Sets}$ is the functor that associates to each $X \in Ob(\mathbf{S})$ the set

$$\mathscr{P}(X) := \Big\{P : \mathscr{E}_X \to [0,1] \mid \sum_{x \in \mathscr{E}_X} P(x) = 1\Big\} \tag{4.1}$$

of all probability laws on $\mathscr{E}_X$. And to each arrow $\varphi : X \to Y$ associates the marginalization of probabilities performed with respect to $\mathscr{E}(\mathrm{f})$. This means that $\mathscr{P}(\mathrm{f}) : P \mapsto P(\mathscr{E}(\mathrm{f})^{-1}(\cdot))$ or equivalently

$$\forall\, P \in \mathscr{P}(X),\ \forall\, y \in \mathscr{E}_Y \quad \mathscr{P}(\mathrm{f})(P)(y) = \sum_{x \in \mathscr{E}(\mathrm{f})^{-1}(y)} P(x). \tag{4.2}$$

Sometimes we will write $Y_* P$ in place of $\mathscr{P}(\mathrm{f})(P)$

Let us fix some preliminary notations. Consider any arrow $\mathrm{f} : X \to Y$ in $\mathbf{S}$, $P \in \mathscr{P}(X)$ and $y \in \mathscr{E}_Y$. We will adopt the notation $P(Y = y) = P(\{Y = y\}) := P(\mathscr{E}(\mathrm{f})^{-1}(y))$, which is very common in probability theory. Similarly, given a diagram $Y \xleftarrow{\mathrm{f}} X \xrightarrow{\mathrm{g}} Z$, the with the expression $P(Y = y, Z = z) = P(\{Y = y\} \cap \{Z = z\})$ we mean $P(\mathscr{E}(\mathrm{f})^{-1}(y) \cap \mathscr{E}(\mathrm{g})^{-1}(z))$, which is also equal to $P(\mathscr{E}(\langle \mathrm{f}, \mathrm{g} \rangle)^{-1}(\omega(y, z)))$, where $\omega(y, z)$ is the unique element of $\mathscr{E}_{YZ}$ mapped to (y, z) by the canonical injection of Definition 2.2.
Furthermore, if $A \subseteq \mathscr{E}_X$ is a subset such that $P(A) > 0$, the conditional law $P_{|A} \in \mathscr{P}(X)$ is defined by

$$P_{|A}(x) = \frac{P(A \cap \{x\})}{P(A)} \tag{4.3}$$

The two conventions explained above may be combined, e.g. if $P(Y = y) > 0$, we have

$$P_{|Y=y}(x) = \frac{P(\{Y = y\} \cap \{x\})}{P(Y = y)} = \begin{cases} P(x)/P(Y = y) & x \in \mathscr{E}(\mathrm{f})^{-1}(y) \\ 0 & x \notin \mathscr{E}(\mathrm{f})^{-1}(y). \end{cases} \tag{4.4}$$

We remark that conditioning commutes with marginalization, meaning that for any diagram like $X \xrightarrow{\mathrm{f}} Y \xrightarrow{\mathrm{h}} Z$, and given any probability law P on $\mathscr{E}_X$, we have

$$\begin{aligned} (Y_* P)_{|Z=z}(y) &= \frac{Y_* P(\{Z = z\} \cap \{y\})}{Y_* P(Z = z)} = \frac{P(\mathscr{E}(\mathrm{f})^{-1}(\mathscr{E}(\mathrm{h})^{-1}(z) \cap \{y\}))}{P(\mathscr{E}(\mathrm{f})^{-1}(\mathscr{E}(\mathrm{h})^{-1}(z)))} \\ &= \frac{P(\mathscr{E}(\mathrm{h} \circ \mathrm{f})^{-1}(z) \cap \mathscr{E}(\mathrm{f})^{-1}(y))}{P(Z = z)} = Y_*(P_{|Z=z}) \end{aligned} \tag{4.5}$$

for all $y \in \mathscr{E}_Y$ and $z \in \mathscr{E}_Z$.

Definition 4.2. An adapted probability functor $\mathscr{Q}$ on an information structure $(\mathbf{S}, \mathscr{E})$ is a subfunctor of $\mathscr{P}$ that is stable under conditioning. This means that whenever there is an arrow $\mathrm{f} : X \to Y$ in $\mathbf{S}$, for any $P \in \mathscr{Q}(X)$, and every $y \in \mathscr{E}_Y$ such that $P(Y = y) > 0$, the conditional probability $P_{|Y=y}$ belongs to $\mathscr{Q}(X)$.

Remark 4.1. The set of all probabilities on a finite set of n elements is in bijection with the $n - 1$-dimensional standard simplex

$$\Delta^{n-1} = \left\{ (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n | \sum_{i=1}^{n} c_i = 1 \text{ and } c_i \geq 0\ \forall\, i : 1, \ldots, n \right\} \subset \mathbb{R}^n. \tag{4.6}$$

Thus, it can be identified with a closed subspace of $\mathbb{R}^n$ equipped with the usual topology, and inherits from $(\mathbb{R}^n, \mathfrak{B}_{\mathbb{R}^n})$[16] the structure of a measurable space. In this way, for any object X of a finite information structure, $\mathscr{P}(X)$ becomes a measurable space, as well as $\mathscr{Q}(X)$ for any probability functor $\mathscr{Q}$.
Moreover, if for all $X \in Ob(\mathbf{S})$, the set $\mathscr{Q}(X)$ is a simplicial subcomplex[17] of $\mathscr{P}(X)$, then $\mathscr{Q}$ is adpted. To see this, take $P \in \mathscr{Q}(X)$, and let $(x_1, x_2, \dots, x_n)$ be the simplex of smallest dimension in which P is contained. Given $\mathrm{f} : X \to Y$, for any $y \in \mathscr{E}_Y$, the law $P_{|Y=y}$ belongs to the boundary of $(x_1, x_2, \dots, x_n)$, unless $\mathscr{E}(\mathrm{f})^{-1}(y) = \mathscr{E}_X$, but in this case we would have $P = P_{|Y=y}$. Thus, since $\mathscr{Q}(X)$ is itself a simplicial complex, $P_{|Y=y} \in \mathscr{Q}(X)$.
Consider a finite information structure $(\mathbf{S}, \mathscr{E})$ and an adapted probability functor defined on it, say $\mathscr{Q}$. Because of to the previous remark, for any X in $\mathbf{S}$, we are allowed to consider the set $\mathrm{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R})$. Is possible to endow this set with a structure of $\mathbb{R}$-vector space, such that the operations are defined pointwise, using the operations in $\mathbb{R}$.

$$\begin{aligned} (v_1 + v_2)(P) &:= v_1(P) + v_2(P) \qquad && v_1, v_2 \in \mathrm{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R}),\ P \in \mathscr{Q}(X) \\ (r \cdot v)(P) &:= rv(P) && v \in \mathrm{Hom}_{Meas}(\mathscr{Q}, \mathbb{R}),\ r \in \mathbb{R} \end{aligned} \tag{4.7}$$

Now let α be a positive real number, we give to the vector space $\mathrm{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R})$ an $\mathscr{A}_X$-module structure depending on the parameter α. First, we assign to each $Y \in \mathscr{S}_X$ an endomorphism $Y. \in End(\mathrm{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R}))$, which operates in the following manner:

$$Y. : v \mapsto Y.v, \qquad Y.v : P \mapsto \sum_{\substack{y \in \mathscr{E}_Y \\ P(Y=y) \neq 0}} P(Y = y)^\alpha v(P_{|Y=y}) \tag{4.8}$$

It can be readily verified that $Y.$ is $\mathbb{R}$-linear. Another aspect to verify is whether $Y.v$ is measurable, a fact that becomes evident once we demonstrate the measurability of $P \mapsto P_{|Y=y}$ for every $y \in \mathscr{E}_Y$. To see this, consider P as a point $(p_1, p_2, \dots, p_n)$, where $n = |\mathscr{E}_X|$ and name f the function that testifies to $Y \in \mathcal{A}_X$. Conditioning with respect to $\{Y = y\}$ maps $(p_1, p_2, \dots, p_n)$ to the point whose i-th component is: $\frac{p_i}{P(Y=y)}$ if $\mathscr{E}(\mathbf{f})(x_i) = y$; otherwise is zero. This function is certainly measurable because it is rational in the variables $p_1, p_2, \dots, p_n$.

Proposition 4.1. *Suppose given a finite information structure $(\mathbf{S}, \mathscr{E})$ together with an adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$. Then, for any $X \in Ob(\mathbf{S})$, the function*

$$\begin{aligned} \mathscr{S}_X &\longrightarrow End(\mathrm{Hom}_{Meas}(\mathscr{Q}, \mathbb{R})) \\ Y &\longmapsto Y.(\cdot) \end{aligned} \tag{4.9}$$

is a morphism of monoids for all the possible values of the parameter α.

[16] We refer to the Borel σ-algebra

[17] Let X be a simplicial complex [6]. A simplicial subcomplex of X is a simplicial complex made up of simplices that belong to X.

Proof. Observe first that $\top.(\cdot) = \text{id}$, in fact $\forall\, v \in \text{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R})$ it holds that

$$\top.v(P) = P(\top = \{*\})^\alpha v(P_{|\top=\{*\}}) = v(P) \qquad \forall\, P \in \mathscr{Q}(X). \tag{4.10}$$

Now, pick $Y, Z \in \mathscr{S}_X$. Then there is a diagram like $Y \xleftarrow{\text{f}} X \xrightarrow{\text{g}} Z$. We have to show that $(YZ).(\cdot) = Y. \circ Z.(\cdot)$ as functions on $\text{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R})$. Let P and v vary as above,

$$\begin{aligned}
(YZ).v(P) &= \sum_{\substack{(y,z)\in\mathscr{E}_{YZ}\\ P(Y=y,Z=z)\neq 0}} P(Y=y, Z=z)^\alpha v(P_{|YZ=(y,z)}) \\
&\overset{1)}{=} \sum_{\substack{y\in\mathscr{E}_Y\\ P(Y=y)\neq 0}} \sum_{\substack{z\in\mathscr{E}_Z\\ P_{|Y=y}(\{Z=z\})\neq 0}} P(Y=y)^\alpha P_{|Y=y}(\{Z=z\})^\alpha v(P_{|YZ=(y,z)}) \\
&\overset{2)}{=} \sum_{\substack{y\in\mathscr{E}_Y\\ P(Y=y)\neq 0}} \sum_{\substack{z\in\mathscr{E}_Z\\ P_{|Y=y}(\{Z=z\})\neq 0}} P(Y=y)^\alpha P_{|Y=y}(\{Z=z\})^\alpha v((P_{|Y=y})_{|Z=z}) \\
&= \sum_{\substack{y\in\mathscr{E}_Y\\ P(Y=y)\neq 0}} P(Y=y)^\alpha Z.v(P_{|Y=y}) \\
&= Y.(Z.v)(P)
\end{aligned} \tag{4.11}$$

In 1) we used the familiar relation $P(Y = y, Z = z) = P(Y = y)P_{|Y=y}(\{Z = z\})$ deduced directly from the definition of conditional probability. For 2), just note that $P(Y = y, Z = z) \neq 0$ implies that $P_{|Y=y}(\{Z = z\}) \neq 0$ and $P(Y = y) \neq 0$, and in this case, for every $B \subset \mathscr{E}_X$, we have

$$P_{|YZ=(y,z)}(B) = \frac{P(B \cap \{Y=y\} \cap \{Z=z\})}{P(\{Y=y\} \cap \{Z=z\})} = \frac{P_{|Y=y}(B \cap \{Z=z\})}{P_{|Y=y}(\{Z=z\})} = (P_{|Y=y})_{|Z=z}(B). \tag{4.12}$$

This concludes the proof. □

Extending by $\mathbb{R}$-linearity the homomorphism of Proposition 4.1, we obtain an homomorphism of $\mathbb{R}$-algebras $\mathscr{A}_X \to End(\text{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R}))$. This yields the stucture of $\mathscr{A}_X$-module we were looking for. The $\mathscr{A}_X$-module $\text{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R})$ will be denoted by $\mathscr{F}_\alpha(X)$ or $\mathscr{F}_\alpha(\mathscr{Q}_X)$, to underline the dependence on the probability functor.

Definition 4.3. Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure together with an adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$. For each $\alpha > 0$, $\mathscr{F}_\alpha = \mathscr{F}_\alpha(\mathscr{Q})$ is the presheaf of $\mathscr{A}$-modules defined by $X \mapsto \mathscr{F}_\alpha(X)$ on objects, and such that $\mathscr{F}_\alpha(\text{f}) = \_ \circ \mathscr{Q}(\text{f})$ for any arrow $\text{f} : X \to Y$ in $\mathbf{S}$.

The maps $\_ \circ \mathscr{Q}(\mathrm{f})$ are $\mathbb{R}$-linear, as we can see by a direct verification

$$\begin{aligned}(r_1 v_1 + r_2 v_2) \circ \mathscr{Q}(\mathrm{f})(P) &= (r_1 v_1 + r_2 v_2)(Y_* P) \\ &= r_1 v_1(Y_* P) + r_2 v_2(Y_* P) \\ &= r_1 v_1 \circ \mathscr{Q}(\mathrm{f})(P) + r_2 v_2 \circ \mathscr{Q}(\mathrm{f})(P),\end{aligned} \tag{4.13}$$

where $r_1, r_2 \in \mathbb{R},\ v_1, v_2 \in \mathscr{F}_\alpha(X),\ P \in \mathscr{Q}(X)$.
To ensure that $\mathscr{F}_\alpha$ fulfills the definition A.1, we must check the commutativity of all the squares like

$$\begin{array}{ccc} \mathscr{A}_Y \times \mathscr{F}_\alpha(Y) & \longrightarrow & \mathscr{F}_\alpha(Y) \\ \big\downarrow{\scriptstyle \mathscr{A}(\mathrm{f}) \times \mathscr{F}_\alpha(\mathrm{f})} & & \big\downarrow{\scriptstyle \mathscr{F}_\alpha(\mathrm{f})} \\ \mathscr{A}_X \times \mathscr{F}_\alpha(X) & \longrightarrow & \mathscr{F}_\alpha(X) \end{array} \tag{4.14}$$

Take (Z, w) in the top left object, and $P \in \mathscr{Q}(X)$.

$$\begin{aligned} Z.w(\mathscr{Q}(\mathrm{f})(P)) &= \sum_{\substack{z \in \mathscr{E}_Z \\ Y_* P(Z=z) \neq 0}} Y_* P(Z = z)^\alpha w((Y_* P)_{|Z=z}) \\ &\overset{4.5}{=} \sum_{\substack{z \in \mathscr{E}_Z \\ P(Z=z) \neq 0}} P(Z = z)^\alpha w(\mathscr{Q}(\mathrm{f})(P_{|Z=z})) = Z.(w \circ \mathscr{Q}(\mathrm{f}))(P). \end{aligned} \tag{4.15}$$

Definition 4.4. The *probabilistic information cohomology* of $(\mathbf{S}, \mathscr{E})$ with respect to the adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Sets}$ and the parameter α is $H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$.

Recall that $H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ is the cohomology of the complex

$$C^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) := (\mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_\bullet, \mathcal{F}_\alpha(\mathscr{Q})), \delta^\bullet). \tag{4.16}$$

We will describe explicitly the cocycles of this complex and, under suitable assumptions we will compute the first cohomology group.
However, before doing that, let us revisit the discussion of section 3 about the functoriality in $\mathbf{S}$ of information cohomology, specializing it to the case of probabilistic information cohomology.
Consider a morphism $(\phi, \hat{\phi}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{S}', \mathscr{E}')$ in **InfoStr** together with two adapted probability functors $\mathscr{Q}, \mathscr{Q}'$, each defined respectively over $\mathbf{S}$ and $\mathbf{S}'$. Let X be a generic object in $\mathbf{S}$; using $\hat{\phi}_X$, a probability law on $\mathscr{E}_X$ uniquely determines a probability law on $\mathscr{E}'_{\phi(X)}$ through marginalization. We will call this operation *external marginalization*, in contrast to the term *internal marginalization*, which we will use to indicate the restriction morphisms of probability functors on an information structure. We remark that it is not guaranteed that if $P \in \mathscr{Q}(X)$ then the marginalized $P\hat{\phi}_X^{-1}$ belongs to $\mathscr{Q}'(X)$, but assuming this yields a function $\mathscr{P}(\hat{\phi}_X) : \mathscr{Q}(X) \to \phi^* \mathscr{Q}'(X)$.

The notation $\mathscr{P}(\hat{\phi}_X)$ can be misleading since in 4.1 we defined $\mathscr{P}$ as a functor on $\mathbf{S}$. However, in general, we can define the functor $\mathscr{P} : \mathbf{Sets}_{fin} \to \mathbf{Meas}$ which associates to each finite set A the simplex of all probabilities on A. Given $g : A \to B$ a function, $\mathscr{P}(g)$ is defined to be the marginalization of probabilities performed with respect to g as in definition 4.1. Then, provided $A = \{a_1, a_2, \dots, a_n\}$ and $B = \{b_1, b_2, \dots, b_m\}$, we have

$$\mathscr{P}(g) : (p_1, \dots, p_n) \mapsto \Big(\sum_{\substack{a_i \in A \\ g(a_i) = b_1}} p_i, \; \dots, \; \sum_{\substack{a_i \in A \\ g(a_i) = b_m}} p_i \Big) \tag{4.17}$$

From this explicit expression of $\mathscr{P}(g)$ as a function of the variables $p_1, \dots, p_n$, it becomes evident that $\mathscr{P}(g)$ is measurable. So when we write $\mathscr{P}(\hat{\phi}_X)$, we are actually referring to the measurable map obtained by applying the functor $\mathscr{P} : \mathbf{Sets}_{fin} \to \mathbf{Meas}$ to $\hat{\phi}_X$.
Let $\mathrm{f} : X \to Y$ be any arrow in $\mathbf{S}$, the following square

$$\begin{array}{ccc} \mathscr{Q}(X) & \xrightarrow{\mathscr{P}(\hat{\phi}_X)} & \mathscr{Q}'(\phi(X)) \\ \big\downarrow {\scriptstyle \mathscr{Q}(\mathrm{f})} & & \big\downarrow {\scriptstyle \mathscr{Q}'(\phi(\mathrm{f}))} \\ \mathscr{Q}(Y) & \xrightarrow{\mathscr{P}(\hat{\phi}_Y)} & \mathscr{Q}'(\phi(Y)) \end{array} \tag{4.18}$$

is commutative because the naturality of $\hat{\phi}$ entails

$$P\,(\mathscr{E}\mathrm{f})^{-1}\,\hat{\phi}_Y^{-1} = P\hat{\phi}_X^{-1}\,(\mathscr{E}'\phi(\mathrm{f}))^{-1} \quad \forall\, P \in \mathscr{Q}(X). \tag{4.19}$$

Then, provided that $\mathscr{P}(\hat{\phi}_X)(\mathscr{Q}_X) \subset \mathscr{Q}'_X$ for each $X \in Ob(\mathbf{S})$, the maps $\big(\mathscr{P}(\hat{\phi}_X)\big)_{X \in Ob(\mathbf{S})}$ join into a natural transformation $\mathscr{Q} \to \phi^*\mathscr{Q}'$. The commutativity of the square (4.18) also shows that the operations of external and internal marginalization commute with each other.
Applying the functor $\mathrm{Hom}_{Meas}(\_, \mathbb{R}) : \mathbf{Meas} \to \mathbf{Mod}_{\mathbb{R}}$[18] to the maps $\mathscr{P}(\hat{\phi}_X)$, we obtain a map of presheaves of $\mathbb{R}$-modules $\Theta : \mathrm{Hom}_{Meas}(\phi^*\mathscr{Q}', \mathbb{R}) \to \mathrm{Hom}_{Meas}(\mathscr{Q}, \mathbb{R})$, whose X component is

$$\begin{aligned} \Theta_X : \mathrm{Hom}_{Meas}(\mathscr{Q}'(\phi(X)), \mathbb{R}) &\longrightarrow \mathrm{Hom}_{Meas}(\mathscr{Q}(X), \mathbb{R}) \\ v' &\longmapsto v' \circ \mathscr{P}(\hat{\phi}_X) \end{aligned} \tag{4.20}$$

As we will see in the next proposition, under certain assumptions on $\hat{\phi}$ and on the probability functors, Θ turns out to be $\mathscr{A}$-linear and allows us to define a map of graded vector spaces between the information cohomology of $\mathbf{S}$ with coefficients in $\mathscr{F}_\alpha(\mathscr{Q})$ and the information cohomology of $\mathbf{S}'$ with coefficients in $\mathscr{F}_\alpha(\mathscr{Q}')$, for each $\alpha > 0$.

