INFORMATION STRUCTURES AND THEIR COHOMOLOGY

JUAN PABLO VIGNEAUX

ABSTRACT. We introduce the category of *information structures*, whose objects are suitable diagrams of measurable sets that encode the possible outputs of a given family of observables and their mutual relationships of refinement; they serve as mathematical models of contextuality in classical and quantum settings. Each information structure can be regarded as a ringed site with trivial topology; the structure ring is generated by the observables themselves and its multiplication corresponds to joint measurement. We extend Baudot and Bennequin's definition of *information cohomology* to this setting, as a derived functor in the category of modules over the structure ring, and show explicitly that the bar construction gives a projective resolution in that category, recovering in this way the cochain complexes previously considered in the literature. Finally, we study the particular case of a one-parameter family of coefficients made of functions of probability distributions. The only 1-cocycles are Shannon entropy or Tsallis α-entropy, depending on the value of the parameter.

Contents

1. Introduction

Entropy plays a fundamental role in several domains of mathematics and physics, and it is natural to ask why it is so important and ubiquitous. A pragmatic answer might

highlight the connections to limiting theorems in probability theory and dynamical systems; examples are the Shannon-Macmillan-Breiman theorem [Cover and Thomas, 2006, Thm. 16.8.1] or the quantification of large deviations in terms of relative entropy [Varadhan, 2003]. But entropy also possess remarkable algebraic properties, as Shannon already pointed out in the foundational article of information theory [Shannon, 1948]. He gave there an "axiomatic characterization" of entropy based on expected ("natural") properties of a measure of uncertainty. Since then, many authors have contributed with similar theorems, proposing alternative axiomatic characterizations of *information functions* (many of them summarized in [Csiszár, 2008]; see also [Khinchin, 1957; Otáhal, 1994; Hatori, 1958]); the list includes a recent category-theoretic article [Baez et al., 2011] that focuses on the *information loss* induced by reductions between finite probability spaces. Some of these works led to the study of functional equations uniquely solved by the entropy [Aczél and Daróczy, 1975], hence to very sophisticated techniques involving real analysis.

The purpose of the present article—which complements a previous one by Pierre Baudot and Daniel Bennequin [2015]—is to develop a new perspective that identifies entropy with a topological invariant of a finite statistical system. In particular, entropy appears as a cohomology class and not merely as a function. We introduce here a new definition of *information structures* (categories of observables) and cohomological invariants associated to certain presheaves on them, using the framework developed by Artin, Grothendieck, Verdier and their collaborators in the SGA 4 [Artin et al., 1972a,b]. We recall that toposes were introduced there as a general foundation of topology, that allowed a unified study of several cohomological invariants involving groups, topological spaces, and schemes. Our results constitute an extension of the field of application of these ideas.

1.1. ENTROPIES AND THEIR ALGEBRAIC CHARACTERIZATION. Shannon [1948] defined the information content of a random variable X, taking values in a finite set $\mathscr{E}_X$, by the formula

$$S_1[X](P) := -\sum_{x \in \mathscr{E}_X} P(X = x) \log P(X = x), \tag{1}$$

where P denotes a probability measure (law) on $\mathscr{E}_X$. The function S_1 is called (Gibbs-Shannon) entropy, and quantifies the uncertainty of a measurement.[1]

Given two random variables X and Y, valued respectively in sets $\mathscr{E}_X$ and $\mathscr{E}_Y$, their joint measurement (X, Y) is also random variable, valued in $\mathscr{E}_{XY} \subset \mathscr{E}_X \times \mathscr{E}_Y$. Following again [Shannon, 1948], a probability law P on $\mathscr{E}_{XY}$ can be represented by a tree as in Fig. 1-(a). The probability of observing $X = x$ is computed as the sum of all the outputs of (X, Y) that contain x in the first component: $X_*P(x) := P(X = x) = \sum_{(x,y) \in \mathscr{E}_{XY}} P(x, y)$. The probability X_*P on $\mathscr{E}_X$ is called *marginal law*. Instead of measuring directly (X, Y) one could measure first X, which constitutes a first random choice; the uncertainty that

---

[1]In information theory, it is customary to write $H(X)$ instead of $S_1[X](P)$. We have decided to reserve the use of H for cohomology, employing instead the letter S, common in physics (although our S_1 is adimensional). In turn, the presence of Tsallis α-entropies S_α justifies the subscript 1, as explained below. Finally, the functional equations involved in this work contain evaluations of S_1 at different laws, making necessary to mention the argument P explicitly.

remains after obtaining the result $X = x_0$ is represented by the *conditional probability law* $P|_{X=x_0} : \mathscr{E}_{XY} \to [0,1]$, given by

$$P|_{X=x_0}(x,y) := \begin{cases} \frac{P(x,y)}{X_*P(x_0)} & \text{if } x = x_0 \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

provided $X_*P(x_0) > 0$ (it remains undefined for x_0 in the maximal X_*P-null set). This iterated choice/measurement can in turn be pictured as a tree, e.g. Fig. 1-(b). The function S_1 satisfies the so-called *chain rule*

$$S_1[(X,Y)](P) = S_1[X](X_*P) + \sum_{\substack{x \in \mathscr{E}_X \\ X_*P(x)>0}} X_*P(x)S_1[Y](Y_*P|_{X=x}) \tag{3}$$