Proposition 4.2. *Let* $(\phi, \hat{\phi}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{S}', \mathscr{E}')$ *a morphism of information structures.*

[18] $\mathbf{Mod}_{\mathbb{R}}$ is the category of $\mathbb{R}$-modules and $\mathbb{R}$-linear functions

Let $\mathscr{Q}$ (resp. $\mathscr{Q}'$) be an adapted probability functor on $\mathbf{S}$ (resp $\mathbf{S}'$). Suppose that

1. *for all $X \in Ob(\mathbf{S})$, the map $\hat{\phi}_X$ is a bijection,*
2. *for all $X \in Ob(\mathbf{S})$ and all $P \in \mathscr{Q}(X)$, the law $\mathscr{P}(\hat{\phi}_X)(P)$ belongs to $\mathscr{Q}'(X)$.*

Then, for every $\alpha > 0$ there exists a cochain map

$$\Gamma^\bullet : \big(C^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')), \delta\big) \to \big(C^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))\big) \tag{4.21}$$

such that the image of $f' \in C^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}'))$ is given by

$$(\Gamma^n f')_X[Y_1|\dots|Y_n](P) := f'_{\phi(X)}[\phi(Y_1)|\dots|\phi(Y_n)](P\hat{\phi}_X^{-1}) \tag{4.22}$$

This cochain map induces a morphism of graded vector spaces in cohomology

$$\Gamma^\bullet : H^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')) \to H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \tag{4.23}$$

Proof. Fix $\alpha > 0$. We refer to the notations of Proposition 3.4, specialized to the case $\mathcal{F}' := \mathscr{F}_\alpha(\mathscr{Q}')$. Recall that we have the chain maps $\Phi_n : \mathscr{B} \to \phi_*\mathscr{B}'$ and $\Psi_\bullet$, whose n-th graded is

$$\begin{aligned} \Psi_n : \mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}'_n, \mathscr{F}_\alpha(\mathscr{Q}')) &\longrightarrow \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_n, \phi_*\mathscr{F}_\alpha(\mathscr{Q}')) \\ f' &\longmapsto \phi_*(f') \circ \Phi^n. \end{aligned} \tag{4.24}$$

Once we have shown that Θ is $\mathscr{A}$-linear, we can just post-compose $\Psi_n(f')$ with Θ to obtain a morphism of $\mathscr{A}$-modules $\mathscr{B}_n \to \mathscr{F}_\alpha(\mathscr{Q})$. Notice that indeed $\phi_*\mathscr{F}_\alpha(\mathscr{Q}')$ and $\mathrm{Hom}_{Meas}(\phi^*\mathscr{Q}', \mathbb{R})$ coincide as abelian presheaves.
Therefore, we need to prove that the square A.1 with reference to $\Theta : \phi_*\mathscr{F}_\alpha(\mathscr{Q}') \to \mathscr{F}_\alpha(\mathscr{Q})$, is commutative. This holds true if and only if for any object X in $\mathbf{S}$,

$$Y.(v' \circ \mathscr{P}(\hat{\phi}_X)) = (\phi(Y).v') \circ \mathscr{P}(\hat{\phi}_X) \quad \forall\, v' \in \mathscr{F}_\alpha(\mathscr{Q}'),\ \forall\, Y \in \mathscr{S}_X. \tag{4.25}$$

Let f : $X \to Y$ be the arrow that makes Y an element of $\mathscr{S}_X$, and pick any $P \in \mathscr{Q}(X)$. Since all the components of $\hat{\phi}$ are assumed to be bijective, $\hat{\phi}_Y^{-1}(y')$ is an element of $\mathscr{E}_Y$, for all $y' \in \mathscr{E}'_Y$. In this case the equation (4.19) implies

$$P(Y = \hat{\phi}_Y^{-1}(y')) = P\hat{\phi}_X^{-1}(\{\phi(Y) = y'\}) \quad \forall\, y' \in \mathscr{E}'_Y, \tag{4.26}$$

which in turn implies

$$
\begin{aligned}
(P\hat{\phi}_X^{-1})_{|\phi(Y)=\hat{\phi}_Y(y)}(x') &= \frac{P(\hat{\phi}_X^{-1}(x') \cap \hat{\phi}_X^{-1}(\mathscr{E}'\phi(\mathrm{f})^{-1}(\hat{\phi}_Y(y)))}{P\hat{\phi}_X^{-1}(\phi(Y) = \hat{\phi}_Y(y))} \\
&= \frac{P(\hat{\phi}_X^{-1}(x')) \cap \mathscr{E}(\mathrm{f})^{-1}(y)}{P(Y = y)} \\
&= P_{|Y=y}(\hat{\phi}_X^{-1}(x')) \qquad \forall\, y \in \mathscr{E}_Y \;\; \forall x' \in \mathscr{E}'_{\phi(X)}.
\end{aligned}
\tag{4.27}
$$

These observation entails

$$
\begin{aligned}
Y.(v' \circ \mathscr{P}(\hat{\phi}_X))(P) &= \sum_{\substack{y \in \mathscr{E}_Y \\ P(Y=y) \neq 0}} P(Y = y) v'(P_{|Y=y}\hat{\phi}_X^{-1}) \\
&\overset{4.26}{\underset{4.27}{=}} \sum_{\substack{y \in \mathscr{E}_Y \\ P(Y=y) \neq 0}} P\hat{\phi}_X^{-1}(\phi(Y) = \hat{\phi}_Y(y)) v'((P\hat{\phi}_X^{-1})_{|\phi(Y)=\hat{\phi}_Y(y)}) \\
&\overset{y' := \hat{\phi}_Y(y)}{=} \sum_{\substack{y' \in \mathscr{E}'_\phi(Y) \\ P\hat{\phi}_X^{-1}(\phi(Y)=y') \neq 0}} P\hat{\phi}_X^{-1}(\{\phi(Y) = y'\}) v'((P\hat{\phi}_X^{-1})_{|\phi(Y)=y'}) = (\phi(Y).v')(P\hat{\phi}_X^{-1})
\end{aligned}
\tag{4.28}
$$

and thus 4.25 is proved. We are now allowed to define

$$
\begin{aligned}
\Gamma^n : \mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}'_n, \mathscr{F}_\alpha(\mathscr{Q}')) &\longrightarrow \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_n, \mathscr{F}_\alpha(\mathscr{Q})) \\
f' &\longmapsto \Theta \circ \phi_*(f') \circ \Phi^n = \Theta \circ \Psi(f'),
\end{aligned}
\tag{4.29}
$$

which is in agreement with (4.22). In fact

$$
\begin{aligned}
\big(\Theta_X \circ \phi_*(f')_X \circ \Phi_X^n([Y_1|\ldots|Y_n])\big)(P) &= \Theta_X(f'_{\phi(X)}[\phi(Y_1)|\ldots|\phi(Y_n)])(P) \\
&= f'_{\phi(X)}[\phi(Y_1)|\ldots|\phi(Y_n)](P\hat{\phi}_X^{-1})
\end{aligned}
\tag{4.30}
$$

just using the the definitions of the objects involved. Observe that $\Gamma^\bullet$ is a cochain map because $\Psi_\bullet$ is. This concludes the proof as in Proposition 3.4. □

The next corollary states that the probabilistic information cohomology is invariant for equivalent information structures.

Corollary 4.1. *If $(\phi, \hat{\phi}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{S}', \mathscr{E}')$ is an isomorphism in **InfoStr** and for every $X \in Ob(\mathbf{S})$, the map $\mathscr{P}(\hat{\phi}_X) : \mathscr{Q}(X) \to \mathscr{Q}'(\phi(X))$ is surjective, then $\Gamma^\bullet$ is an isomorphism of cochain complexes.*

Proof. There exists a morphism $(\phi^{-1}, \widehat{\phi^{-1}}) : \mathbf{S}' \to \mathbf{S}$ such that

$$
\begin{aligned}
\phi \circ \phi^{-1} = \mathbb{1}_{\mathbf{S}'} \quad & \phi^{-1} \circ \phi = \mathbb{1}_{\mathbf{S}} \\
\phi^*(\widehat{\phi^{-1}}) \circ \hat{\phi} = \mathrm{id}_{\mathscr{E}} \quad & \left(\phi^{-1}\right)^*(\hat{\phi}) \circ \widehat{\phi^{-1}} = \mathrm{id}_{\mathscr{E}'}
\end{aligned}
\tag{4.31}
$$

Thus, for any X in $\mathbf{S}$ we have $\hat{\phi^{-1}}_{\phi(X)} \circ \hat{\phi}_X = \mathrm{id}_{\mathscr{E}_X}$ and $\hat{\phi}_X \circ \hat{\phi^{-1}}_{\phi(X)} = \mathrm{id}_{\mathscr{E}'_{\phi(X)}}$. This implies that all $\hat{\phi}_X$ are bijective, but also all $\hat{\phi^{-1}}_{X'}$ for $X' \in Ob(\mathbf{S}')$ are bijective, since ϕ is bijective as a function between the object of $\mathbf{S}$ and those of $\mathbf{S}'$. Therefore, we can apply the previous Proposition 4.2 to both ϕ and ϕ^{-1}, so as to find two cochain maps $\Gamma^\bullet$ and ${}^{-1}\Gamma^\bullet$. We still have to verify that these two maps are inverses of each other. Before proceeding with the proof, we note that the map $\mathscr{P}(\hat{\phi}_X) : \mathscr{Q}_X \to \mathscr{Q}'_{\phi(X)}$ is actually bijective: it is surjective by assumption, and it is injective because marginalizing a probability with respect to a bijective function (like $\hat{\phi}_X$) does not change the law but only the names of possible outcomes. Therefore, $P\hat{\phi}_X^{-1}$ uniquely determines P.
Now take any $n \in \mathbb{N}_0$ and any $f' \in \mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}'_n, \mathscr{F}_\alpha(\mathscr{Q}'))$;

$$
\begin{aligned}
\left(({}^{-1}\Gamma)^n \Gamma^n f'\right)_{\phi(X)}&[\phi(Y_1)|\dots|\phi(Y_n)](P\hat{\phi}_X^{-1}) = \\
&= (\Gamma^n f')_{\phi^{-1}(\phi(X))}\,[\phi^{-1}\phi(Y_1)|\dots|\phi^{-1}\phi(Y_n)](P\hat{\phi}_X^{-1}\hat{\phi^{-1}}_X^{-1}) \\
&= (\Gamma^n f')_X\,[Y_1|\dots|Y_n](P) \\
&= f'_{\phi(X)}[\phi(Y_1)|\dots|\phi(Y_n)](P\hat{\phi}_X^{-1})
\end{aligned}
\tag{4.32}
$$

$$\forall\, X \in Ob(\mathbf{S}) \quad \forall\,(Y_1,\dots,Y_n) \in \mathscr{S}_X^n \quad \forall\, P \in \mathscr{Q}(X)$$

Since both $\phi : Ob(\mathbf{S}) \to Ob(\mathbf{S}')$ and $\mathscr{P}(\hat{\phi}_X)$ are bijective, the quantifiers on X, $(Y_1,\dots,Y_n)$, and P in (4.32) work as if they were quantifiers on $X' \in Ob(\mathbf{S}')$, $(Y_1,\dots,Y_n) \in {\mathscr{S}'_{X'}}^n$ and $P' \in \mathscr{Q}'(X')$ respectively. Hence (4.32) shows that ${}^{-1}\Gamma^\bullet \circ \Gamma^\bullet = \mathrm{id}$. The other equality can be proven analogously. □

4.1 $H^0(\mathbf{S}, \mathscr{F}_\alpha)$

Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure equipped with an adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$.
A 0-cochain of $C^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ is a map of $\mathscr{A}$-modules $f : \mathscr{A}[\,] \to \mathscr{F}_\alpha(\mathscr{Q})$. As we saw in Proposition 3.3, such a map is uniquely determined by a morphism of presheaves of sets $f : \underline{\{[\,]\}} \to U \circ \mathscr{F}_\alpha(\mathscr{Q})$, where U denotes the forgetful functor. So it can be viewed as a compatible family $(f_X[\,])_{X \in Ob(\mathbf{S})} \in \lim_{\mathbf{S}} U \circ \mathscr{F}_\alpha(\mathscr{Q})$, and since $\mathbf{S}$ admits a terminal object $\top$, the whole family is determined by $f_\top[\,]$.

$$X \xrightarrow{\tau_X} \top, \quad f_X[\,](P) = f_\top[\,](\mathscr{Q}(\tau_X)(P)) = f_\top[\,](1) := K \in \mathbb{R} \qquad \forall\, P \in \mathscr{Q}_X, \tag{4.33}$$

where 1 indicates the only possible probability law on the set $\mathscr{E}_\top = \{*\}$. The constant K is independent of both X and P, and fully specifies the 0-cochain f. This implies the existence of an isomorphism of abelian groups

$$C^0(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \simeq (\mathbb{R}, +, 0) \quad \forall\, \alpha > 0 \tag{4.34}$$

A 0-cochain f, represented by the constant K according to (4.34), is a 0-cocycle if $\delta^0(f) = f \circ \partial_1 = 0$. Let X be an object of $\mathbf{S}$, we compute the value $\delta^0(f)_X$ on $P \in \mathscr{Q}_X$:

$$\begin{aligned}(f_X \circ \partial_1(X))[Y](P) &= f_X(Y[\,] - [\,])(P) \\ &= Y.f_X([\,])(P) - f_X[\,](P) \\ &= \sum_{y \in \mathscr{E}_Y} P(Y=y)^\alpha f_X[\,](P_{|Y=y}) - K \\ &= K\left(\sum_{y \in \mathscr{E}_Y} P(Y=y)^\alpha - 1\right) \qquad Y \in \mathscr{S}_X,\ P \in \mathscr{Q}\end{aligned} \tag{4.35}$$

If $\alpha = 1$, then $\forall\, X \in Ob(\mathbf{S})$, $\forall\, Y \in \mathscr{S}_X$, $\forall\, P \in \mathscr{Q}$ we have $\left(\sum_{y \in \mathscr{E}_Y} P(Y=y)^\alpha - 1\right) = 0$, so the last term in (4.35) vanishes with K free to vary in $\mathbb{R}$, or equivalently $\delta^0(f) = 0$ for any 0-cochain f. This can be expressed also as $H^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) = Z^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) = C^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q}))$.

In the case of $\alpha \neq 1$, we have to distinguish between two cases:

-if the probability functor satisfies

$$\exists X \in Ob(\mathbf{S})\ , \exists\, p \in \mathscr{Q}(X) \ \text{ that is not an atomic probability} \tag{4.36}$$

then for such a p, it holds that $\left(\sum_{x \in \mathscr{E}_X} p(X=x)^\alpha - 1\right) \neq 0$. This implies that $\delta^0(f)_X$ vanishes if and only if $K = 0$. Hence in this case $Z^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) = 0$.

-However if $\mathscr{Q}$ does not satisfies (4.36), we obtain again $\left(\sum_{y \in \mathscr{E}_Y} P(Y=y)^\alpha - 1\right) = 0$ independent of $X \in Ob(\mathbf{S})$, $Y \in \mathscr{S}_X$, $P \in \mathscr{Q}$. Whence we conclude as before that $Z^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) = C^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q}))$. This may be considered a degenerate case, but the definitions we provided do not forbid it.

4.2 1-cocycles

Keeping the same setting of the previous subsection, the 1-cochain are the elements of $\mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_1, \mathscr{F}_\alpha(\mathscr{Q}))$. According to Proposition 3.2. a 1-cochain f is in 1:1 correspondence with the map of presheaves of sets $\mathscr{S} \to U \circ \mathscr{F}_\alpha(\mathscr{Q})$ obtained by restricting f to $\mathscr{S}$. Moreover, the naturality of f implies that for any diagram like $X \to Y \to Z$ in $\mathbf{S}$ the equality

$$f_X[Z](P) = f_Y[Z](Y_* P) \quad \forall\, P \in \mathscr{Q}(X) \tag{4.37}$$

holds. Using (4.37), we see that the map $f_Y[Y]$ fully specifies the map $f_X[Y]$ any time there is an arrow $X \to Y$ is $\mathbf{S}$. For this reason we will often adopt the notation $f[Y] := f_Y[Y]$. Note that f is uniquely determined by the collection of measurable functions

$(f[X])_{X\in Ob(\mathbf{S})}$, which is clearly determined by f itself. This defines a bijection

$$C^1(\mathbf{S},\mathscr{F}_\alpha(\mathscr{Q})) \simeq \prod_{X\in Ob(s)} \mathscr{F}_\alpha(\mathscr{Q}(X)). \tag{4.38}$$

The α-entropy can be used to define an element of $C^1(\mathbf{S},\mathscr{F}_\alpha(\mathscr{Q}))$.

Definition 4.5. Given $\alpha > 0$, we define $S_\alpha : \mathscr{B}_1 \to \mathscr{F}_\alpha$ to be the 1-cochain such that

$$\forall\, X \in Ob(\mathbf{S}),\ \forall\, P \in \mathscr{Q}(X) \quad S_\alpha[X](P) := S_\alpha(P) \tag{4.39}$$

This particular type of 1-cochains will play a crucial role in this section. A first instance of this claim is given by the structure of the 1-coboundaries. In the previous subsection we saw that if $\alpha = 1$ all 0-cochains are also cocycles, thus $\delta^0 C^0(\mathbf{S},\mathscr{F}_1(\mathscr{Q})) = 0$. The same happens if $\mathscr{Q}$ contains only atomic probabilities. Otherwise, named K the 0-cochain such that $K_T[\,](1) = K \in \mathbb{R}$, we proved that

$$\left(\delta^0(K)\right)[Y] : P \mapsto K\left(\sum_{y\in\mathscr{E}_Y} P(Y=y)^\alpha\ -1\right) = KS_\alpha(P) \qquad Y \in Ob(\mathbf{S}), \tag{4.40}$$

which means that $\delta^0(K) = K\cdot S_\alpha$ as 1-cochains. Therefore, we have $\delta^0 C^0(\mathbf{S},\mathscr{F}_\alpha(\mathscr{Q})) = \mathbb{R}\cdot S_\alpha \simeq \mathbb{R}$, i.e. the real vector space of the 1-coboundaries is 1-dimensional and generated by S_α.
As for 1-cocycles, we consider $f \in C^0(\mathbf{S},\mathscr{F}_\alpha(\mathscr{Q}))$ and compute the X component of $\delta^0(f)$

$$\begin{aligned}\left(\delta^1(f)\right)_X [Y|Z] = f_X\left(Y[Z] - [YZ] + [Y]\right) = \\ Y.f_X[Z] - f_X[YZ] + f_X[Y] \qquad \forall\, Y,Z \in \mathscr{S}_X.\end{aligned} \tag{4.41}$$

The 1-cochain f belongs to $Ker(\delta^1)$ if and only if for any object X in $\mathbf{S}$, $(\delta^1(f))_X = 0$. Representing f as an element of $\prod_{X\in Ob(s)} \mathscr{F}_\alpha(\mathscr{Q}(X))$, the cocycle condition becomes

$$\begin{gathered}\forall\, X,Y \in Ob(\mathbf{S}) \text{ such that the product } XY \text{ exists} \\ f[X] + X.f[Y] = f[XY] = f[Y] + Y.f[X]\end{gathered} \tag{4.42}$$

The symmetry of the equation is due to the fact that the product in $\mathscr{S}_{XY}$ is commutative.

Proposition 4.3. *Let $f \in Z^1(\mathbf{S},\mathscr{F}_\alpha(\mathscr{Q}))$, then for any $X \in Ob(\mathbf{S})$ it holds that*

$$f[X](\delta_x) = 0 \quad \forall\, x \in \mathscr{E}_X. \tag{4.43}$$

Where δ_x denotes the atomic probability with support in $\{x\}$.

Proof. From the cocycle condition (4.42) we obtain

$$f[X] = f[X \cdot X] = X.f[X] + f[X] \implies X.f[X] = 0. \tag{4.44}$$

On the other hand $X.f[X](\delta_x) = f[X](\delta_x)$ because x is the only element of $\mathscr{E}_X$ such that $\delta_x(X = x) \neq 0$. Hence (4.43) follows. $\square$

Observe that for every $X \in Ob(\mathbf{S})$ such that $\mathscr{E}_X$ is a singleton, we can apply the proposition and conclude that $f[X] \equiv 0$ for all 1-coycles, since the only possible probability law on $\mathscr{E}_X$ is atomic.

Example 4.1. We consider a very simple information structure

$$\mathbf{S}: \quad \begin{array}{ccc} \bot & \xrightarrow{\tau} & \top \\ \downarrow{\scriptstyle \mathscr{E}} & & \downarrow{\scriptstyle \mathscr{E}} \\ \{a,b\} & \xrightarrow{\mathscr{E}(\tau)} & \{*\} \end{array} \tag{4.45}$$

In this example, we consider the probabilistic information cohomology constructed using the probability functor $\mathscr{P} : \mathbf{S} \to \mathbf{Meas}$. In this setting, a 1-cocycle f is determined by the measurable function $f[\bot] : \mathscr{P}(\{a,b\}) \to \mathbb{R}$, in fact $f[\top] \equiv 0$ by Proposition 4.43. Being f a cocycle, $f[\bot]$ must satisfy the equation $\bot .f[\bot] \equiv 0$ which holds if and only if $f[\bot](0,1) = 0 = f[\bot](1,0)$. Note that $(1,0)$ and $(0,1)$ are the atomic probabilities with support respectively in $\{a\}$ and $\{b\}$. All the other relations derived by the cocycle condition become tautological. Therefore, $f[\bot]$ is free to vary among the measurable functions between $\Delta(1) = \{(p_a, p_b) \in \mathbb{R}^2 | p_a + p_b = 1\}$ and the real number. This means that there is an isomorphism of $\mathbb{R}$-vector spaces $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \simeq \mathrm{Hom}_{Meas}(\Delta(1), \mathbb{R})$. Then

$$H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \simeq \frac{\mathrm{Hom}_{Meas}(\Delta(1), \mathbb{R})}{\mathbb{R} \cdot S_\alpha} \tag{4.46}$$

is an infinite dimensional vector space, since $\mathrm{Hom}_{Meas}(\Delta(1), \mathbb{R})$ is.

Proposition 4.4. *Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure and let $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$ be an adapted probability functor. Then for any positive real number α, the 1-cochain S_α (def.4.5) is a cocycle.*

Proof. Consider a generic diagram in $\mathbf{S}$ of the form $X \xleftarrow{\pi_X} XY \xrightarrow{\pi_Y} Y$. The cocycle condition reads $S_\alpha[XY] = S_\alpha[Y] + Y.S_\alpha[X]$ which means that, for any $P \in \mathscr{Q}(XY)$,

$$S_\alpha(P) = S_\alpha(Y_* P) + \sum_{y \in \mathscr{E}_Y} P(Y = y) S_\alpha(X_*(P_{|Y=y})). \tag{4.47}$$

This equation holds because it is exactly the chain rule of α-entropy.
However, if $\alpha \neq 1$, this proposition follows immediately because S_α is even a 1-coboundary (4.40). $\square$

Example 4.2. Here is another simple but much more interesting example. Let $(\mathbf{S}, \mathscr{E})$ be the information structure pictured as

$$\mathbf{S}: \quad \begin{array}{ccccc} & & X & & \\ & \overset{\pi_X}{\nearrow} & & \overset{\tau_X}{\searrow} & \\ XY & & & & \top \\ & \underset{\pi_Y}{\searrow} & & \underset{\tau_Y}{\nearrow} & \\ & & Y & & \end{array} \tag{4.48}$$

We set $\mathscr{E}_X = \mathscr{E}_Y := \{0,1\}$ and $\mathscr{E}_{XY} := \{(0,0),(1,0),(0,1)\}$; the restriction morphisms are defined to be the projections from the cartesian product. As in the previous example, we take $\mathscr{P} : \mathbf{S} \to \mathbf{Meas}$ as the adapted probability functor. Consider now $f \in Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$, the equation (4.42) implies that

$$\begin{gathered} \forall (p_1,p_2,p_3) \in \mathbb{R}^3 \text{ such that } p_1+p_2+p_3 = 1, \\ X.f_{XY}[Y](p_1,p_2,p_3) + f[X](p_1+p_3,p_2) = Y.f_{XY}[X](p_1,p_3,p_2) + f[Y](p_1+p_2,p_3), \end{gathered} \tag{4.49}$$

where the triple (p_1,p_2,p_3) represents the probability distribution P on $\mathscr{E}_{XY}$ such that $P(\{(0,0)\}) = p_1$, $P(\{(1,0)\}) = p_2$ and $P(\{(0,1)\}) = p_3$. Following this notation, the marginalized probabilities are represented as $\mathscr{P}(\pi_X)(P) = X_*P = (p_1+p_3,p_2)$ and $\mathscr{P}(\pi_Y)(P) = Y_*P = (p_1+p_2,p_3)$. Similarly, we can compute the triples representing the conditioned probabilities which appear in equation (4.49) because of the action of $\mathscr{A}$ on $\mathscr{F}_\alpha$. Note that $P_{|X=1} = (0,1,0)$ and $P_{|Y=1} = (0,0,1)$ are atomic, thus we know, by Proposition 4.43, that both $f_{XY}[X](P_{|Y=1})$ and $f_{XY}[Y](P_{|X=1})$ are equal to 0. Furthermore, whenever P is an atomic probability, the equation (4.49) results trivially satisfied. Assuming this is not the case, i.e assuming there is no $p_i = 1$ for $i : 1,2,3$, the equation (4.49) becomes

$$\begin{aligned} &(p_1+p_3)^\alpha f[Y]\left(\frac{p_1}{p_1+p_3}, \frac{p_3}{p_1+p_3}\right) + f[X](p_1+p_3,p_2) \\ &= (p_1+p_2)^\alpha f[X]\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right) + f[Y](p_1+p_2,p_3). \end{aligned} \tag{4.50}$$

Observe that setting $p_1 = 0$ in the last equation, we obtain

$$f[X](p_3,p_2) = f[Y](p_2,p_3) \quad \text{with } \; p_3 = 1-p_2. \tag{4.51}$$

Therefore, defining the measurable function $u : [0,1) \to \mathbb{R}$ by $u(p) := f[X](p, 1-p)$, equation (4.51) implies that $f[Y](p, 1-p) = u(1-p)$ and (4.50) becomes

$$\begin{gathered} \forall\ (p_2, p_3) \in [0,1)^2 \text{ ; such that } p_2 + p_3 \in [0,1] \\ (1-p_2)^\alpha u\left(\frac{p_3}{1-p_2}\right) + u(1-p_2) = (1-p_3)^\alpha u\left(\frac{1-p_2-p_3}{1-p_3}\right) + u(p_3) \end{gathered} \tag{4.52}$$

The equation (4.52) is strictly related with the so called "fundamental equation of information theory"(FEITH). In the next subsection we will prove that the only measurable solution of (4.52) is the α-entropy, following the proofs in [12] and [7] for the case $\alpha = 1$, and in [3] for $\alpha \neq 1$.

4.3 Functional equation

This entire subsection is devoted to solving the functional equation displayed in the next theorem.

Theorem 4.1. *Let $\alpha > 0$ and $u : [0,1] \to \mathbb{R}$ be a measurable[19] function which satisfies*

$$\begin{gathered} \forall\ (x, y) \in [0,1)^2 \textit{ such that } x + y \in [0,1] \\ u(1-x) + (1-x)^\alpha u\left(\frac{y}{1-x}\right) = u(y) + (1-y)^\alpha u\left(\frac{1-x-y}{1-y}\right) \end{gathered} \tag{4.53}$$

Then exists a real number λ such that $\forall\, x \in [0,1]$,

$$u(x) = \begin{cases} \lambda\left(-x\log_2(x) - (1-x)\log_2(1-x)\right) = \lambda S_1(x, 1-x) & \textit{if } \alpha = 1 \\ \lambda\left(\frac{1}{1-\alpha}\left(x^\alpha + (1-x)^\alpha - 1\right)\right) = \lambda S_\alpha(x, 1-x) & \textit{if } \alpha \neq 1 \end{cases} \tag{4.54}$$

With the conventions $0\log_2 0 := \lim_{x\to 0} x\log_2 x = 0$

The proof is based on the two propositions presented below, and each of them is achieved through several steps and lemmas.

Proposition 4.5 (Regularity)**.** *Any measurable solution of (4.53) belongs to $C^\infty((0,1))$.*

We start the proof of this first proposition with the following

Lemma 4.1. *Any measurable solution of (4.53) is locally bounded [20] on $(0,1)$.*

Proof. Let $y_0 \in (0,1)$, we are going to prove that

$$\begin{aligned} &\exists\, N = N(y_0) \in \mathbb{N},\ \exists\, U \ni y_0 \textit{ and } U \overset{\circ}{\subset} (0,1) \textit{ s.t. } \forall\, y \in U\ \exists\, x = x(y) \in (0,1) \textit{ s.t.} \\ &\quad 1-x, \frac{y}{1-x}, \frac{1-x-y}{1-y} \in u^{-1}([-N, N]). \end{aligned} \tag{4.55}$$

[19] Measurable with respect to the Lebesgue measure on $\mathbb{R}$ and on $[0,1)$

[20] A function $u : (0,1) \to \mathbb{R}$ is called locally bounded if for every $x \in (0,1)$, there exists an open neighborhood U of x such that u is bounded on U. This is equivalent to saying that u is bounded on every compact subset of $(0,1)$.

This would imply that u is locally bounded, because, using (4.53), we would have

$$|u(y)| \leq |u(\frac{y}{1-x})| + |u(1-x)| + |u(\frac{1-x-y}{1-y})| \leq 3N \quad \forall\, y \in U \tag{4.56}$$

From now on, we will denote the Lebesgue measure by μ.
Let $A_n := u^{-1}([-n,n])$; since u is measurable, $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of measurable subsets of $(0,1)$ which is increasing to $(0,1)$. The σ-additivity of μ implies that $\lim_{n\to\infty} \mu(A_n) = \mu((0,1)) = 1$, thus

$$\forall\, \varepsilon > 0,\, \exists N = N(\varepsilon)\ s.t.\ \forall n \geq N,\ \mu((0,1) \setminus A_n) < \varepsilon. \tag{4.57}$$

Let us fix some useful notations: given a real number z and a measurable subset $C \subset \mathbb{R}$, we define

$$\begin{aligned} z + C &:= \{z + c \in \mathbb{R} | c \in C\} \\ zC &:= \{zc \in \mathbb{R} | c \in C\} \\ C^{-1} &:= \left\{c^{-1} \in \mathbb{R} | c \in C\right\} \end{aligned} \tag{4.58}$$

With these notations for all $x, y \in \mathbb{R}$ we have:

$$\begin{aligned} 1 - x \in A_N &\iff x \in\ 1 - A_N, \\ \frac{y}{1-x} \in A_N &\iff x \in\ 1 - yA_N^{-1}, \\ \frac{1-x-y}{1-y} \in A_N &\iff x \in (1-y)(1-A_N). \end{aligned} \tag{4.59}$$

We can ensure that $1 - A_N\, \cap\, 1 - yA_N^{-1}\, \cap\, (1-y)(1-A_N) \neq \emptyset$ by proving that it has a positive measure. Let $z_0 := \min(y_0, 1-y_0)$. Using simple set operations and the additivity of μ, we can estimate

$$\begin{aligned} &\mu(1 - A_N\, \cap\, 1 - y_0A_N^{-1}\, \cap\, (1-y_0)(1-A_N)) \\ &\qquad\qquad \geq \mu(1 - A_N\, \cap\, 1 - y_0A_N^{-1}\, \cap\, (1-y_0)(1-A_N)\, \cap (0, z_0)) \\ &\qquad \geq z_0 - \underbrace{\mu((0,z_0) \setminus\ 1 - A_N)}_{i} - \underbrace{\mu((0,z_0) \setminus\ 1 - y_0A_N^{-1})}_{ii} - \underbrace{\mu((0,z_0) \setminus\ (1-y)(1-A_N))}_{iii} \end{aligned} \tag{4.60}$$

We proceed by providing a lower estimate for $i)$, $ii)$, and $iii)$. To do this, we use a very special case of [5] Thm. 11.25.

$$i) \quad \int 1_{\mu((0,z_0)\setminus 1-A_N)}(x)d\mu(x) \overset{f(t)=1-t}{=} \int \underbrace{1_{\mu((0,z_0)\setminus 1-A_N)}(1-t)}_{=1_{(1-z_0,1)\setminus A_N}(t)} \underbrace{J(D_t(1-t))}_{=1} dt \\ = \mu((1-z_0,1)\setminus A_N) \le \mu((0,1)\setminus A_N) \tag{4.61}$$

$$ii) \quad \mu((0,z_0)\setminus 1-y_0A_N^{-1}) \overset{i)}{=} \mu((1-z_0,1)\setminus y_0A_N^{-1}) \overset{f(t)=y_0t}{=} y_0\mu((\frac{1-z_0}{y_0},\frac{1}{y_0})\setminus A_N^{-1}) \\ \overset{f(t)=\frac{1}{t}}{\le} y_0\cdot\frac{1}{y_0^2}\mu((y_0,\frac{y_0}{1-z_0})\setminus A_N) \le \frac{1}{y_0}\mu((0,1)\setminus A_N) \tag{4.62}$$

$$iii) \quad \mu((0,z_0)\setminus (1-y_0)(1-A_N)) \overset{f(t)=(1-y_0)t}{=} \mu((0,\frac{z_0}{1-y_0})\setminus 1-A_N) \\ \overset{i)}{\le} \mu((0,1)\setminus A_N) \tag{4.63}$$

If we choose N big enough to have $\varepsilon < \frac{z_0}{2+1/y_0}$, then (4.61), (4.62) and (4.63) lead to

$$\mu(1-A_N \cap 1-y_0A_N^{-1} \cap (1-y_0)(1-A_N)) \ge z_0 - 2\varepsilon - \frac{\varepsilon}{y_0} > 0 \tag{4.64}$$

Notice that the choice of ε affects the value of N according to (4.57).
Since the map

$$y \mapsto \mu(1-A_N \cap 1-yA_N^{-1} \cap (1-y)(1-A_N)), \quad y\in\mathbb{R} \tag{4.65}$$

is continuous [7], there exists an open neighborhood $U \ni y_0$ such that $\forall\, y\in U$, $\mu(1-A_N \cap 1-yA_N^{-1} \cap (1-y)(1-A_N)) > 0$, so the initial claim (4.55) follows. □

We have shown that u is locally bounded on $(0,1)$, this implies that it is also integrable on any compact subset of $(0,1)$. Consider as above $y_0\in(0,1)$ arbitrary but fixed, and $\delta>0$ such that $(y_0-\delta,y_0+\delta)\subset(0,1)$. If we choose $s,t\in\mathbb{R}$ such that $0<s<t<1-y_0-\delta$, then

$$\forall\, y\in(y_0-\delta,y_0+\delta) \qquad 0<\frac{y}{1-s}z<\frac{y}{1-t}<1 \\ 0<\frac{1-y-t}{1-y}<\frac{1-y-s}{1-y}<1 \tag{4.66}$$

Integrating (4.53) with respect to x from s to t, we get

$$(t-s)u(y) = y^{\alpha+1}\int_{\frac{y}{1-s}}^{\frac{y}{1-t}} \frac{u(z)}{z^{\alpha+2}}dz - \int_{1-t}^{1-s} u(z)dz - (1-y)^{\alpha+1}\int_{\frac{1-y-t}{1-y}}^{\frac{1-y-s}{1-y}} u(z)dz \tag{4.67}$$

The right side of (4.67) is continuous at y_0 as a function of y, then u is continuous at y_0 too. Since y_0 was arbitrary, it follows that u is continuous on $(0,1)$. But this implies that the right side of (4.67) is differentiable at y_0 and hence also u is. Iteration of this argument shows that u is differentiable infinitely many times on $(0,1)$.

Proposition 4.6 (Simmetry)**.** *Any measurable solution of (4.53) satisfies* $u(x) = u(1-x)\ \forall\, x \in (0,1)$

To prove this second proposition, we will adopt a more algebraic approach, following [3]. Let $\mathbb{RP}^1$ be the projective line on $\mathbb{R}$. Recall that its points are in bijective correspondence with the 1-dimensional vector subspaces of $\mathbb{R}^2$. In particular, the homogeneous coordinates $(x:y) \in \mathbb{RP}^1$, where $x, y \in \mathbb{R}$, represent the line joining the origin with the point $(x,y) \in \mathbb{R}^2$. Therefore the homogeneous coordinates are defined up to a multiplicative real constant: $(x:y) = (cx:cy),\ \forall\, c \in \mathbb{R}$.

Definition 4.6 (Modular Group)**.** Consider the action of the group $SL_2(\mathbb{Z})$ on $\mathbb{RP}^1$ given by

$$\begin{aligned} \varrho : SL_2(\mathbb{Z}) &\longrightarrow End(\mathbb{RP}^1) \\ A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &\longmapsto A \cdot \_ : (x:y) \mapsto (a_{11}x + a_{21}y : a_{12}x + a_{22}y) \end{aligned} \tag{4.68}$$

Since $Ker(\varrho) = \{-\mathrm{Id}, \mathrm{Id}\}$, is defined a canonical action of the modular group $G := SL_2(\mathbb{Z})/\{-\mathrm{Id}, \mathrm{Id}\}$ on $\mathbb{RP}^1$.

Theorem 4.2. *The group G admits the presentation $G \simeq \langle S, T | S^2 = 1, (ST)^3 = 1 \rangle$, where*

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{4.69}$$

We omit the proof, which can be found in [9].
Define the function

$$\begin{aligned} h : [0,1] &\longrightarrow \mathbb{R} \\ x &\longmapsto u(x) - u(1-x) \end{aligned} \tag{4.70}$$

Observe that for any $x \in [0,1]$, $h(x) = -h(1-x)$, which means that h is antisymmetric around $1/2$. Then we immediately obtain $h(1/2) = 0$. Now let $z \in (1/2, 1)$, so that we can substitute $x = y = 1 - z$ in (4.53); using antisymmetry we get:

$$h(z) = z^\alpha h\left(\frac{2z-1}{z}\right) = z^\alpha h\left(\frac{1-z}{z}\right) \quad for\ 1/2 \le z \le 1 \tag{4.71}$$

Moreover, setting $x = 0$ in (4.53) we get $u(1)\,(1-(1-y)^\alpha) = 0$, then $u(1) = 0$. Similarly, setting $y = 0$, we get $u(0) = 0$. Therefore, the function h is subject to the conditions

$h(0) = h(1) = h(1/2) = 0$. Our goal is proving

$$h(q) = 0, \quad \forall\, q \in \mathbb{Q} \cap [0,1] \tag{4.72}$$

because once we have shown this, the proposition follows because h is continuous and $\mathbb{Q}$ is dense in $\mathbb{R}$. In order to achieve (4.72), we extend the domain of h to the whole real line imposing:

$$h(x) = h(1+x) \quad \forall\, x \in (-\infty, +\infty). \tag{4.73}$$

Lemma 4.2. *For all $x \in \mathbb{R}$, $h(x) = -h(1-x)$*

Proof. We have already observed that $h(x) = -h(1-x)$ holds for $x \in [1/2, 1]$ (and hence for $x \in [0, 1/2]$).
Let $x \in [1, 2]$,

$$h(x) = h(\underbrace{x-1}_{\in[0,1]}) = -h(1-(x-1)) = -h(2-x) = -h(1-x). \tag{4.74}$$

Proceed by induction: suppose that $h(x) = -h(1-x)$ holds for $x \in [n-1, n]$. If $x \in [n, n+1]$, then

$$h(x) = h(\underbrace{\ x-1\ }_{\in[n-1,n]}) = -h(2-x) = -h(1-x). \tag{4.75}$$

□

Lemma 4.3. *The extended function $h : \mathbb{R} \to \mathbb{R}$, satisfies the following equation for all $x \in \mathbb{R}$.*

$$h(x) = |x|^\alpha h\left(\frac{2x-1}{x}\right) \overset{4.2}{=} -|x|^\alpha h\left(\frac{1-x}{x}\right) \tag{4.76}$$

Proof. Equation (4.76) holds for $x \in [1, 2]$:

$$h\left(\frac{1}{x} - 1\right) = h\underbrace{\left(\frac{1}{x}\right)}_{\in[1/2,1]} \overset{4.71}{=} -\left(\frac{1}{x}\right)^\alpha h(x-1) = -\frac{1}{|x|^\alpha} h(x) \tag{4.77}$$

Equation (4.76) holds for $x \in [2, +\infty)$:
By induction; the base case has just been proven. Suppose that (4.76) holds on $[n-1, n]$, then for any $x \in [n, n+1]$ we have

$$\begin{aligned} h(x) = h(x-1) = -(x-1)^\alpha h\left(\frac{1}{1-x} - 1\right) &= -(x-1)^\alpha h\left(\frac{1}{1-x}\right) \\ h\left(2 - \frac{1}{x}\right) = h\left(1 - \frac{1}{x}\right) = -\left(1 - \frac{1}{x}\right)^\alpha h\left(\frac{x}{x-1} - 1\right) &= -\left(\frac{x-1}{x}\right)^\alpha h\left(\frac{1}{x-1}\right) \end{aligned} \tag{4.78}$$

Combining these two equations yields $h(x) = x^\alpha h\left(2 - \frac{1}{x}\right)$.
Equation (4.76) holds for $x \in [0, 1/2)$:

$$h\left(\frac{1}{x} - 1\right) = h\underbrace{\left(\frac{1}{x}\right)}_{\in[2,+\infty)} = -\left(\frac{1}{x}\right)^\alpha h(x-1) = -\frac{1}{|x|^\alpha}h(x) \tag{4.79}$$

Equation (4.76) holds for $x \in (-\infty, 0)$:
It turns out that h is also an odd function, in fact $h(x) = -h(1-x) = -h(-x)$. Then we easily compute

$$h(x) = -h(-x) = +(-x)^\alpha h\left(\frac{1}{-x} - 1\right) = |x|^\alpha h\left(2 - \frac{1}{x}\right) \tag{4.80}$$

This concludes the proof. □

Recall that $\mathbb{R}$ can be seen as a topological subspace of the real projective line, through the section

$$\begin{aligned} j : \mathbb{R} &\to \mathbb{RP}^1 \\ x &\to (x : 1) \end{aligned} \tag{4.81}$$

whose image is $\mathbb{RP}^1 \setminus \{(1 : 0)\}$, the point left is said point at infinity. We called the map j a section because there is also a retraction $e : (x : y) \mapsto \frac{x}{y}$. In this way, any point $(x : y)$ of the projective line can be represented either with a real number $\frac{x}{y}$ or with the additional symbol ∞, if happens that $y = 0$. Observe that the matrices

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.82}$$

are elements of $PGL_2(\mathbb{R})$, which is the group of automorphisms of the projective line. Moreover, composing with j and e, we get

$$\begin{aligned} x \mapsto (x : 1) &\overset{A\cdot}{\longmapsto} (2x - 1 : x) \mapsto \frac{2x-1}{x} \quad x \in \mathbb{R} \setminus \{0\} \\ x \mapsto (x : 1) &\overset{B\cdot}{\longmapsto} (-x + 1 : x) \mapsto \frac{1-x}{x} \quad x \in \mathbb{R} \setminus \{0\} \end{aligned} \tag{4.83}$$

Thus,

$$h(x) = |x|^\alpha h(A \cdot (x : 1)) = -|x|^\alpha h(B \cdot (x : 1)) \quad x \in \mathbb{R} \setminus \{0\}. \tag{4.84}$$

Lemma 4.4. *The matrices A and B^2 generate the group G.*

For the proof we refer to [3].

Lemma 4.5.

The orbit of $(1/2:1)$ *under the action of* G *is* $\mathbb{RP}^1 \cap \mathbb{Q} := \{(p:q) \in \mathbb{RP}^1 | p,q \in \mathbb{Z}\}$.

Proof. Is known that if p,and q are relatively prime numbers, exist integers $n, m \in \mathbb{Z}$ such that $pn + qm = 1$. Thus,

$$det \begin{pmatrix} n & p - \frac{1}{2}n \\ m & q - \frac{1}{2}m \end{pmatrix} = 1 \quad \text{and} \quad \begin{pmatrix} n & p - \frac{1}{2}n \\ m & q - \frac{1}{2}m \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \tag{4.85}$$

This proves that a generic $(p:q)$ belongs to the orbit of $(1/2:1)$ □

From (4.84), we deduce that $h(x) = 0$ if $(x:1)$ can be written as $(x:1) = \omega \cdot (1/2:1)$, where ω is a finite string of symbols from the alphabet $\{A, A^{-1}, B, B^{-1}\}$ such that neither 0 nor ∞ occur in the process of computation of $(x:1)$ starting from $(1/2:1)$.

Lemma 4.6.
For any nonzero rational number r*, exists a finite sequence* $\omega_n\omega_{n-1}\dots\omega_1$ *in the free monoid over* $\{A, A^{-1}, B, B^{-1}\}$*, such that:*

- $\omega_n\omega_{n-1}\dots\omega_1 \cdot (1/2:1) = (r:1)$
- $x_i := \omega_i\omega_{i-1}\dots\omega_1 \cdot (1/2:1)$ *is neither* 0 *nor* ∞*, for all* $i : 1, \dots, n$

Proof. Lemma 4.5 implies the existence of a matrix $M \in G$ such that $(r:1) = M \cdot (1/2:1)$, and by lemma 4.4 we can write $M = \omega_n\omega_{n-1}\dots\omega_1$, for some $n \in \mathbb{N}$ and with the ω_i's in $\{A, A^{-1}, B, B^{-1}\}$. However, it is not excluded that some x_i may be 0 or ∞. In this case, it is possible to eliminate the occurrences of 0 and ∞ by modifying the sequence without affecting the result. Let $\bar{i}$ be the greatest index with the property: $x_i \in \{0, \infty\}$. Notice that $i < n$ because $x_n = r$.
If $x_{\bar{i}} = 0$:

$$x_{\bar{i}+1} = \omega_{\bar{i}+1} \cdot x_i = \begin{cases} 1/2 & if\ \omega_{\bar{i}+1} = A^{-1} \\ 1 & if\ \omega_{\bar{i}+1} = B^{-1} \end{cases} \tag{4.86}$$

In the first case, it is enough to delete from the sequence all the ω_i's up to $\bar{i}+1$, and thus arrive at an expression $r = \omega_n \dots \omega_{\bar{i}+2} \cdot (1/2:0)$ which fulfills the conditions in Lemma 4.6. While in the second case, we can write $r = \omega_n \dots \omega_{\bar{i}+2} B \cdot (1/2:0)$. The cases $\omega_{\bar{i}+1} = A$ and $\omega_{\bar{i}+1} = B$ are not possible because $\bar{i}$ is maximal and both A and B map 0 to ∞.
If $x_{\bar{i}} = \infty$:

$$x_{\bar{i}+1} = \omega_{\bar{i}+1} \cdot x_i = \begin{cases} 2 & if\ \omega_{\bar{i}+1} = A \\ -1 & if\ \omega_{\bar{i}+1} = B \end{cases} \tag{4.87}$$

In the first case, write $r = \omega_n \dots \omega_{\bar{i}+2} BAB^{-1}B^{-1} \cdot (1/2:0)$, while in the second one, write $r = \omega_n \dots \omega_{\bar{i}+2} AAB^{-1}B^{-1} \cdot (1/2:0)$. As above, the cases $\omega_{\bar{i}+1} = A^{-1}$ and $\omega_{\bar{i}+1} = B^{-1}$ are not allowed by maximality of $\bar{i}$. □

This lemma concludes the proof of $h(r) = 0, \ \forall r \in \mathbb{Q}$. As we have already argued, this implies that $u(x) = u(1-x)$ for every $x \in [0,1]$, so the proof of Proposition 4.5 is concluded.
By now, we have shown that any measurable solution u of (4.53) satisfies the conditions:

$$\begin{aligned} &- \ u \in C^\infty((0,1)) \\ &- \ u(0) = 0 = u(1) \\ &- \ u(x) = u(1-x) \ \ \forall x \in [0,1] \end{aligned} \tag{4.88}$$

In particular the third conditions implies that the functional equation in (4.53) coincides with

$$u(x) + (1-x)^\alpha u\left(\frac{y}{1-x}\right) = u(y) + (1-y)^\alpha u\left(\frac{x}{1-y}\right) \tag{4.89}$$

which is the (FEITH).
We are ready to solve the functional equation. Consider first the case $\alpha = 1$, so we start from

$$u(1-x) + (1-x)u\left(\frac{y}{1-x}\right) = u(y) + (1-y)u\left(1 - \frac{x}{1-y}\right). \tag{4.90}$$

Suppose $x, y \in (0,1), x+y \in (0,1)$. Differentiating with respect to x, we get

$$-u'(1-x) + \frac{y}{1-x}u'\left(\frac{y}{1-x}\right) - u\left(\frac{y}{1-x}\right) = -u'\left(1 - \frac{x}{1-y}\right). \tag{4.91}$$

Differentiating (4.91) with respect to y we arrive to

$$\frac{y}{(1-x)^2}u''\left(\frac{y}{1-x}\right) = \frac{x}{(1-y)^2}u''\left(1 - \frac{x}{1-y}\right). \tag{4.92}$$

Consider the substitutions

$$\begin{cases} t = \frac{y}{1-x} \\ s = 1 - \frac{x}{1-y} \end{cases} \iff \begin{cases} y = \frac{ts}{1-t+ts} \\ x = \frac{(1-t)(1-s)}{1-t+ts}. \end{cases} \tag{4.93}$$

As we can see, these substitutions are bijective between the sets

$$\left\{(x,y) \in (0,1)^2 | x+y \in (0,1)\right\} \quad \text{and} \quad (0,1)^2. \tag{4.94}$$

Therefore, performing the substitutions, (4.92) becomes

$$t(1-t)u''(t) = s(1-s)u''(s) \qquad \forall (s,t) \in (0,1) \times (0,1) \tag{4.95}$$

which means that both members are constant, i.e

$$t(1-t)u''(t) = k \quad t \in (0,1), k \in \mathbb{R}. \tag{4.96}$$

We can easily solve the Cauchy problem given by (4.96) together with the boundary conditions $u(0) = u(1) = 0$. In fact by a double integration of (4.96), we obtain

$$u(t) = k\left(t \ln t + (1-t)\ln(1-t)\right) + ct + d \quad c, d \in \mathbb{R} \tag{4.97}$$

and $u(0) = u(1) = 0$ imply that $c = d = 0$, then

$$u(t) = \frac{-k}{\log_2 e} S_1(t) \qquad \forall\, t \in [0,1]. \tag{4.98}$$

We turn now to the case $\alpha \neq 1$. Repeating the steps followed in the previous case leads to the differential equation

$$s^{2-\alpha}\left[(1-s)u''(s) + (\alpha-1)u'(s)\right] = (1-t)^{2-\alpha}\left[tu''(t) + (1-\alpha)u'(t)\right] \quad \forall\,(s,t) \in (0,1)^2 \tag{4.99}$$

Unlike the previous case, the two sides of this equality do not have the same form. However, fixing $s = \bar{s}$ and letting t vary, we note that the second side is constant in t, to the real number $k = \bar{s}^{2-\alpha}\left[(1-\bar{s})u''(\bar{s}) + (\alpha-1)u'(\bar{s})\right]$. Thereafter, letting s vary, we see that the first side is also constant with value k. Hence (4.99) is equivalent to the system

$$\begin{cases} s^{2-\alpha}\left[(1-s)u''(s) + (\alpha-1)u'(s)\right] = k & \forall\, s \in (0,1) \\ (1-t)^{2-\alpha}\left[tu''(t) + (1-\alpha)u'(t)\right] = k & \forall\, t \in (0,1) \end{cases} \tag{4.100}$$

Nevertheless, if we perform the bijective substitution $t = 1 - s$ in the first equation, we obtain the second one, and vice versa. It follows that these two equations are equivalent to each other, so we can consider only one of them and solve it.

$$tu''(t) + (1-\alpha)u'(t) = k(1-t)^{\alpha-2} \qquad k \in \mathbb{R}, \forall\, t \in (0,1) \tag{4.101}$$

Through an integration, using in particular integration by parts on the first term, we arrive to a linear ordinary differential equation of the first order.

$$u'(t) - \frac{\alpha}{t}u(t) = \frac{1}{t}\left(\frac{k}{\alpha-1}(1-t)^{\alpha-1} + c\right) \qquad c \in \mathbb{R} \tag{4.102}$$

The general solution of this equation is given by

$$\begin{aligned} u(t) &= e^{\int -\frac{\alpha}{z}dz}\left[c_1 + \int \frac{1}{t}\left(\frac{k}{\alpha-1}(1-t)^{\alpha-1} + c\right)e^{-\int -\frac{\alpha}{z}dz}dt\right] \\ &= t^\alpha c_1 + \frac{k}{\alpha(\alpha-1)}(1-t)^\alpha - \frac{c}{\alpha} \qquad c_1 \in \mathbb{R}. \end{aligned} \tag{4.103}$$

Imposing the boundary conditions $u(0) = u(1) = 0$, we obtain

$$u(t) = \frac{k}{\alpha(\alpha-1)}t^\alpha + \frac{k}{\alpha(\alpha-1)}(1-t)^\alpha - \frac{k}{\alpha(\alpha-1)} = \frac{k}{\alpha(\alpha-1)}S_\alpha(t) \quad \forall\, t \in (0,1). \tag{4.104}$$

4.4 Local structure of 1-cocycles

Let us come back to Example (4.2). We saw that if f is a 1-cocycle, then $f[X](p, 1-p) = f[Y](1-p, p)$ is a measurable function in the variable $p \in [0,1]$ which satisfies a functional equation like (4.53). Thus there exists a real number λ such that $f[X] = \lambda S_\alpha[X]$ and $f[Y] = \lambda S_\alpha[Y]$. Moreover, the chain rule for α-entropies entails

$$f[XY](P) = \lambda S_\alpha[X](X_*P) + X.\lambda S_\alpha[Y](P) = \lambda S_\alpha[XY](P) \qquad \forall\, P \in \mathscr{P}_{XY}, \tag{4.105}$$

then $f = \lambda S_\alpha$. Since f was arbitrary, we obtain $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) = \mathbb{R}\cdot S_\alpha$. In the next two subsections we try to generalize these results to a generic finite information structure $(\mathbf{S}, \mathscr{E})$ equipped with a generic probability functor $\mathscr{Q}$. In particular, in this subsection we will prove that, given a diagram $X \xleftarrow{\pi_X} XY \xrightarrow{\pi_Y} Y$ in $\mathbf{S}$, and under certain conditions on $\mathscr{E}$ and $\mathscr{Q}$, the cocycle condition and the functional equation (4.53) determine the form of the components in X,Y and XY of the 1-cocycles, up to a multiplicative constant. Some nontrivial additional conditions are needed, in fact in Example 4.1 we could have considered the product of $\perp$ with itself, but, as we have shown, the component in $\perp$ of 1-cocycles is free to vary among all the measurable functions $\Delta(1) \to \mathbb{R}$. On the other hand, the conditions $\mathscr{E}_{XY} = \mathscr{E}_X \times \mathscr{E}_Y$ and $\mathscr{Q} = \mathscr{P}$ turn out to be too restrictive. We present here a weakened version but still sufficient to prevent the degeneration of the cocycle condition. In order to simplify notation, we will identify the set $\mathscr{E}_{XY}$ with a subset of $\mathscr{E}_X \times \mathscr{E}_Y$ through the canonical injection mentioned in Definition 2.2.

Definition 4.7. Consider $(\mathbf{S}, \mathscr{E})$ a finite information structure and $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$ an adapted probability functor. Let $X, Y \in Ob(\mathbf{S})$ such that $\mathbf{S}$ admits the product XY. Denote $k := |\mathscr{E}_X|$ and $l := |\mathscr{E}_Y|$. The product XY is said *nondegenerate* if $k, l \geq 2$, and exist numerations $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_l\}$ of $\mathscr{E}_X$ and $\mathscr{E}_Y$ together with a north-east lattice path[21] $(\gamma_i)_{i=1}^{k+l-3}$ on $\mathbb{Z}^2$, starting from $(1,1)$ to $(k-1, l-1)$ such that:

[21] Is a sequence of points of $\mathbb{Z}^2$ such that $\gamma_1 = (1,1)$, $\gamma_{k+l-3} = (k-1, l-1)$ and $\gamma_{i+1} - \gamma_i \in \{(1,0),(0,1)\}$ for every $i : 1, \dots, k+l-3$

1. If, for some $i \in \{1, 2, \ldots, k+l-4\}$, $\gamma_i = (a, b)$ and $\gamma_{i+1} = (a+1, b)$, then for every $P \in \mathscr{Q}_X$ with support in $\{x_j | a \le j \le k\}$ there exists a probability $\tilde{P} \in \mathscr{Q}_{XY}$ such that $\mathscr{Q}(\pi_X)(\tilde{P}) = P$ and

$$supp(\tilde{P}) \subset \{(x_a, y_b)\} \cup \{(x_j, y_{b+1}) | a+1 \le j \le k\} \tag{4.106}$$

or

$$supp(\tilde{P}) \subset \{(x_a, y_{b+1})\} \cup \{(x_j, y_b) | a+1 \le j \le k\} \tag{4.107}$$

2. If, for some $i \in \{1, 2, \ldots, k+l-4\}$, $\gamma_i = (a, b)$ and $\gamma_{i+1} = (a, b+1)$, then for every $P \in \mathscr{Q}_Y$ with support in $\{x_j | b \le j \le l\}$, there exists a probability $\tilde{P} \in \mathscr{Q}_{XY}$ such that $\mathscr{Q}(\pi_Y)(\tilde{P}) = P$ and

$$supp(\tilde{P}) \subset \{(x_a, y_b)\} \cup \{(x_{a+1}, y_j) | b+1 \le j \le l\} \tag{4.108}$$

or

$$supp(\tilde{P}) \subset \{(x_{a+1}, y_b)\} \cup \{(x_a, y_j) | b+1 \le j \le l\} \tag{4.109}$$

3. If, for some $i \in \{1, 2, \ldots, k+l-3\}$, $\gamma_i = (a, b)$, then exist at least three elements $\{z_1, z_2, z_3\}$ in $\{x_a, x_{a+1}\} \times \{y_b, y_{b+1}\}$ such that $\mathscr{Q}_{XY}$ contains all the possible probability laws with support in $\{z_1, z_2, z_3\}$. Or equivalently, the convex hull $[\delta_{z_1}, \delta_{z_2}, \delta_{z_3}]$ is a subset of $\mathscr{Q}_{XY}$

According to this definition, the product of any object with itself is always degenerate: as an instance, the third condition can not be verified since $\mathscr{E}_{XX} = \mathscr{E}_X \simeq \{(x, x) | x \in \mathscr{E}_X\} \subset \mathscr{E}_X \times \mathscr{E}_X$, so for any (a, b), the set $\{x_a, x_{a+1}\} \times \{y_b, y_{b+1}\}$ has only two elements. We remark also that if a product is nondegenerate, then both $\mathscr{Q}(\pi_X)$ and $\mathscr{Q}(pi_Y)$ must be surjective because of the first and the second condition.

Theorem 4.3. *Let* $(\mathbf{S}, \mathscr{E})$ *be a finite information structure with* $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$*, an adapted probability functor. Consider a generic* $f \in Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$*, where* α *can be any positive real number. Let* $X, Y \in Ob(\mathbf{S})$ *such that* XY *is a nondegenerate product. Then there exist* $\lambda \in \mathbb{R}$ *such that*

$$f[X] = \lambda S_\alpha[X] \quad f[Y] = \lambda S_\alpha[Y] \quad f[XY] = \lambda S_\alpha[XY] \tag{4.110}$$

Proof. We introduce firstly a notation for probability laws. Le P be a probability on a finite set $A = \{a_1, a_2, \ldots, a_n\}$. Then we will denote

$$P = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ p_1 & p_2 & \ldots & p_n \end{pmatrix} \quad \text{if } \; P(a_i) = p_i \; \forall \, i : 1, \ldots, n \tag{4.111}$$

with the convention that if some element of A does not appear in the matrix, then its probability of occurrence is zero. The cocycle condition (4.42) entails

$$f[X] - Y.f[X] = f[Y] - Y.f[Y]. \tag{4.112}$$

Consider now the enumerations $\{x_1, \ldots, x_k\}$, $\{y_1, \ldots, y_l\}$ of $\mathscr{E}_X$ and $\mathscr{E}_Y$ and the lattice path $(\gamma_i)_i$ given by Definition 4.36. There must exist an index i_1 and an integer b_1 between 1 and $l-1$, such that $\gamma_{i_1} = (1, b_1)$ and $\gamma_{i_1+1} = (2, b_1)$. Thus for any

$$P = \begin{pmatrix} x_1 & \ldots & x_k \\ p_1 & \ldots & p_k \end{pmatrix} \in \mathscr{Q}_X,$$

the first point of Definition 4.36 provides a probability law $\tilde{P} \in \mathscr{Q}_{XY}$ such that

$$supp(\tilde{P}) \subset \{(x_1, y_{b_1})\} \cup \{(x_j, y_{b_1+1}) | 2 \leq j \leq k\} \tag{4.113}$$

or

$$supp(\tilde{P}) \subset \{(x_1, y_{b_1+1})\} \cup \{(x_j, y_{b_1}) | 2 \leq j \leq k\}, \tag{4.114}$$

and $X_*\tilde{P} = P$. The equation (4.112), when applied to $\tilde{P}$ reads:

$$\begin{aligned} f[X] \begin{pmatrix} x_1 & \ldots & x_k \\ p_1 & \ldots & p_k \end{pmatrix} =& f[Y]\sigma \begin{pmatrix} y_{b_1} & y_{b_1+1} \\ p_1 & 1-p_1 \end{pmatrix} \\ &+ (1-p_1)^\alpha f[X] \begin{pmatrix} x_2 & \ldots & x_k \\ p_2/(1-p_1) & \ldots & p_k/(1-p_1) \end{pmatrix} \end{aligned} \tag{4.115}$$

where σ is an element of the symmetric group Σ_2. The permutation σ acts on the position of the symbols p_1 and $1-p_1$ in the bottom row of the matrix. Its presence is due to the fact that it is not known a priori in which of the two possible sets the support of $\tilde{P}$ is contained. Observe that for all $x \in \mathscr{E}_X$, the conditional probability $\tilde{P}_{|X=x}$ is atomic, thus the term $-X.f[Y](\tilde{P})$ vanishes. The same holds for one of the two probabilities $\tilde{P}_{|Y=y_{b_1}}$ and $\tilde{P}_{|Y=y_{b_1+1}}$. Which one depends on the support of $\tilde{P}$, however this does not affect the form of (4.115). Furthermore, by point 3. of Definition 4.36, we can find a probability $p \in \mathscr{Q}_{XY}$, with support contained in $\{x_1, x_{2+1}\} \times \{y_{b_1}, y_{b_1+1}\}$ and such that

$$X_*p = \begin{pmatrix} x_1 & x_2 \\ p_1 & 1-p_1 \end{pmatrix}. \tag{4.116}$$

Now applying point 2., we obtain a probability $\tilde{X_*p} \in \mathscr{Q}_{XY}$ which, once inserted in (4.112), yields

$$f[X] \begin{pmatrix} x_1 & x_2 \\ p_1 & 1-p_1 \end{pmatrix} = f[Y]\sigma \begin{pmatrix} y_{b_1} & y_{b_1+1} \\ p_1 & 1-p_1 \end{pmatrix}. \tag{4.117}$$

Combining with (4.115) we get

$$f[X]\begin{pmatrix} x_1 & \dots & x_k \\ p_1 & \dots & p_k \end{pmatrix} = f[X]\begin{pmatrix} x_1 & x_2 \\ p_1 & 1-p_1 \end{pmatrix} + (1-p_1)^{\alpha} f[X]\begin{pmatrix} x_2 & \dots & x_k \\ p_2/(1-p_1) & \dots & p_k/(1-p_1) \end{pmatrix} \tag{4.118}$$

The computations that led to this equation can be repeated for each increase in the x-coordinate of the point γ_i in the north-east lattice path. The formula corresponding to the passage from (a, b) to $(a + 1, b)$, for a generic $1 < a \leq k - 2$, would be

$$\begin{aligned} f[X]\begin{pmatrix} x_a & \dots & x_k \\ p_a & \dots & p_k \end{pmatrix} =& f[X]\begin{pmatrix} x_a & x_{a+1} \\ p_a & 1-p_a \end{pmatrix} \\ &+ (1-p_a)^{\alpha} f[X]\begin{pmatrix} x_{a+1} & \dots & x_k \\ p_{a+1}/(1-p_a) & \dots & p_k/(1-p_a) \end{pmatrix} \end{aligned} \tag{4.119}$$

for any probability law in $\mathscr{Q}_X$ whose support is contained in $\{x_j \in \mathscr{E}_X | a \leq j \leq k\}$. Then we substitute sequentially these equations in (4.118) so that $f[X]$ results determined by the value it takes on probabilities with support in a couple of consecutive points (according to the fixed ordering) of $\mathscr{E}_X$. We present e.g. the first step of this process: using (4.119) with $a = 2$, we get

$$\begin{aligned} f[X]\begin{pmatrix} x_2 & \dots & x_k \\ \frac{p_2}{1-p_1} & \dots & \frac{p_k}{1-p_1} \end{pmatrix} =& f[X]\begin{pmatrix} x_2 & x_3 \\ \frac{p_2}{1-p_1} & \frac{1-p_1-p_2}{1-p_1} \end{pmatrix} \\ &+ (\frac{1-p_1-p_2}{1-p_1})^{\alpha} f[X]\begin{pmatrix} x_3 & \dots & x_k \\ \frac{p_3}{1-p_1-p_2} & \dots & \frac{p_k}{1-p_1-p_2} \end{pmatrix}, \end{aligned} \tag{4.120}$$

thus after the substitution, (4.118) becomes

$$\begin{aligned} f[X]\begin{pmatrix} x_1 & \dots & x_k \\ p_1 & \dots & p_k \end{pmatrix} = f[X]\begin{pmatrix} x_1 & x_2 \\ p_1 & 1-p_1 \end{pmatrix} &+ f[X]\begin{pmatrix} x_2 & x_3 \\ \frac{p_2}{1-p_1} & \frac{1-p_1-p_2}{1-p_1} \end{pmatrix} \\ &+ (\frac{1-p_1-p_2}{1-p_1})^{\alpha} f[X]\begin{pmatrix} x_3 & \dots & x_k \\ \frac{p_3}{1-p_1-p_2} & \dots & \frac{p_k}{1-p_1-p_2} \end{pmatrix} \end{aligned} \tag{4.121}$$

When $a = k - 1$ is reached, the process terminates, and $f[X]$ is expressed by

$$f[X]\begin{pmatrix} x_1 & \dots & x_k \\ p_1 & \dots & p_k \end{pmatrix} = \sum_{j=1}^{k-1} \left(1 - \sum_{r=1}^{j-1} p_r\right)^{\alpha} f[X]\begin{pmatrix} x_j & x_{j+1} \\ \frac{p_j}{1-\sum_{r=1}^{j-1} p_r} & \frac{1-\sum_{r=1}^{j} p_r}{1-\sum_{r=1}^{j-1} p_r} \end{pmatrix} \tag{4.122}$$

Observe that the definition of nondegenerate product remains the same if the roles of X and Y are reversed. Therefore, we can compute the value of $f[Y]$ on $Q \in \mathscr{Q}_Y$ in complete analogy to what was done for $f[X]$.

We are left to determinate the functions

$$\begin{aligned} \varphi_a : [0,1] &\to \mathbb{R} & \psi_b : [0,1] &\to \mathbb{R} \\ p &\mapsto f[X]\begin{pmatrix} x_a & x_{a+1} \\ p & 1-p \end{pmatrix} & p &\mapsto f[Y]\begin{pmatrix} y_b & y_{b+1} \\ p & 1-p \end{pmatrix}. \end{aligned} \tag{4.123}$$

If the north-east lattice path passes through the point (a, b), then condition 3. ensures the existence of all probabilities $p \in \mathscr{Q}_{XY}$ whose support is contained in a certain subset $\{z_1, z_2, z_3\}$ of $\{x_a, x_{a+1}\} \times \{y_b, y_{b+1}\}$. Among these three elements, there must be one, say z_X, whose first component is different from that of the other two. Similarly, another element z_Y, different from z_X, must have the second component different from that of the other two. Denote $\mu_X := p(z_X)$ and $\mu_Y := p(z_Y)$. The equation (4.42) reads

$$\begin{aligned} &\forall \mu_X, \mu_Y \in [0,1) \;\; s.t. \;\; \mu_X + \mu_Y \in [0,1] \\ &f[X]\sigma\begin{pmatrix} x_a & x_{a+1} \\ \mu_X & 1-\mu_X \end{pmatrix} + (1-\mu_X)^\alpha f[Y]\sigma'\begin{pmatrix} y_b & y_{b+1} \\ \frac{\mu_Y}{1-\mu_X} & \frac{1-\mu_X-\mu_Y}{1-\mu_X} \end{pmatrix} = \\ &f[Y]\sigma'\begin{pmatrix} y_b & y_{b+1} \\ \mu_Y & 1-\mu_Y \end{pmatrix} + (1-\mu_Y)^\alpha f[X]\sigma\begin{pmatrix} x_a & x_{a+1} \\ \frac{\mu_X}{1-\mu_Y} & \frac{1-\mu_X-\mu_Y}{1-\mu_Y} \end{pmatrix} \end{aligned} \tag{4.124}$$

where $\sigma, \sigma' \in \Sigma_2$ account the missing knowledge about which element among $\{x_a, x_{a+1}\}$ (respectively among $\{y_b, y_{b+1}\}$) appears only once. This equation is equivalent to (4.50), thus we can conclude that exists a real number $\lambda_{(a,b)}$ such that

$$f[X]\sigma\begin{pmatrix} x_a & x_{a+1} \\ p & 1-p \end{pmatrix} = f[Y]\sigma'\begin{pmatrix} y_b & y_{b+1} \\ 1-p & p \end{pmatrix} = \lambda_{(a,b)} S_\alpha(p, 1-p) \quad \forall\, p \in [0,1] \tag{4.125}$$

Since the α-entropy is symmetric in its arguments, we can drop the symbols σ and σ'. Therefore, for any pair $(a, b) \in \mathbb{Z}^2$ that belongs to the north east lattice path, we found

$$\varphi_a(p) = \psi_b(p) = \lambda_{(a,b)} S_\alpha(p, 1-p) \quad \forall\, p \in [0,1]. \tag{4.126}$$

It is important to prove that $\lambda_{(a,b)}$ is constant in (a, b), i.e. along the lattice path. To achieve this, suppose $\gamma_i = (a, b)$; If $\gamma_{i+1} = (a + 1, b)$, then

$$\begin{aligned} &\begin{pmatrix} (x_{a+1}, y_b) & (x_{a+2}, y_{b+1}) \\ p & 1-p \end{pmatrix} \in \mathscr{Q}_{XY}, \; \forall p \in [0,1] \\ or \; &\begin{pmatrix} (x_{a+1}, y_{b+1}) & (x_{a+2}, y_b) \\ p & 1-p \end{pmatrix} \in \mathscr{Q}_{XY}, \; \forall p \in [0,1] \end{aligned} \tag{4.127}$$

by point 3. of Definition 4.7. Applying the cocycle condition (4.42) to these probabilities, we get, $\forall\, p \in [0,1]$

$$\lambda_{(a+1,b)} S_\alpha(p, 1-p) = f[X] \begin{pmatrix} x_{a+1} & x_{a+2} \\ p & 1-p \end{pmatrix} = f[Y]\sigma \begin{pmatrix} y_b & y_{b+1} \\ p & 1-p \end{pmatrix} = \lambda_{(a,b)} S_\alpha(p, 1-p) \tag{4.128}$$

This implies that $\lambda_{(a+1,b)} = \lambda_{(a,b)}$. Moreover, $\lambda_{(a,b+1)} = \lambda_{(a,b)}$ is proven in the same way. Hence, we can define $\lambda := \lambda_{(1,1)} = \cdots = \lambda_{(k-1,l-1)}$. Combining equations (4.122) and (4.126), we arrive to

$$\begin{aligned} f[X] \begin{pmatrix} x_1 & \dots & x_k \\ p_1 & \dots & p_k \end{pmatrix} &= \sum_{j=1}^{k-1} \left(1 - \sum_{r=1}^{j-1} p_r\right)^\alpha \lambda S_\alpha \left(\frac{p_j}{1 - \sum_{r=1}^{j-1} p_r}, \frac{1 - \sum_{r=1}^{j} p_r}{1 - \sum_{r=1}^{j-1} p_r} \right) \\ &= \begin{cases} \lambda \sum_{i=1}^{k} -p_j \log_2(p_j) & if\ \alpha = 1 \\ \lambda \sum_{i=1}^{k} p_j^\alpha - 1 & if\ \alpha \neq 1 \end{cases} \end{aligned} \tag{4.129}$$

The last equality can be proven by a direct calculation, but it is also a consequence of the chain rule.
This shows that for any $\alpha > 0$, there exist a multiplicative constant λ such that $f[X] = \lambda S_\alpha[X]$ and $f[Y] = \lambda S_\alpha[Y]$, so the proof can be concluded as in (4.105). □

Remark 4.2. The proof suggests that having enumerations of $\mathscr{E}_X$ and $\mathscr{E}_Y$ such that there exists a north-east lattice path (with the properties described in Definition 4.7) which reaches $k-1$ on the coordinate x, is enough to infer $f[X] = \lambda S_\alpha[X]$, for some $\lambda \in \mathbb{R}$. The determination of $f[Y]$ can be carried out similarly, with enumerations that may be different from those used for $f[X]$, but such that there exists a north-east lattice path that reaches $k-1$ on the y-axis. However, the multiplicative constant λ' such that $f[Y] = \lambda' S_\alpha[Y]$ is in general different from λ. This problem can be overcome for example assuming that there exists $(a,b) \in \{1, \dots k-1\} \times \{1, \dots, l-1\}$ such that

$$\mathscr{Q}_{XY} \supset \mathscr{P}(\{(x_a, y'_b), (x_{a+1}, y'_{b+1})\}) \ \ or\ \ \mathscr{Q}_{XY} \supset \mathscr{P}(\{(x_a, y'_{b+1}), (x_{a+1}, y'_b)\}), \tag{4.130}$$

where x_a denotes the a-th element in the numeration of $\mathscr{E}_X$ related to $f[X]$ and y'_b denotes the b-th element in the numeration of $\mathscr{E}_Y$ related to $f[Y]$. Indeed, with this hypothesis , we can use (4.42) to derive

$$f[X] \begin{pmatrix} x_a & x_{a+1} \\ p & 1-p \end{pmatrix} = f[Y]\tau \begin{pmatrix} y'_b & y'_{b+1} \\ p & 1-p \end{pmatrix} \quad \forall\, p \in [0,1]. \tag{4.131}$$

Then $\lambda = \lambda'$ follows. Therefore, these modified hypotheses can replace the condition "XY is nondegenerate" in Theorem 4.3.

4.5 $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$

As previously mentioned, in this subsection we will prove that, under suitable assumptions, every 1-cocycle of the complex $C^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ is locally a multiple of the cocycle defined by the α-entropy, and consequently we will determine the first cohomology group.

Definition 4.8. Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure. An object Z is called *reducible* if there exist objects $X, Y \in Ob(\mathbf{S})$, such that $Z = XY$, otherwise Z is said to be *irreducible.* Moreover, we will use the term *non-trivially reducible* to indicate that the product XY is non-degenerate (with reference to a previously fixed probability functor).

Definition 4.9. Given an information structure $(\mathbf{S}, \mathscr{E})$ and an adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$, we define $\mathbf{S}^*_{\mathscr{Q}}$ as the full subcategory of $\mathbf{S}$, such that

$$Ob(\mathbf{S}^*_{\mathscr{Q}}) = \{X \in Ob(\mathbf{S}) |\, \exists P \in \mathscr{Q}(X) \text{ non atomic}\}.$$

Recall that a category is connected if for every two objects there is a finite sequence of arrows (not necessarily composable) which connects them. Moreover, any category $\mathcal{C}$ is a disjoint union (the coproduct in **Cat**) of connected categories, called connected components. Denote $\pi_0(\mathcal{C})$ the set of connected components of $\mathcal{C}$.

Theorem 4.4. *Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure equipped with an adapted probability functor $\mathscr{Q} : \mathbf{S} \to \mathbf{Meas}$ whose restriction maps (internal marginalizations) are all surjective. Assume that for any object $X \in Ob(\mathbf{S})$, exists an arrow $Z \to X$ in $\mathbf{S}$ such that Z is non-trivially reducible. Then, for each $\alpha > 0$, there is an isomorphism of real vector spaces*

$$\begin{gathered}\chi_\alpha : \prod_{C \in \pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R} \longrightarrow Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \quad \textit{given by} \\ \chi((\lambda_C)_{C \in \pi_0(\mathbf{S}^*_{\mathscr{Q}})})[X] = \lambda_C S_\alpha[X] \quad \forall C \in \pi_0(\mathbf{S}^*), \forall\, X \in C\end{gathered} \tag{4.132}$$

Under this isomorphism, the subgroup of 1-coboundaries is identified with the diagonal subspace $\Delta = \mathbb{R} \cdot (1, \dots, 1, \dots)$ if $\alpha \neq 1$, otherwise is the zero subspace. Therefore,

$$H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) = \begin{cases} \prod_{\pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R} & if \ \ \alpha = 1 \\ \left(\prod_{\pi_0(\mathbf{S}^*_{\mathscr{Q}})} \mathbb{R}\right) / \Delta & if \ \ \alpha \neq 1 \end{cases} \tag{4.133}$$

Proof. Observe first that if $W \in Ob(\mathbf{S}) \setminus Ob(\mathbf{S}^*_{\mathscr{Q}})$, then for any 1-cocycle f, it holds that $f[W] \equiv 0$. Hence to characterize a 1-cocycle f, is sufficient to specify the functions $f[X]$ for all $X \in Ob(\mathbf{S}^*_{\mathscr{Q}})$. Now pick one of those objects X, by hypothesis there exist $t : Z \to X$ with Z non-trivially reducible. Consider the diagram $Z \xleftarrow{id} ZX \xrightarrow{t} X$; since $XZ = Z$, the cocycle condition (4.42) entails

$$f[X] = f[Z] - X.f[Z]. \tag{4.134}$$

In Theorem 4.3 we proved that exists a real number, say λ_Z, such that $f[Z] = \lambda_Z S_\alpha[Z]$. Moreover, our assumptions on $\mathscr{Q}$ ensure that for any $P \in \mathscr{Q}(X)$ there is a probability $\bar{P} \in \mathscr{Q}(Z)$ such that $X_*\bar{P} = P$. Applying the equation 4.134 to $\bar{P}$ we get

$$\begin{aligned} f[X](P) &= \lambda_Z S_\alpha[Z](\bar{P}) - \sum_{x \in \mathscr{E}_X} P(x)\lambda_Z S_\alpha[Z](\bar{P}_{|X=x}) \\ &= \lambda_Z \left(S_\alpha(\bar{P}) - \sum_{x \in \mathscr{E}_X} X_*\bar{P}(x) S_\alpha(\bar{P}_{|X=x}) \right) = \lambda_Z S_\alpha(X_*\bar{P}) = \lambda_Z S_\alpha[X](P) \end{aligned} \tag{4.135}$$

for all $P \in \mathscr{Q}(X)$. This implies $f[X] = \lambda_Z S_\alpha[X]$. In case there is another non-trivially reducible object Z' with a map on X, then exists another constant $\lambda_{Z'}$ such that

$$\lambda_Z S_\alpha[X] = f[X] = \lambda_{Z'} S_\alpha[X]. \tag{4.136}$$

Since X is an object in $\mathbf{S}^*_{\mathscr{Q}}$, is possible to find a probability $P \in \mathscr{Q}(X)$ with nonzero entropy, then equation (4.136) yields $\lambda_Z = \lambda_{Z'}$. Now consider two objects $X_1, X_2 \in Ob(\mathbf{S}^*_{\mathscr{Q}})$ lying in the same connected component C. This means that there is a zig-zag diagram in $\mathbf{S}^*_{\mathscr{Q}}$ like

$$X_1 \longrightarrow Y_1 \longleftarrow Y_2 \longrightarrow \ldots \longleftarrow Y_n \longrightarrow X_2. \tag{4.137}$$

Any object of this sequence is the target of a map from a non-trivially reducible object.

$$\begin{array}{ccccccccc} X_1 & \longrightarrow & Y_1 & \longleftarrow & Y_2 & \longrightarrow \ldots \longleftarrow & Y_n & \longrightarrow & X_2. \\ \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\ Z_1 & & H_1 & & H_2 & & H_n & & Z_2 \end{array} \tag{4.138}$$

This diagram can be read as

$$Z_1 \longrightarrow Y_1 \longleftarrow H_1 \longrightarrow Y_1 \longleftarrow H_2 \longrightarrow \ldots \longleftarrow H_n \longrightarrow X_1 \longleftarrow Z_2$$

from which becomes clear that $\lambda_{Z_1} = \lambda_{H_1} = \lambda_{H_2} = \cdots = \lambda_{H_n} = \lambda_{Z_2}$. Thus we can define λ_C to be the constant such that $f[X] = \lambda_C S_\alpha[X]$ for all X in C. Observe that if two objects X_1, X_2 lie in different connected components C_1, C_2, it is not possible to use the cocycle condition to derive relations between $f[X_1]$ and $f[X_2]$, so λ_{C_1} and λ_{C_2} are independent. Then the function χ is actually a bijection. $\square$

Corollary 4.2. *Let* $\{(\mathbf{S}_i, \mathscr{E}_i)\}_{i\in I}$ *be a family of information structures, each of them equipped with an adapted probability functor* $\mathscr{Q}_i$ *so that the hypotheses of the previous theorem are verified separately for all* $i \in I$*. Denote* $\bigsqcup_{i\in I} \mathscr{Q} : \bigsqcup_{i\in I} \mathbf{S}_i \to \mathbf{Meas}$ *the functor that coincides with* $\mathscr{Q}_i$ *on* $\mathbf{S}_i$*. Then this functor is an adapted probability functor on the coproduct of the* $(\mathbf{S}_i, \mathscr{E}_i)$ *in* ***InfoStr*** *and*

$$Z^1(\bigsqcup_{i\in I} \mathbf{S}_i, \mathscr{F}_\alpha(\bigsqcup_{i\in I} \mathscr{Q})) \simeq \prod_{i\in I} Z^1(\mathbf{S}_i, \mathscr{F}_\alpha(\mathscr{Q}_i)) \tag{4.139}$$

Proof. The pair $\left(\bigsqcup_{i\in I} \mathbf{S}_i, \bigsqcup_{i\in I} \mathscr{Q}\right)$ satisfies the hypotheses of Theorem 4.4 because they are satisfied by each pair $(\mathbf{S}_i, \mathscr{Q}_i)$. Moreover, the category $(\bigcup_{i\in I} \mathbf{S}_i)^*$ is the disjoint union of $(\mathbf{S}_i)^*$. Now (4.139) follows applying Theorem 4.4. □

Let us now examine the case in which the hypotheses of the Theorem 4.4 are not satisfied. If the structure $(\mathbf{S}, \mathscr{E})$ is bounded, asking that any object is the target of an arrow starting from a non-trivially reducible object, is equivalent to asking that all the minimal objects of the poset $\mathbf{S}$ are non-trivially reducible. In the next theorem we will prove that if one of them is irreducible, then $H^1(\mathbf{S}, \mathscr{F}_\alpha)$ has infinite dimensions as a vector space on $\mathbb{R}$.

Theorem 4.5. *Let* $(\mathbf{S}, \mathscr{E})$ *be a bounded, finite information structure such that* $\mathscr{E}$ *is conservative. If there exists an object in* $\mathbf{S}$ *which is irreducible and minimal , then*

$$dim_{\mathbb{R}} H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) = \infty \tag{4.140}$$

Proof. Call M the irreducible minimal object in question. We claim that the subcategory of $\mathbf{S}$ whose set of objects is

$$T = \{X \in Ob(\mathbf{S}) | \exists\, \pi : M \to X,\ M \neq X\} = \mathscr{S}_M \setminus \{M\} \tag{4.141}$$

has a minimum $\tilde{X}$. By contradiction, if there were two distinct minimal objects X_1 and X_2, then the product X_1X_2, which of course exists in $\mathbf{S}$, could not belong to T, so we would have $M = X_1X_2$, in contradiction with the irreducibility of M. Consider $f \in Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$ such that $f[Y] \equiv 0$ for all $Y \in Ob(\mathbf{S})$, $Y \neq M$. Note that if M was reducible, the cocycle condition (4.42) would entail $f[M] \equiv 0$. But in this case, the only nontrivial relation that we can derive from (4.42) is

$$f[M] = \underbrace{f[\tilde{X}]}_{\equiv 0} + \tilde{X}.f[M], \tag{4.142}$$

obtained applying (4.42) to the product $M\tilde{X} = M$. Thus, for every $P \in \mathscr{P}_M$,

$$f[M](P) = \sum_{x\in\mathscr{E}_{\tilde{X}}} P(\tilde{X} = x) f[M](P_{|\tilde{X}=x}). \tag{4.143}$$

Since $\mathscr{E}$ is conservative, the map $\mathscr{E}(t : M \to \tilde{X}) : \mathscr{E}_M \twoheadrightarrow \mathscr{E}_{\tilde{X}}$ is not an isomorphism. Then exists at least one $\bar{x} \in \mathscr{E}_{\tilde{X}}$ such that $m := |\mathscr{E}(t)^{-1}(\bar{x})| > 1$. The conditional probability $P_{|\tilde{X}=\bar{x}}$ is an element of $\Delta(m)$ seen as a face of $\Delta(|\mathscr{E}_M|)$. Thus we can choose any measurable function $g : \Delta(m) \to \mathbb{R}$ that vanishes on the vertices (because of the condition $M.f[M] = 0$) and define $f[M](P_{|\tilde{X}=\bar{x}}) := g(P_{|\tilde{X}=\bar{x}})$. This yields infintely many linearly independent cocycles. $\square$

We close this section with some examples concerning information structure where all minimal objects are reducible, but not necessarily can be written as a non-degenerate product.

Example 4.3. Let $(\mathbf{S}, \mathscr{E})$ be an information structure whose underlying poset is

$$\begin{array}{ccccc} & & X & & \\ & \overset{\pi_X}{\nearrow} & & \overset{\tau_X}{\searrow} & \\ XY & & & & T \\ & \underset{\pi_Y}{\searrow} & & \underset{\tau_Y}{\nearrow} & \\ & & Y & & \end{array} \tag{4.144}$$

as in Example 4.2. The functor $\mathscr{E}$ is specified by $\mathscr{E}_X := \{x_1, x_2, x_3, x_4\}$, $\mathscr{E}_Y := \{y_1, y_2, y_3, y_4\}$ and

$$\mathscr{E}_{XY} = \{x_1, x_2\} \times \{y_1, y_2\} \cup \{x_3, x_4\} \times \{y_3, y_4\}. \tag{4.145}$$

The restriction maps of $\mathscr{E}$ are the canonical projections. Consider the probability functor such that

$$\begin{aligned} \mathscr{Q}_{XY} &= \mathscr{P}(\{x_1, x_2\} \times \{y_1, y_2\}) \cup \mathscr{P}(\{x_3, x_4\} \times \{y_3, y_4\}), \\ \mathscr{Q}_X &= \mathscr{P}(\{x_1, x_2\}) \cup \mathscr{P}(\{x_3, x_4\}), \\ \mathscr{Q}_Y &= \mathscr{P}(\{y_1, y_2\}) \cup \mathscr{P}(\{y_3, y_4\}). \end{aligned} \tag{4.146}$$

Note that $\mathscr{Q}$ is adapted and its marginalization morphisms are surjective. In this setting, the product XY does not satisfy the Definition 4.7 since neither $(2, 1)$ nor $(1, 2)$ can be a point of the north-east lattice path because of condition 3. This argument is relative to the enumerations displayed before, but a similar problem arises in any case. To specify a 1-cocycle f are sufficient the following functions of the variable $p \in [0, 1]$

$$f[X]\begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} \quad f[X]\begin{pmatrix} x_3 & x_4 \\ p & 1-p \end{pmatrix} \quad f[X]\begin{pmatrix} y_1 & y_2 \\ p & 1-p \end{pmatrix} \quad f[X]\begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix} \tag{4.147}$$

because $f[XY]$ is determined by (4.42). Moreover, applying again (4.42) to the probabilities

$$\begin{pmatrix} (x_1, y_1) & (x_2, y_1) & (x_1, y_2) \\ p_1 & p_2 & p_3 \end{pmatrix} \quad (p_1, p_2, p_3) \in \Delta(2) \tag{4.148}$$

we obtain, as in Example 4.2 that

$$f[X]\begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} = f[X]\begin{pmatrix} y_1 & y_2 \\ p & 1-p \end{pmatrix} = \lambda_1 S_\alpha(p, 1-p) \tag{4.149}$$

for some real number λ_1. Similarly exists λ_2 such that

$$f[X]\begin{pmatrix} x_3 & x_4 \\ p & 1-p \end{pmatrix} = f[X]\begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix} = \lambda_2 S_\alpha(p, 1-p) \tag{4.150}$$

Then any 1-cocycle is determined by the pair (λ_1, λ_2). Note that the structure of $\mathscr{Q}$ prevents both λ_1 and λ_2 from appearing in the equations derived from the cocycle condition, so they are independent. These considerations outline an isomorphism of real vector spaces $\mathbb{R}^2 \simeq Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$.

Example 4.4. This is a variation of the previous example. The information structure remains the same but we consider $\mathscr{P}$ as probability functor. Consider f as above and apply (4.42) to the probability

$$\begin{pmatrix} (x_1, y_1) & (x_2, y_1) & (x_3, y_3) & (x_4, y_3) \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix} \in \mathscr{P}_{XY}. \tag{4.151}$$

We have

$$\begin{gathered} f[X]\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix} = (p_1+p_2)^\alpha f[X]\begin{pmatrix} x_1 & x_2 \\ \frac{p_1}{p_1+p_2} & \frac{p_2}{1+p_2} \end{pmatrix} \\ (p_3+p_4)^\alpha f[X]\begin{pmatrix} x_3 & x_4 \\ \frac{p_3}{p_3+p_4} & \frac{p_4}{p_3+p_4} \end{pmatrix} + f[Y]\begin{pmatrix} y_1 & y_3 \\ p_1+p_2 & p_3+p_4 \end{pmatrix} = \\ \lambda_1 S_\alpha(\frac{p_1}{p_1+p_2}) + \lambda_2 S_\alpha(\frac{p_3}{p_3+p_4}) + f[Y]\begin{pmatrix} y_1 & y_3 \\ p_1+p_2 & p_3+p_4 \end{pmatrix} \end{gathered} \tag{4.152}$$

But the function of p

$$f[X]\begin{pmatrix} x_1 & x_3 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_1 & y_3 \\ p & 1-p \end{pmatrix} \tag{4.153}$$

can not be determined using the (FEITH) because $\mathscr{E}_{XY} \cap \{x_1, x_3\} \times \{y_1, y_3\}$ has only 2 elements. Thus, any choice of a measurable function $g : \Delta(1) \to \mathbb{R}$ such that $g(0,1) = g(1,0) = 0$, yields a 1-cocycle defining

$$f[X]\begin{pmatrix} x_1 & x_3 \\ p & 1-p \end{pmatrix} := g(p, 1-p). \tag{4.154}$$

Hence in this case the vector space of 1-cocycles, is infinite dimensional.

Example 4.5. We present another variation of Example 4.3, which shows that the conditions of Theorem 4.4 are only sufficient. Compared to the previous example, we only change the value of $\mathscr{E}$ on XY.

$$\mathscr{E}_{XY} := \{(x_1, y_1), (x_1, y_2), (x_1, y_3), (x_2, y_2), (x_3, y_3), (x_4, y_3), (x_4, y_4)\} \tag{4.155}$$

In this case, we have at our disposal all the probability laws with support in $\{(x_1, y_1), (x_1, y_3), (x_3, y_3)\}$, thus we can deduce

$$f[X]\begin{pmatrix} x_1 & x_3 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_1 & y_3 \\ p & 1-p \end{pmatrix} = \lambda_3 S_\alpha(p) \tag{4.156}$$

for some real number λ_3. Analogously, using respectively the 2-simplex of the probabilities on $\{(x_1, y_1), (x_1, y_2), (x_2, y_2)\}$ and $\{(x_3, y_3), (x_4, y_3), (x_4, y_4)\}$ we obtain

$$\begin{aligned} f[X]\begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_1 & y_2 \\ p & 1-p \end{pmatrix} = \lambda_1 S_\alpha(p) \\ f[X]\begin{pmatrix} x_3 & x_4 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix} = \lambda_2 S_\alpha(p) \end{aligned} \tag{4.157}$$

for some $\lambda_1, \lambda_2 \in \mathbb{R}$.
All the equalities presented below are achieved by applying the cocycle condition to the appropriate probability law (or are direct consequence of the previous ones).

$$f[X]\begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_2 & y_3 \\ p & 1-p \end{pmatrix} = f[X]\begin{pmatrix} x_1 & x_3 \\ p & 1-p \end{pmatrix} \tag{4.158}$$

This equation implies $\lambda_3 = \lambda_1$.

$$f[Y]\begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix} = f[X]\begin{pmatrix} x_1 & x_4 \\ p & 1-p \end{pmatrix} = f[Y]\begin{pmatrix} y_1 & y_3 \\ p & 1-p \end{pmatrix} \tag{4.159}$$

This implies $\lambda_3 = \lambda_2$. We are ready to determine $f[X]$ and $f[Y]$:

$$
\begin{aligned}
&f[X]\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix} = f[Y]\begin{pmatrix} y_1 & y_3 \\ p_2 & 1-p_2 \end{pmatrix} + (1-p_2)^\alpha f[X]\begin{pmatrix} x_1 & x_3 & x_4 \\ \frac{p_1}{1-p_2} & \frac{p_3}{1-p_2} & \frac{p_4}{1-p_2} \end{pmatrix} \\
&= (1-p_2)^\alpha \left[\left(\frac{(1-p_2-p_1)}{1-p_2} \right)^\alpha f[X]\begin{pmatrix} x_3 & x_4 \\ \frac{p_3}{1-p_1-p_2} & \frac{p_4}{1-p_1-p_2} \end{pmatrix} + f[Y]\begin{pmatrix} y_1 & y_3 \\ \frac{p_1}{1-p_2} & \frac{1-p_1-p_2}{1-p_2} \end{pmatrix} \right] \\
&+ f[Y]\begin{pmatrix} y_1 & y_3 \\ p_2 & 1-p_2 \end{pmatrix} = \lambda_3 \left(S_\alpha(p_2) + (1-p_2-p_1)^\alpha S_\alpha\left(\frac{p_3}{1-p_2-p_3} \right) + (1-p_2)^\alpha S_\alpha\left(\frac{p_1}{1-p_2} \right) \right) \\
&= \lambda_3 S_\alpha[X]\begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix}.
\end{aligned}
\tag{4.160}
$$

$$
\begin{aligned}
&f[Y]\begin{pmatrix} y_1 & y_2 & y_3 & y_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix} = f[X]\begin{pmatrix} x_1 & x_4 \\ 1-p_4 & p_4 \end{pmatrix} + (1-p_4)^\alpha f[Y]\begin{pmatrix} y_1 & y_2 & y_3 \\ \frac{p_1}{1-p_4} & \frac{p_2}{1-p_4} & \frac{p_3}{1-p_4} \end{pmatrix} \\
&= \lambda_3 S_\alpha(p_4) + (1-p_4-p_3)^\alpha f[Y]\begin{pmatrix} y_1 & y_2 \\ \frac{p_1}{1-p_4-p_3} & \frac{p_2}{1-p_4-p_3} \end{pmatrix} + (1-p_4)^\alpha f[X]\begin{pmatrix} x_1 & x_4 \\ \frac{1-p_4-p_3}{1-p_4} & \frac{p_3}{1-p_4} \end{pmatrix} \\
&= \lambda_3 \left(S_\alpha(p_4) + (1-p_4-p_3)^\alpha S_\alpha\left(\frac{p_1}{1-p_4-p_3} \right) + (1-p_4)^\alpha S_\alpha\left(\frac{p_3}{1-p_4} \right) \right) \\
&= \lambda_3 S_\alpha[Y]\begin{pmatrix} y_1 & y_2 & y_3 & y_4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix}.
\end{aligned}
\tag{4.161}
$$

Moreover, $f[X]$ and $f[Y]$ determine $f[XY]$ as we saw in (4.105). Therefore any 1-cocycle is a multiple of S_α: $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \simeq \mathbb{R}$. Nevertheless, the product XY is degenerate. In fact, if the probability functor involved is $\mathscr{P}$, one can verify that in order for a lattice path to exist that satisfies Definition 4.7, the set $\mathscr{E}_{XY}$ must contain at least 8 elements. Observe that the method used to determine $f[X]$ and $f[Y]$ is an instance of the Remark 4.2: for $f[X]$ we used the enumerations $\{x_2, x_1, x_3, x_4\}$, $\{y_2, y_3, y_1, y_4\}$; while for $f[Y]$ we used the enumerations $\{x_4, x_1, x_2, x_3\}$ and $\{y_4, y_3, y_1, y_2\}$.

A Presheaves of modules

The aim of this section is to prove that, given any presheaf of rings $\mathcal{O}$, the category $\mathbf{PMod}(\mathcal{O})$ of presheaves of $\mathcal{O}$-modules is abelian and has enough injective objects. First, we recall what a presheaf of modules is.

Definition A.1. Let $\mathcal{C}$ be a small category and $\mathcal{O} : \mathcal{C}^{\text{op}} \to \mathbf{Rings}$ be a presheaf of rings. A presheaf of $\mathcal{O}$-modules is an abelian presheaf $\mathcal{F}$, together with a map $\mu : \mathcal{O} \times \mathcal{F} \to \mathcal{F}$ of presheaves of sets such that for any object X in $\mathcal{C}$, the map μ_X defines on $\mathcal{F}(X)$ a structure of $\mathcal{O}(X)$-module.

A morphism $\alpha : (\mathcal{F}, \mu) \to (\mathcal{G}, \nu)$ of $\mathcal{O}$-modules, is a natural transformation such that the diagram

$$\begin{array}{ccc} \mathcal{O} \times \mathcal{F} & \xrightarrow{\mu} & \mathcal{F} \\ {\scriptstyle \mathrm{id}\times\alpha}\downarrow & & \downarrow{\scriptstyle \alpha} \\ \mathcal{O} \times \mathcal{G} & \xrightarrow{\nu} & \mathcal{G} \end{array} \tag{A.1}$$

is commutative.

Notation. $\mathrm{Hom}_{PMod(\mathcal{O})}(\mathcal{F}, \mathcal{G}) := \mathrm{Hom}_{\mathcal{O}}(\mathcal{F}, \mathcal{G})$

Proposition A.1. $\mathbf{PMod}(\mathcal{O})$ *is an abelian category.*

Proof. We demonstrate first that the category is pre-additive.
Given two morphisms $\mathcal{F} \overset{\alpha}{\underset{\beta}{\rightrightarrows}} \mathcal{G}$ of presheaves of modules, we consider the sum $\alpha + \beta$ as morphisms of abelian presheaves. This natural transformation, as well as the zero morphism $0 : \mathcal{F} \to \mathcal{G}$, makes the diagram A.1 commutative. Hence, $\mathrm{Hom}_{\mathcal{O}}(\mathcal{F}, \mathcal{G})$ inherits the structure of abelian group from $\mathrm{Hom}_{Psh(\mathcal{C},Ab)}(\mathcal{F}, \mathcal{G})$, and becomes a subgroup.
We now claim that $\mathbf{PMod}(\mathcal{O})$ is also additive. The constant presheaf associated to the abelian group (0), is the zero object in $\mathbf{Psh}(\mathcal{C}, \mathbf{Ab})$. But, if equipped with the trivial action of $\mathcal{O}$, it can be seen as a presheaf of $\mathcal{O}$-modules which is clearly both initial and terminal in $\mathbf{PMod}(\mathcal{O})$. Let $(\mathcal{F}, \mu), (\mathcal{G}, \nu)$ be presheaves of $\mathcal{O}$-modules. On the biproduct of $(\mathcal{F}, \mu)$ and $(\mathcal{G}, \nu)$ as abelian presheaves, is defined an action $\mu \oplus \nu$ described by

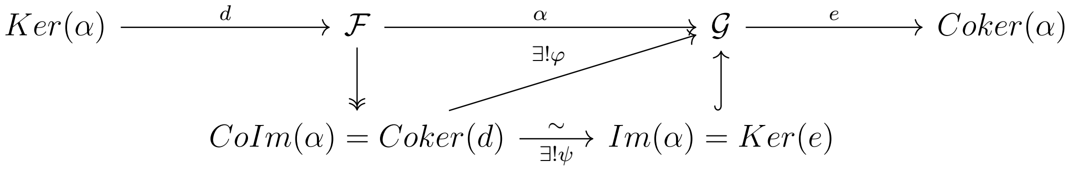
$$\begin{aligned} (\mu \oplus \nu)_X : \mathcal{O}(X) \times (\mathcal{F}(X) \oplus \mathcal{G}(X)) &\longrightarrow (\mathcal{F}(X) \oplus \mathcal{G}(X)) \\ (a, (s, t)) &\longmapsto (\mu_X(a, s), \nu_X(a, t)), \end{aligned} \tag{A.2}$$

for any $X \in Ob(\mathcal{C})$. The naturality in X of $\mu \oplus \nu$ follows by the naturality in X of μ and ν. Therefore, $(\mathcal{F} \oplus \mathcal{G}, \mu \oplus \nu)$ is a presheaf of -modules. It is possible to check that it is actually the biproduct in in $\mathbf{PMod}(\mathcal{O})$, but we omit the details.
Now we prove that any map in $\mathbf{PMod}(\mathcal{O})$ has a kernel end a cokernel. Consider a map $\alpha : (\mathcal{F}, \mu) \to (\mathcal{G}, \nu)$. In the category of abelian presheaves we have

$$Ker(\alpha) \xrightarrow{d} \mathcal{F} \xrightarrow{\alpha} \mathcal{G} \xrightarrow{e} Coker(\alpha)$$

Since α is $\mathcal{O}$-linear, i.e. makes the corresponding diagram (A.1) commutative, the action of $\mathcal{O}$ on $\mathcal{F}$ induces an action of $\mathcal{O}$ on $ker(\alpha)$ such that the map d turns out to be $\mathcal{O}$-linear, and then $ker(\alpha)$ with this action satisfies the universal property of kernel in $\mathbf{PMod}(\mathcal{O})$. Similarly, the action of $\mathcal{O}$ on $\mathcal{G}$ induces an action on $Coker(\alpha)$ which makes the map e $\mathcal{O}$-linear.
Finally, we remind that in $\mathbf{Psh}(\mathcal{C}, \mathbf{Ab})$ there is a canonical isomorphism $\psi : CoIm(\alpha) \xrightarrow{\sim} Im(\alpha)$. We claim that ψ is an isomorphism also in $\mathbf{PMod}(\mathcal{O})$. Notice that $Im(\alpha)$ and

$CoIm(\alpha)$ are defined by means of kernels and cokernels. Thus, by the above discussion, $Im(\alpha)$ and $CoIm(\alpha)$ are the image and coimage of α as abelian presheaves, and their $\mathcal{O}$-modules structures are induced by ν and μ respectively. Recall how ψ is obtained:



The map φ is induced by the universal property of the cokernel, thus is $\mathcal{O}$-linear, analogously ψ is induced by the universal property of the kernel, thus is $\mathcal{O}$-linear too. Hence the category is abelian. □

Let us recall some useful concepts concerning the functoriality between presheaves categories.

Definition A.2. Let $u : \mathcal{C} \to \mathcal{D}$ be a functor between small categories.

$$\begin{aligned} u^p : \mathbf{Psh}(\mathcal{D}, \mathbf{Sets}) &\longrightarrow \mathbf{Psh}(\mathcal{C}, \mathbf{Sets}) \\ \mathcal{F} &\longmapsto u^p(\mathcal{F}) := \mathcal{F} \circ u, \end{aligned} \tag{A.3}$$

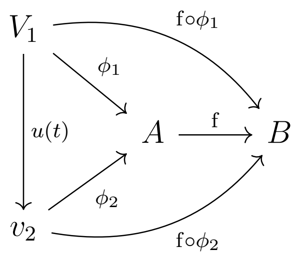
and for any arrow $\alpha : \mathcal{F} \to \mathcal{F}'$, $u^p(\alpha)$ is the natural transformation defined by $u^p(\alpha)_X := \alpha_{u(X)}$ for all $X \in Ob(\mathcal{C})$.

Proposition A.2. *The functor u^p has a left adjoint.*

Proof. The proof is by construction. Let $\mathcal{F}$ be a presheaf on $\mathcal{C}$. Let A be an object in $\mathcal{D}$. Define the category ${}_A\mathcal{I}$ whose objects are pairs (V, ϕ), where $V \in Ob(\mathcal{C})$ and $\phi : u(V) \to A$. A morphism $(V_1, \phi_1) \xrightarrow{t} (V_2, \phi_2)$ is given by a morphism $t : V_1 \to V_2$ in $\mathcal{C}$ such that $\phi_2 \circ u(t) = \phi_1$. On this category we consider the diagram ${}_A\mathcal{F} : {}_A\mathcal{I}^{\text{op}} \to \mathbf{Sets}$, defined on objects by ${}_A\mathcal{F}((V, \phi)) := \mathcal{F}(V)$, and if t is an arrow, ${}_A\mathcal{F}(t) := \mathcal{F}(t)$. Define

$${}_p u\mathcal{F}(A) := \lim_{{}_A\mathcal{I}^{\text{op}}} {}_A\mathcal{F}. \tag{A.4}$$

This construction in functorial in A, indeed, given $\mathrm{f} : A \to B$ in $\mathcal{D}$, we have that ${}_p u\mathcal{F}(B)$ is a cone over the diagram ${}_A\mathcal{F}$. We take as ${}_p u(\mathcal{F})(\mathrm{f})$ the unique map ${}_p u\mathcal{F}(B) \to {}_p u\mathcal{F}(A)$ induced by the universal property of the limit.



(A.5)

$$\begin{array}{ccc} \lim_{{}_A\mathcal{I}^{\mathrm{op}}} {}_A\mathcal{F} & \xrightarrow{p_{(V_1,\phi_1)}} & \mathcal{F}(V_1) \\ \Big\uparrow \exists!\, {}_pu(\mathcal{F})(\mathrm{f}) & \nearrow_{p_{(V_1,\mathrm{f}\circ\phi_1)}} & \\ \lim_{{}_B\mathcal{I}^{\mathrm{op}}} {}_B\mathcal{F} & & \end{array} \tag{A.6}$$

Furthermore, we claim that the construction is also functorial in $\mathcal{F}$. Given a natural transformation $\alpha : \mathcal{F} \to \mathcal{F}'$, we obtain, for any object A, a map between the diagrams ${}_A\mathcal{F}$ and ${}_A\mathcal{F}'$, which in turn induces a map ${}_pu(\mathcal{F})(A) \to {}_pu(\mathcal{F}')(A)$. These maps join to form a natural transformation ${}_pu(\alpha) : {}_pu(\mathcal{F}) \to {}_pu(\mathcal{F}')$. Once ${}_pu : \mathbf{Psh}(\mathcal{C}, \mathbf{Sets}) \to \mathbf{Psh}(\mathcal{D}, \mathbf{Sets})$ is defined, we proceed to prove the adjunction. We have to show that there exist isomorphisms

$$\bar{\ }: \mathrm{Hom}_{\mathbf{Psh}(\mathcal{C})}(u^p(\mathcal{G}), \mathcal{F}) \simeq \mathrm{Hom}_{\mathbf{Psh}(\mathcal{D})}(\mathcal{G}, {}_pu\mathcal{F}) \tag{A.7}$$

functorial both in $\mathcal{F}$ and $\mathcal{G}$.
For any $A \in Ob(\mathcal{D})$, $\mathcal{G}(A)$ is a cone over ${}_A(u^p\mathcal{G})$, hence there exists a unique map $\eta_{\mathcal{G}}(A) : \mathcal{G}(A) \to {}_puu^p\mathcal{G}(A)$ such that for every $(V, \phi) \in Ob({}_A\mathcal{I})$, it holds that $\mathcal{G}(\phi) = p_{(V,\phi)} \circ \eta_{\mathcal{G}}(A)$.
$\eta_{\mathcal{G}} : \mathcal{G} \to {}_uu^p\mathcal{G}$ is a natural transformation: pick $\mathrm{f} : A \to B$

$$\begin{array}{ccccc} \mathcal{G}(B) & \xrightarrow{\eta_{\mathcal{G}}(B)} & {}_puu^p\mathcal{G}(B) & & \\ \Big\downarrow \mathcal{G}(\mathrm{f}) & & \Big\downarrow {}_puu^p\mathcal{G}(\mathrm{f}) & \searrow^{p_{(V,\mathrm{f}\circ\phi)}} & \\ \mathcal{G}(A) & \xrightarrow[\eta_{\mathcal{G}}(A)]{} & {}_puu^p\mathcal{G}(A) & \xrightarrow[p_{(V,\phi)}]{} & \mathcal{G}(u(V)) \end{array} \tag{A.8}$$

The triangle on the right is commutative, then ${}_puu^p\mathcal{G}(\mathrm{f}) \circ \eta_{\mathcal{G}}(B)$ is the unique map $g : \mathcal{G}(B) \to {}_puu^p\mathcal{G}(A)$ such that for every (V, ϕ), it holds that $p_{(V,\phi)} \circ g = \mathcal{G}(\mathrm{f} \circ \phi)$. Since $\eta_{\mathcal{G}}(A) \circ \mathcal{G}(\mathrm{f})$ has this property too, they coincide.
$\eta : \mathbb{1}_{Psh(\mathcal{D})} \to {}_puu^p$ is natural: consider an arrow $\delta : \mathcal{G} \to \mathcal{G}'$, we have that

$$\begin{array}{ccc} \mathcal{G} & \xrightarrow{\eta_{\mathcal{G}}} & {}_puu^p\mathcal{G} \\ \Big\downarrow \delta & & \Big\downarrow {}_puu^p(\delta) \\ \mathcal{G}' & \xrightarrow{\eta_{\mathcal{G}'}} & {}_puu^p\mathcal{G}' \end{array} \tag{A.9}$$

is commutative iff, for any $A \in Ob(\mathcal{D})$, the left square of

$$\begin{array}{ccccc} \mathcal{G}(A) & \xrightarrow{\eta_{\mathcal{G}}(A)} & {}_puu^p\mathcal{G}(A) & \xrightarrow{p_{(V,\phi)}} & \mathcal{G}(u(V)) \\ \Big\downarrow \delta(A) & & \Big\downarrow {}_puu^p(\delta)(A) & & \Big\downarrow \delta(u(V)) \\ \mathcal{G}'(A) & \xrightarrow{\eta_{\mathcal{G}'(A)}} & {}_puu^p\mathcal{G}'(A) & \xrightarrow{p'_{(V,\phi)}} & \mathcal{G}'(u(V)) \end{array} \tag{A.10}$$

is commutative. The right hand square commutes by definition of ${}_pu$. Using the naturality of δ we get

$$p'_{(V,\phi)} \circ {}_p u u^p(\delta)(A) \circ \eta_{\mathcal{G}}(A) = \delta(u(V)) \circ \mathcal{G}(\phi) = \\ \mathcal{G}'(\phi) \circ \delta(A) = p'_{(V,\phi)} \circ \eta_{\mathcal{G}'}(A) \circ \delta(A) \qquad \forall\, (V,\phi). \tag{A.11}$$

As above we conclude that the square is commutative.
The arrow η just defined will serve as the unit of the adjunction. We now turn to find a counit. Let $X \in Ob(\mathcal{C})$ and $\mathcal{F}$ be a presheaf on $\mathcal{C}$, and consider the map

$$u^p{}_p u\mathcal{F}(X) \xrightarrow{p_{(u(X),id_{u(X)})}} \mathcal{F}(X). \tag{A.12}$$

It is maybe convenient introduce the notation $p^{\mathcal{F}}_{(V,\phi)}$ for the projections from the limit, to keep track of the diagram involved. We define $\varepsilon : u^p{}_p u \to \mathbb{1}_{Psh(\mathcal{C})}$ by setting $\varepsilon_{\mathcal{F}}(X) := p^{\mathcal{F}}_{(u(X),id_{u(X)})}$. The naturality of ε in both X and $\mathcal{F}$ follows directly from the definitions.
At this point we can prove the adjunction. Given $\alpha : u^p\mathcal{G} \to \mathcal{F}$, we define $\bar{\alpha} := {}_p u(\alpha) \circ \eta_{\mathcal{G}}$, as if η were the unit. Vice versa, given $\beta : \mathcal{G} \to \mathcal{F}$, we define $\bar{\beta} := \varepsilon_{\mathcal{F}} \circ u^p(\beta)$. These assignments are mutual inverse. In fact for any $X \in Ob(\mathcal{C})$

$$\bar{\bar{\alpha}}_X = p^{\mathcal{F}}_{(X,id_{u(X)})} \circ \bar{\alpha}_{u(X)} = \alpha_X \circ \mathcal{G}(id_{u(X)}) = \alpha_x. \tag{A.13}$$

$$\begin{array}{ccccc} & & \mathcal{G}(u(X)) & \xrightarrow{\alpha_X} & \mathcal{F}(X) \\ & \overset{\text{id}}{\nearrow} & \uparrow{\scriptstyle p^{u^p\mathcal{G}}} & & \uparrow{\scriptstyle p^{\mathcal{F}}(u(X))} \\ \mathcal{G}(u(X)) & \xrightarrow[\eta_{\mathcal{G}}(u(X))]{} & {}_p u u^p \mathcal{G}(u(X)) & \xrightarrow[{}_p u(\alpha)_{(u(X))}]{} & {}_p u\mathcal{F} \end{array} \tag{A.14}$$

And for any $A \in Ob(\mathcal{D})$ and any (V,ϕ) in ${}_A\mathcal{I}$, we have

$$p^{\mathcal{G}}_{(V,\phi)} \circ \bar{\bar{\beta}}_A = \bar{\beta}_V \circ \mathcal{G}(\phi) = p^{\mathcal{F}}_{(V,\text{id})} \circ \beta_{u(V)} \circ \mathcal{G}(\phi) = p^{\mathcal{F}}_{(V,\text{id})} \circ {}_p u\mathcal{F}(\phi) \circ \beta_A = p^{\mathcal{G}}_{(V,\phi)} \circ \beta_A \tag{A.15}$$

This concludes the proof, because the naturality in $\mathcal{F}$ and $\mathcal{G}$ of the bijections is a consequence of the naturality of ε and η, and of the functoriality of ${}_p u$ and u^p. $\square$

Remark A.1. Let $(\mathcal{A}, U)$ be a type of algebraic structure. [11] We can define, keeping the notations of proposition A.2, $u^p : \mathbf{Psh}(\mathcal{D},\mathcal{A}) \to \mathbf{Psh}(\mathcal{C},\mathcal{A})$ by $u^p(\mathcal{G})(X) := \mathcal{G}(u(X))$. We can also define ${}_p u$ between categories of presheaves on $\mathcal{A}$, using the same limit used for **Sets**. This is possible since $\mathcal{A}$ admits all limits. Given a presheaf of algebraic structures, either defined on $\mathcal{C}$ or $\mathcal{D}$, we can compose it with U and obtain its underlying presheaf of sets. Moreover, this operation, commutes with both u^p and ${}_p u$: it is clear for u^p; while for ${}_p u$ it follows from the fact that U commutes with limits in $\mathcal{A}$. Observe that the previous proof would still hold even if, in place of the category of sets, we had any algebraic structure. Indeed, the only property of **Sets** used in the proof, is completeness. Hence there is an adjunction $u^p \dashv {}_p u$ even between presheaves of algebraic structures.

Examples of types of algebraic structures are $(\mathbf{Ab}, U)$ and $(\mathbf{Ring}, U)$, where U denotes the forgetful functor in both situations.

This last remark enables us to define the functors u^p and ${}_pu$ even between categories of presheaves of modules.
Let $u : \mathcal{C} \to \mathcal{D}$ a covariant functor between small categories. Let $\mathcal{O} \in \mathbf{Psh}(\mathcal{D}, \mathbf{Ring})$. Suppose that $\mathcal{G}$ is a presheaf of $\mathcal{O}$-modules, this amounts to say that it is an abelian presheaf with a natural transformation $\nu : U \circ \mathcal{O} \times U \circ \mathcal{G} \longrightarrow U \circ \mathcal{G}$, satisfying properties in Definition A.1. We can compute $u^p\mathcal{G}$ as an abelian presheaf, and endow it with the action

$$\begin{array}{ccc} u^p(U \circ \mathcal{O} \times U \circ \mathcal{G}) & \xrightarrow{u^p(\nu)} & u^p(U \circ \mathcal{G}) \\ \Vert & & \Vert \\ U \circ (u^p\mathcal{O}) \times U \circ (u^p\mathcal{G}) & & U \circ (u^p\mathcal{G}) \end{array} \tag{A.16}$$

Here we used that U commutes with u^p, and u^p commutes with limits. This last property can be checked directly, but it is also a consequence of the fact that u^p has a left adjoint u_p [11]. Observe that, for any $X \in Ob(\mathcal{C})$, this new action actually turns $u^p\mathcal{G}(X) = \mathcal{G}(u(X))$ into an $\mathcal{O}(X)$ module, simply because $u^p(\mu)_X = \mu_{u(X)}$. Furthermore, for any arrow $\delta \in \mathrm{Hom}_{\mathcal{O}}(\mathcal{G}, \mathcal{G}')$, the diagram

$$\begin{array}{ccc} u^p\mathcal{O} \times u^p\mathcal{G} & \xrightarrow{u^p(\nu)} & \mathcal{G} \\ \downarrow{\scriptstyle id \times u^p(\delta)} & & \downarrow{\scriptstyle u^p(\delta)} \\ u^p\mathcal{O} \times u^p\mathcal{G}' & \xrightarrow{u^p(\nu')} & \mathcal{G}' \end{array} \tag{A.17}$$

is commutative, because it is obtained applying u^p to the diagram encoding the $\mathcal{O}$-linarity of δ. Hence $u^p(\nu)$ is $\mathcal{O}$-linear.
This discussion yields a functor $\mathbf{PMod}(\mathcal{O}) \to \mathbf{PMod}(u^p\mathcal{O})$, which we call again u^p, since it is compatible with the definition of u^p on abelian presheaves.
Now, we move to definition of ${}_pu : \mathbf{PMod}(u^p\mathcal{O}) \to \mathbf{PMod}(\mathcal{O})$.
Pick an $u^p\mathcal{O}$-module $(\mathcal{F}, \mu)$, the abelian presheaf ${}_pu\mathcal{F}$ can be endowed with a structure of $\mathcal{O}$-module in this way :

$$\begin{array}{ccccc} U \circ ({}_puu^p\mathcal{O}) \times U \circ ({}_pu\mathcal{F}) & \xrightarrow{\sim} & {}_pu(U \circ (u^p\mathcal{O}) \times U \circ \mathcal{F}) & \xrightarrow{{}_pu(\mu)} & {}_pu(U \circ \mathcal{F}) \\ {\scriptstyle U(\eta_{\mathcal{O}}) \times \mathrm{id}}\uparrow & & & & \Vert \\ U \circ \mathcal{O} \times U \circ ({}_pu\mathcal{F}) & & & & U \circ ({}_pu\mathcal{F}) \end{array} \tag{A.18}$$

This was possible since also ${}_pu$ preserves limits, being a right adjoint. For any $A \in Ob(\mathcal{G})$, ${}_pu(\mu)_A$ defines a structure of ${}_puu^p\mathcal{O}(A)$-module on ${}_pu\mathcal{F}(A)$, which in turn becomes a $\mathcal{O}$-module by restriction of scalars, trough $\eta_{\mathcal{O}}(A)$. As before, we consider an arrow $\gamma : \mathcal{F} \to$

$\mathcal{F}'$ in $\mathbf{PMod}(u^p\mathcal{O})$, the commutative diagram

$$\begin{array}{ccccc} \mathcal{O} \times {}_pu\mathcal{F} & \xrightarrow{\eta_{\mathcal{O}}\times \mathrm{id}} & {}_puu^p\mathcal{O} \times {}_pu\mathcal{F} & \xrightarrow{{}_pu(\mu)} & \mathcal{F} \\ \big\downarrow{\scriptstyle \mathrm{id}\times {}_pu(\gamma)} & & \big\downarrow{\scriptstyle id\times {}_pu(\gamma)} & & \big\downarrow{\scriptstyle {}_pu(\gamma)} \\ \mathcal{O} \times {}_pu\mathcal{F}' & \xrightarrow{\eta_{\mathcal{O}}\times \mathrm{id}} & {}_puu^p\mathcal{O} \times {}_pu\mathcal{F}' & \xrightarrow{{}_pu(\mu')} & \mathcal{F}' \end{array} \tag{A.19}$$

reveals that ${}_pu(\gamma)$ is $\mathcal{O}$-linear.

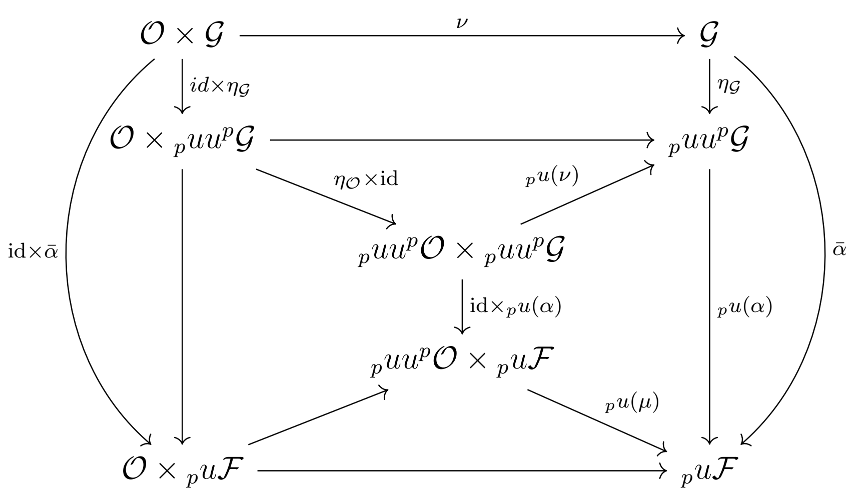
Proposition A.3.

$$\mathbf{PMod}(\mathcal{O}) \quad \underset{{}_pu}{\overset{u^p}{\perp}} \quad \mathbf{PMod}(u^p\mathcal{O}) \tag{A.20}$$

Proof. Let $(\mathcal{F}, \mu)$ be an $u^p\mathcal{O}$-module and $(\mathcal{G}, \nu)$ be an $\mathcal{O}$-module. Since $\mathrm{Hom}_{\mathcal{O}}(\mathcal{G}, {}_pu\mathcal{F}) \subset \mathrm{Hom}_{Psh(\mathcal{D},\mathbf{Ab})}(\mathcal{G}, {}_pu\mathcal{F})$ and $\mathrm{Hom}_{u^p\mathcal{O}}(u^p\mathcal{G}, \mathcal{F}) \subset \mathrm{Hom}_{Psh(\mathcal{C},\mathbf{Ab})}(u^p\mathcal{G}, \mathcal{F})$, it is enough to show

$$\alpha : u^p\mathcal{G} \to \mathcal{F} \,\text{is}\, u^p\mathcal{O} - \text{linear} \iff \bar{\alpha} : \mathcal{G} \to {}_pu\mathcal{F} \,\text{is}\, \mathcal{O} - \text{linear}.$$

It is all about writing down the suitable diagrams.
$\Rightarrow$) We have to prove the commutativity of the external rectangle of



The upper rectangle commutes by naturality of η as the unit of the adjunction $u^p \dashv {}_pu$ within presheaves of sets, while the lower commutes because of the hypothesis on α.
The proof of $\Leftarrow$) is analogous. □

Proposition A.4. *Let $\mathcal{C}$ be a small category and $\mathcal{O} : \mathcal{C} \to$ **Rings** a presheaf of rings. Then $\mathbf{PMod}(\mathcal{O})$ has enough injective objects.*

Proof. Consider the discrete category $Ob(\mathcal{C})$. There is a canonical functor $i : Ob(\mathcal{C}) \to \mathcal{C}\,, X \mapsto i(X) = X$. Because of the discussion above, we have $i^p \dashv {}_pi : \mathbf{PMod}(\mathcal{O}) \to$

PMod$(i^p\mathcal{O})$. Let us provide an explicit description of i^p, η, ε in relation to this specific situation. Let $\mathcal{F}$ an $i^p\mathcal{O}$-module,

$$
{}_pi\mathcal{F}(X) = \lim_{X\mathcal{I}} {}_X\mathcal{F} = \prod_{\substack{\mathrm{dom}(\varphi)=Y \\ \varphi\in h_X}} \mathcal{F}(Y). \tag{A.21}
$$

Where h_X denotes the maximal sieve[22] on X. The limit coincides with the product because ${}_X\mathcal{I}$ is discrete, since in $Ob(\mathcal{C})$ there are no maps except for the identities. The counit $\varepsilon : i^p{}_pi \Rightarrow \mathbb{1}_{PMod(i^p\mathcal{O})}$ is specified by:

$$
\forall\ \mathcal{F}, \forall\ X \in Ob(\mathcal{C}), \prod_{\substack{\mathrm{dom}(\varphi)=Y \\ \varphi\in h_X}} \mathcal{F}(Y) \overset{\varepsilon_{\mathcal{F}}(X)=\pi_{\mathrm{id}_X}}{\longrightarrow} \mathcal{F}(X). \tag{A.22}
$$

For $\mathcal{G}$ in **PMod**$(\mathcal{O})$, the component X of $\eta_{\mathcal{G}}$ is induced by universal property of the product.

$$
\begin{array}{ccc}
\mathcal{G}(X) & \xrightarrow{\exists!\,\eta_{\mathcal{G}}(X)} & \prod_{\substack{\mathrm{dom}(\varphi)=Y \\ \varphi\in h_X}} \mathcal{G}(Y) \\
& \searrow^{\mathcal{G}(\varphi)} & \big\downarrow \pi_{\varphi} \\
& & \mathcal{G}(Y)
\end{array} \tag{A.23}
$$

Observe for later use that $\eta_{\mathcal{G}}(X)$ is a monomorphism (in the category of $\mathcal{O}(X)$-modules). To see this, suppose the existence of $\chi : A \to \mathcal{G}(X)$ such that $\eta_{\mathcal{G}}(X) \circ \chi = 0$, then $\mathcal{G}(\varphi) \circ \chi = 0\ \forall\, \varphi \in h_X$. In particular if $\varphi = \mathrm{id}_X$, we have $\chi = 0$.
An object $\mathcal{F}$ in **PMod**$(i^p\mathcal{O})$ is just a collection $\big(\mathcal{F}(X)\big)_{X\in Ob(\mathcal{C})}$, where $\mathcal{F}(X)$ is an $\mathcal{O}(X)$-module. Similarly, a morphism $\mathcal{F} \to \mathcal{F}'$ is a collection of morphisms of $\mathcal{O}(X)$-modules, indexed by $X \in Ob(\mathcal{C})$. Hence, a presheaf of $i^p\mathcal{O}$-modules $\mathcal{H}$ is injective iff $\mathcal{H}(X)$ is injective as $\mathcal{O}(X)$-module, for all $X \in Ob(\mathcal{C})$. Suppose for a moment that for any $\mathcal{F}$ as above we have a monomorphism $\gamma_{\mathcal{F}} : \mathcal{F} \to \mathrm{I}(\mathcal{F})$ into an injective object. Since ${}_pi$ has a left adjoint which is exact, ${}_pi\big(\mathrm{I}(\mathcal{F})\big)$ is still an injective object [14], and of course ${}_pi(\gamma)$ remains a monomorphism. Now, for any $\mathcal{G}$ in **PMod**$(\mathcal{O})$, we follow the previous argument with $i^p\mathcal{G}$, so that we obtain the monomorphism ${}_pi(\gamma_{i^p\mathcal{G}})$. Then

$$
\mathcal{G} \xrightarrow{\eta_{\mathcal{G}}} {}_pii^p\mathcal{G} \xrightarrow{{}_pi(\gamma_{i^p\mathcal{G}})} {}_pi\big(\mathrm{I}(i^p\mathcal{G})\big) \tag{A.24}
$$

is a monomorphism from $\mathcal{G}$ into an injective object.
So we are left to find $\gamma_{\mathcal{F}}$. Recall that if R is any ring, and M is an R-module, we have an injection

$$
e_M : M \to \prod_{\mathrm{Hom}_{\mathbf{Ab}}(M,\mathbb{Q}/\mathbb{Z})} \mathrm{Hom}_{\mathbf{Ab}}(R, \mathbb{Q}/\mathbb{Z}), \tag{A.25}
$$

[22]It means that h_X is the set of all the functions with codomain X

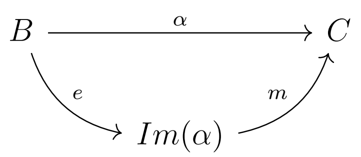
and the target is an injective R-module. Thus, we can define $\gamma_{\mathcal{F}}(X) := e_{\mathcal{F}(X)}$ and $\mathrm{I}(\mathcal{F})(X) := \prod_{\mathrm{Hom}_{\mathbf{Ab}}(\mathcal{F}(X),\mathbb{Q}/\mathbb{Z})} \mathrm{Hom}_{\mathbf{Ab}}(\mathcal{O}(X), \mathbb{Q}/\mathbb{Z})$ $\square$

B Relative bar resolution

Definition B.1. A relative abelian category consists in a pair of abelian categories $(\mathcal{A}, \mathcal{M})$ together with a functor $\square : \mathcal{A} \to \mathcal{M}$ which is additive, exact and faithful.

Definition B.2. A short exact sequence $0 \to K \xrightarrow{\chi} B \xrightarrow{\sigma} C \to 0$ in $\mathbf{A}$, also denoted $\chi||\sigma$, is said relatively split, or $\square$-split if the sequence $\square\chi||\square\sigma$ splits in $\mathbf{M}$.
A monomorphism χ is said allowable, or $\square$-allowable, if $\chi||coker(\chi)$ is relatively split.
Dually, an epimorphism σ is said allowable if $ker(\sigma)||\sigma$ is relatively split.
In general, a morphism α is said allowable if: $Im(\alpha)$ is an allowable monomorphism and $Coim(\alpha)$ is an allowable epimorphism.

Recall that an exact sequence $0 \to M \xrightarrow{\mathrm{f}} N \xrightarrow{\mathrm{g}} L \to 0$ in an abelian category, is split if and only if g has a right inverse (section) or f has a left inverse (retraction). In light of this, we show that a morphism $\alpha : B \to C$ in $\mathbf{A}$ is $\square$-allowable if exists $s : \square C \to \square B$ such that $\square\alpha = \square\alpha \circ \square s \circ \square\alpha$. Consider an epi-mono factorization of α:



(B.1)

We have $\square(m \circ e) = \square(m \circ e) \circ s \circ \square(m \circ e)$. Since $\square$ is exact, $\square e$ is an epimorphism and $\square m$ is a monomorphism. Then we obtain $\mathrm{id} = \square e \circ s \circ \square m$. Therefore, e has a right inverse and m has a left inverse, which implies that $e = Coim(\alpha)$ and $m = Im(\alpha)$ are allowable.

Definition B.3. A resolvent pair is a relative abelian category such that $\square$ has a left adjoint.

We fix some notations for later use: Denote F the left adjoint of $\square$; $\eta : \mathbb{1}_{\mathbf{M}} \to \square F$ the unit of the adjunction; $\epsilon : F\square \to \mathbb{1}_{\mathbf{A}}$ is the counit.

Example B.1. Let $\iota : S \to R$ an homomorphism of rings. An R-module M, can be seen as an S-module by restriction of scalars, i.e. defining $s \cdot m := \iota(s) \cdot m$ for $s \in S$ and $m \in M$. Moreover any morphism $\mathrm{f} : M \to N$ between R-modules, is also S-linear since $\mathrm{f}(s \cdot m) = \mathrm{f}(\iota(s) \cdot m) = \iota(s) \cdot \mathrm{f}(m) = s \cdot \mathrm{f}(m)$. Is then defined a functor $\iota^* : \mathbf{Mod}(R) \to \mathbf{Mod}(S)$ which is clearly additive and faithful. The functor ι^* is also exact because the kernel and the cokernel of a morphism between R-modules are unchanged if it is viewed as a morphism between S-modules. Therefore, ι^* defines a relative abelian

category.
Now consider the functor

$$\begin{aligned} \iota_! : \mathbf{Mod}(S) &\to \mathbf{Mod}(R) \\ B &\mapsto R \otimes_S B, \end{aligned} \tag{B.2}$$

where the R-module structure on $\iota_!(B)$ is defined by $r_1 \cdot (r \otimes b) = (rr_1) \otimes b$, for all $r, r_1 \in R$ and $b \in B$. For any arrow $\mathrm{g} : B \to A$, $\iota_!(\mathrm{g}) := id_R \otimes \mathrm{g}$. We have an adjunction $\iota_! \dashv \iota^*$, witnessed by the functorial isomorphism

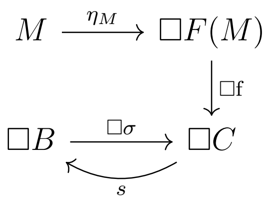
$$\begin{aligned} \mathrm{Hom}_S(B, \iota^*(M)) &\simeq \mathrm{Hom}_R(R \otimes_S B, M) \\ \mathrm{f} &\leftrightarrow \bar{\mathrm{f}} : r \otimes b \;\mapsto\; r \cdot \mathrm{f}(b) \\ \big(\bar{\alpha} : b \mapsto \alpha(b \otimes 1_R)\big) &\leftrightarrow\; \alpha \end{aligned} \tag{B.3}$$

Therefore $(\mathbf{Mod}(R), \mathbf{Mod}(S), \iota^*, \iota_!)$ is a resolvent pair.

Definition B.4. A relative projective object P, is an object in $\mathbf{A}$ such that for any allowable epimorphism $\sigma : B \to C$ and any morphism $\mathrm{f} : P \to C$, there exists a morphism $\mathrm{f}' : P \to B$ such that $\sigma \circ \mathrm{f}' = \mathrm{f}$.

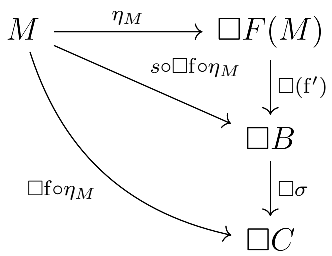
Proposition B.1. *All object of the form $F(M)$, for some M in $\boldsymbol{M}$, are relative projective.*

Proof. For any $\mathrm{f} : F(M) \to C$ and any allowable epimorphism $\sigma : B \to C$, we have to find an arrow $\mathrm{f}' : F(M) \to B$ such that $\mathrm{f} = \sigma \circ \mathrm{f}'$.



(B.4)

Here, s is a right inverse of $\square\sigma$, since σ is an allowable epimorphism. Moreover, by the universal property of unit, there exist a unique map $\mathrm{f}' : F(M) \to B$ such that $s \circ \square\mathrm{f} \circ \eta_M = \square\mathrm{f}' \circ \eta_M$.



(B.5)

The external diagram commutes, then both f and $\sigma \circ \mathrm{f}'$ fit in $\square(\_) \circ \eta_M = \square\mathrm{f} \circ \eta_M$, hence they must coincide, again by the universal property of η.

□

Definition B.5. Let A be an object of $\mathbf{A}$. A relatively projective resolution of A is a complex $\cdots \to P_2 \to P_1 \to P_0$ in $\mathbf{A}$ together with a map $\varepsilon : P_0 \to A$ such that all the P_n's are relative projective and $\cdots \to P_2 \to P_1 \to P_0 \xrightarrow{\varepsilon} A$ is a long exact sequence. Besides, a resolution is said allowable if both ε and all differentials of the complex $P_\bullet$ are allowable morphisms. Finally, a resolution is said relatively free if $P_n = F(M_n)$ $\forall n \in \mathbb{N}$ for some M_n in $\mathbf{M}$.

We claim that any object in $\mathbf{A}$ has a relatively free (and hence relatively projective) resolution.
Let C in $\mathrm{Ob}(\mathbf{A})$. Define

$$\beta_n(C) := \begin{cases} C & n = -1 \\ (F\Box)^{n+1}(C) & n \geq 0, \end{cases} \tag{B.6}$$

where $(F\Box)^n$ indicates the (n+1)-fold iteration of the functor $F\Box : \mathbf{A} \to \mathbf{A}$. The image of these objects trough $\Box$ can be arranged in a sequence, setting $s_{-1} := \eta_{\Box C}$ and $s_n := \eta_{\Box\beta_n(C)}$ for $n \geq 0$.

$$\Box C \xrightarrow{s_{-1}} \Box\beta_0(C) \xrightarrow{s_0} \Box\beta_1(C) \xrightarrow{s_1} \dots \tag{B.7}$$

Proposition B.2. *There are unique* $\boldsymbol{A}$*-morphisms*

$$\varepsilon : \beta_0(C) \to C \qquad \partial_n : \beta_n(C) \to \beta_{n-1}(C) \quad for\ n \in \mathbb{N}, \tag{B.8}$$

which make $\beta_\bullet(C)$ *a relatively free allowable resolution of* C *with* $s_\bullet$ *as a contracting homotopy in* $\boldsymbol{M}$*. This resolution, with its contracting homotopy, is a covariant functor of* C*.*

Proof. $s_\bullet$ is a contracting homotopy if $id_{\Box\beta_n(C)} = \Box\partial_{n+1} \circ s_n + s_{n-1} \circ \Box\partial_n$ holds for all $n \geq -1$, with $\partial_{-2} = 0$. For $n = -1$, it has to be $\Box\varepsilon \circ \eta_{\Box C} = \mathrm{id}_{\Box C}$, thus, by universal property of the unit, ε must be equal to ϵ_C. From this we deduce also that ϵ_C is allowable, since it has a right inverse. The successive differentials are now defined by recursion, so that s will be a contracting homotopy: assume that $\partial_0, \partial_1, \dots, \partial_n$ have already been defined, ∂_{n+1} must be the unique arrow such that $\Box\partial_{n+1} \circ \eta_{\Box\beta_n(C)} = \mathrm{id}_{\Box\beta_n(C)} - s_{n-1} \circ \Box\partial_n$ holds.

$$\begin{array}{ccccccccc} \longrightarrow & \Box\beta_{n+1}(C) & \xrightarrow{\Box\partial_{n+1}} & \Box\beta_n(C) & \xrightarrow{\Box\partial_n} & \Box\beta_{n-1} & \xrightarrow{\Box\partial_{n-1}} \\ & \Big\| \mathrm{id} & \swarrow s_n & \Big\| \mathrm{id} & \swarrow s_{n-1} & \Big\| \mathrm{id} & \\ \longrightarrow & \Box F\Box\beta_n(C) & \xrightarrow{\Box\partial_{n+1}} & \Box\beta_n(C) & \xrightarrow{\Box\partial_n} & \Box\beta_{n-1} & \xrightarrow{\Box\partial_{n-1}} \end{array} \tag{B.9}$$

We prove by induction on n that $\partial_{n-1} \circ \partial_n = 0,\ \forall\, n \geq 0$.
$n = -1$). We have $\Box\epsilon_C \circ \Box\partial_1 \circ s_0 = \Box\epsilon_C - \Box\epsilon_C \circ s_{-1} \circ \Box\epsilon_C = \Box\epsilon_C - \Box\epsilon_C = 0$, thus

$\epsilon_C \circ \partial_1 = 0$ again by the universal property of η.

$$\begin{array}{ccccccc} \longrightarrow & \Box\beta_1(C) & \xrightarrow{\Box\partial_1} & \Box\beta_0(C) & \xrightarrow{\Box\epsilon_C} & \Box C \\ & \Big\| \text{id} & \swarrow s_0 & \Big\| \text{id} & \swarrow s_{-1} & \Big\| \text{id} \\ \longrightarrow & \Box\beta_1(C) & \xrightarrow{\Box\partial_1} & \Box\beta_0(C) & \xrightarrow{\Box\epsilon_C} & \Box C \end{array} \tag{B.10}$$

$n \Rightarrow n+1$). $\Box\partial_n \circ \Box\partial_{n+1} \circ s_n = \Box\partial_n - \Box\partial_n \circ s_{n-1} \circ \Box\partial_n = s_{n-2} \circ \Box\partial_{n-1} \circ \Box\partial_n = 0$. We deduce that $\Box\partial_n \circ \Box\partial_{n+1} = 0$ since s_n is a component of the unit of the adjunction. Furthermore, all differentials are allowable, in fact, $\forall\, n \in \mathbb{N}_0$, it holds that

$$\Box\partial_{n+1} \circ s_n \circ \Box\partial_{n+1} = \Box\partial_{n+1} - s_n \circ \Box\partial_n \circ \Box\partial_{n+1} = \Box\partial_{n+1}. \tag{B.11}$$

Finally we prove functoriality. Given $\mathrm{f} : C \to D$ in $\mathbf{A}$, we have to define $\beta_\bullet(\mathrm{f})$. A natural choice is $\beta_n(\mathrm{f}) := (F\Box)^n(\mathrm{f})$ for $n \geq 0$. We now check that these maps yield a morphism of resolutions, which means that

$$(F\Box)^n(\mathrm{f}) \circ \partial_n^C = \partial_n^D \circ (F\Box)^n(\mathrm{f}) \quad \forall\, n \in \mathbb{N}, \text{ and } \quad \Box\mathrm{f} \circ \epsilon_C = \epsilon_D \circ F\Box(\mathrm{f}). \tag{B.12}$$

The last equality in (B.12) is just a consequence of naturality of ϵ.
Let $n = 1$ for simplicity, all the equalities with $n \neq 1$ can be proved in an analogous way. The adjunction $F \dashv \Box$ provides the following commutative diagram:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathbf{M}}(\Box\beta_0(C), \Box\beta_0(C)) & \xrightarrow{\overline{\ \ }} & \mathrm{Hom}_{\mathbf{A}}(\beta_1(C), \beta_0(C)) \\ \Big\downarrow {\scriptstyle \Box\beta_0(\mathrm{f})\circ\_} & & \Big\downarrow {\scriptstyle \beta_1(\mathrm{f})\circ\_} \\ \mathrm{Hom}_{\mathbf{M}}(\Box\beta_0(C), \Box\beta_0(D)) & \longrightarrow & \mathrm{Hom}_{\mathbf{M}}(\beta_1(C), \beta_0(D)) \\ {\scriptstyle \_\circ\Box\beta_0(\mathrm{f})} \Big\uparrow & & {\scriptstyle \_\circ\beta_1(\mathrm{f})} \Big\uparrow \\ \mathrm{Hom}_{\mathbf{M}}(\Box\beta_0(D), \Box\beta_0(D)) & \longrightarrow & \mathrm{Hom}_{\mathbf{M}}(\beta_1(D), \beta_0(D)) \end{array} \tag{B.13}$$

Using the naturality of both η and ϵ we get $\Box\beta_0(\mathrm{f}) \circ (\mathrm{id} - s_{-1}^C \circ \Box\epsilon_C) = (\mathrm{id} - s_{-1}^D \circ \Box\epsilon_D) \circ \Box\beta_0(\mathrm{f})$. This implies that

$$\begin{aligned} \beta_1(\mathrm{f}) \circ \partial_1^C =& \beta_1(\mathrm{f}) \circ \overline{(\mathrm{id} - s_{-1}^C \circ \Box\epsilon_C)} = \\ \overline{\Box\beta_0(\mathrm{f}) \circ (\mathrm{id} - s_{-1}^C \circ \Box\epsilon_C)} =& \overline{(\mathrm{id} - s_{-1}^D \circ \Box\epsilon_D) \circ \Box\beta_0(\mathrm{f})} = \\ \overline{(\mathrm{id} - s_{-1}^D \circ \Box\epsilon_D)} \circ \beta_1(\mathrm{f}) = \partial_1^D \circ \beta_1(\mathrm{f}) \end{aligned} \tag{B.14}$$

□

The relatively free resolution just presented, is named *unnormalized bar resolution.*

Example B.2. We present a particular resolvent pair that is relevant to our initial pur-

pose. Let $\mathcal{C}$ be a small category, and $\iota : \mathcal{O} \to \mathcal{R}$ a morphism of presheaves of rings. Take $\mathbf{A} := \mathbf{PMod}(\mathcal{R})$ and $\mathbf{M} := \mathbf{PMod}(\mathcal{O})$. A relative abelian category arises considering the functor $\Box : \mathbf{PMod}(\mathcal{R}) \to \mathbf{PMod}(\mathcal{O})$ described by $\Box\mathcal{F}(X) = \iota_X^*(\mathcal{F}(X))$, whenever $\mathcal{F}$ is an $\mathcal{R}$-module and X is an object of $\mathcal{C}$. Similarly to ι^* of the Example B.1, the functor $\Box$ has a left adjoint F, which associates to $\mathcal{N}$ in $\mathbf{PMod}(\mathcal{O})$ the presheaf $\mathcal{R} \otimes_{\mathcal{O}} \mathcal{N} : X \mapsto \mathcal{R}(X) \otimes_{\mathcal{O}(X)} \mathcal{N}(X)$.[11]
In order to construct relatively free resolutions, it will be necessary to have a description of the unit and counit of the adjunction. The $(\mathcal{N}, X)$-component of the unit is

$$\begin{aligned} \eta_{\mathcal{N}}(X) : \mathcal{N}(X) &\to \mathcal{R}(X) \otimes_{\mathcal{O}(X)} \mathcal{N}(X) \\ b &\to 1 \otimes b \end{aligned} \tag{B.15}$$

while the $(\mathcal{F}, X)$-component of the counit is

$$\begin{aligned} \epsilon_{\mathcal{F}}(X) : \mathcal{R}(X) \otimes_{\mathcal{O}(X)} \mathcal{F}(X) &\to \mathcal{F}(X) \\ r \otimes c &\to r \cdot c \end{aligned} \tag{B.16}$$

Once $\mathcal{N}, \mathcal{F}$ are fixed, also the isomorphism $\bar{\ }: \mathrm{Hom}_{\mathcal{O}}(\mathcal{N}, \Box\mathcal{F}) \simeq \mathrm{Hom}_{\mathcal{R}}(\mathcal{R} \otimes_{\mathcal{O}} \mathcal{N}, \mathcal{F})$ can be described using Example B.1. Indeed, given $\alpha : \mathcal{N} \to \Box\mathcal{F}$, we have $\bar{\alpha}_X = \overline{(\alpha_X)}$, where the second bar is referred to the adjunction $(\iota_X)_! \dashv (\iota_X)^*$. Fix an $\mathcal{R}$-module $\mathcal{F}$. We compute the differentials of $\beta_\bullet(\mathcal{F})$ in this specific resolvent pair.

$$\dots \xrightarrow{\partial_3} \beta_3(\mathcal{F}) \xrightarrow{\partial_2} \mathcal{R} \otimes_{\mathcal{O}} (\Box(\mathcal{R} \otimes_{\mathcal{O}} \Box\mathcal{F})) \xrightarrow{\partial_1} \mathcal{R} \otimes_{\mathcal{O}} \Box\mathcal{F} \xrightarrow{\epsilon_{\mathcal{F}}} \mathcal{F} \longrightarrow 0$$

We have shown that $\partial_1 = \overline{\mathrm{id}_{\beta_0(\mathcal{F})} - \eta_{\Box\mathcal{F}} \circ \Box\epsilon_{\mathcal{F}}}$. Therefore, for X a generic object in $\mathcal{C}$

$$\begin{aligned} \partial_1(X) : \mathcal{R}(X) \otimes_{\mathcal{O}(X)} (\Box(\mathcal{R}(X) \otimes_{\mathcal{O}(X)} \Box\mathcal{F}(X))) &\to \Box(\mathcal{R}(X) \otimes_{\mathcal{O}(X)} \Box\mathcal{F}(X)) \\ r \otimes (r_1 \otimes c) &\mapsto rr_1 \otimes c - r \otimes r_1 c \end{aligned} \tag{B.17}$$

The higher differentials are determined by induction; assume that the form of ∂_{n-1} is

$$\partial_{n-1}(X) : r \otimes r_1 \otimes \dots \otimes r_{n-1} \otimes c \longmapsto \sum_{i=0}^{n-1} (-1)^i r \otimes \dots \otimes r_i r_{i+1} \otimes \dots \otimes r_{n-1} \otimes c$$

Observe that $\partial_1(X)$ actually has this form, whence the basic step is true.

Using $\partial_n = \overline{\mathrm{id}_{\beta_{n-1}(\mathcal{F})} - s_{n-2} \circ \Box\partial_{n-1}}$, we obtain

$$\begin{aligned}
&\partial_n(X)(r \otimes r_1 \otimes \cdots \otimes r_n \otimes c) \\
&= r \cdot \left(r_1 \otimes \cdots \otimes r_n \otimes c - s_{n-1}\big(\sum_{i=1}^{n} (-1)^i r_1 \otimes \cdots \otimes r_i r_{i+1} \otimes \cdots \otimes r_n \otimes c\big) \right) \\
&= r r_1 \otimes \cdots \otimes r_n \otimes c - \sum_{i=1}^{n} (-1)^i r \otimes r_1 \otimes \cdots \otimes r_i r_{i+1} \otimes \cdots \otimes r_n \otimes c \\
&= \sum_{i=0}^{n} (-1)^i r \otimes r_1 \otimes \cdots \otimes r_i r_{i+1} \otimes \cdots \otimes r_n \otimes c
\end{aligned} \tag{B.18}$$

References

[1] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. 2015.

[2] Pierre Baudot and Daniel Bennequin. The homological nature of entropy. *Entropy*, 17(5):3253–3318, 2015.

[3] Daniel Bennequin and Juan Pablo Vigneaux. A functional equation related to generalized entropies and the modular group. *Aequationes mathematicae*, 94(6):1201–1212, March 2020.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006.

[5] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[6] A. Hatcher. *Algebraic Topology*. Algebraic Topology. Cambridge University Press, 2002.

[7] Pl Kannappan and Che Tat Ng. Measurable solutions of functional equations related to information theory. *Proceedings of The American Mathematical Society - PROC AMER MATH SOC*, 38:303–303, 04 1973.

[8] S. MacLane. *Homology*. Classics in Mathematics. Springer Berlin Heidelberg, 2012.

[9] J.P. Serre. *A Course in Arithmetic*. Graduate texts in mathematics. Springer, 1973.

[10] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[11] The Stacks project authors. The stacks project. https://stacks.math.columbia.edu, 2024.

[12] Helge Tverberg. A new derivation of the information function. *MATHEMATICA SCANDINAVICA*, 6:297–298, Dec. 1958.

[13] Juan Pablo Vigneaux. Information structures and their cohomology, 2021.

[14] Charles A Weibel. *An introduction to homological algebra*. Number 38. Cambridge university press, 1994.