Evidently, if the measurement of Y is performed first, we obtain another tree, Fig. 1-(c), that corresponds to

$$S_1[(X,Y)](P) = S_1[Y](Y_*P) + \sum_{\substack{y \in \mathscr{E}_Y \\ Y_*P(y)>0}} Y_*P(y)S_1[X](X_*P|_{Y=y}) \tag{4}$$

Shannon [1948, p. 392-393] gave an algebraic characterization of

$$H_n : \Delta^n \to \mathbb{R}, \quad (p_0, ..., p_n) \mapsto -\sum_{i=0}^n p_i \log p_i \tag{5}$$

as the only family of continuous functions that satisfies the chain rule (3) for any possible tree—this is, for arbitrary pairs (X,Y), setting $S_1[X] = H_{|\mathscr{E}_X|}$ and so on—and such that $H_n(1/n, ..., 1/n)$ is monotonic in n.

It is worth noticing that several generalizations of entropy play a role in information theory and statistical mechanics. One of them is the structural α-entropy, defined for each $\alpha \in]0, \infty[\backslash\{1\}$ as

$$S_\alpha[X](P) = \frac{1}{1-\alpha}\left(\sum_{x \in \mathscr{E}_X} P(x)^\alpha - 1\right), \tag{6}$$

It was introduced axiomatically in 1967 by Havrda and Charvát [1967] (who characterized it up to a multiplicative constant). The use of α-entropies in statistical mechanics was proposed by Tsallis [1988], and the most common name for S_α is Tsallis α-entropy. This function satisfies the deformed equation

$$S_\alpha[(X,Y)](P) = S_\alpha[Y](Y_*P) + \sum_{\substack{y \in \mathscr{E}_Y \\ Y_*P(y)>0}} (Y_*P(y))^\alpha S_\alpha[X](X_*P|_{Y=y}). \tag{7}$$
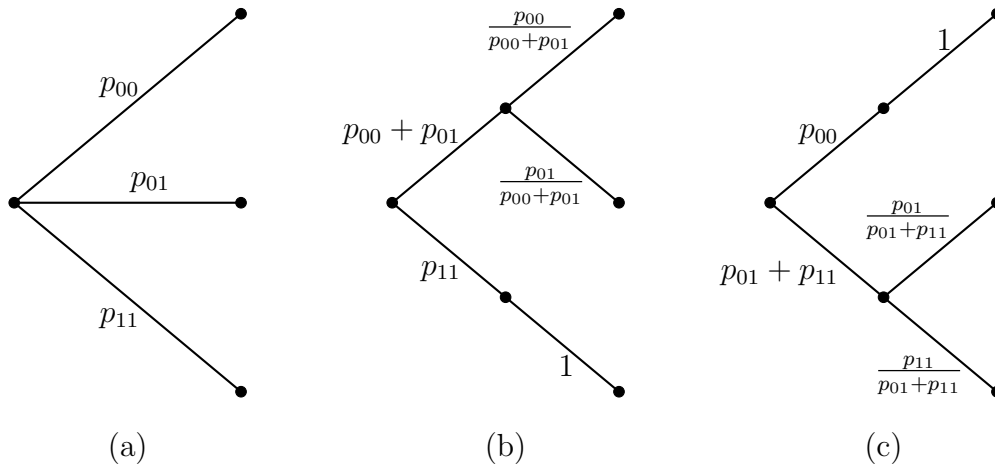
(a)   (b)   (c)

Figure 1: Different groupings when $\mathscr{E}_X = \mathscr{E}_Y = \{0,1\}$ and $\mathscr{E}_{XY} = \{(0,0),(0,1),(1,1)\}$. We denote by p_{ij} the probability of the point $(i,j) \in \mathscr{E}_{XY}$. In (b) and (c), the probabilities to the left are the marginals X_*P and Y_*P, respectively, and those to the right are the conditional laws on the appropriate subset of $\mathscr{E}_{XY}$.

Tverberg [1958] was the first to deduce from the chain rule a simple functional equation that characterized Shannon entropy, called "fundamental equation of information theory":

$$f(x) + (1-x)f\left(\frac{y}{1-x}\right) = f(y) + (1-y)f\left(\frac{x}{1-y}\right), \tag{8}$$

where $f : [0,1] \to \mathbb{R}$ is an unknown function such that $f(0) = f(1) = 1$, and $x, y \in [0,1)$ are such that $x + y \in [0,1]$. The only symmetric, measurable solutions of this equation are the real multiples of $s_1(x) := -x\ln(x) + (1-x)\ln(1-x)$ [Lee, 1964]. The result is quite remarkable, because Shannon's characterization of the functions H_n requires an *infinite* number of equations—for *any* random variable and *any* possible grouping of its outcomes—along with a strong regularity of H_n. Daróczy [1970] proposed a similar equation solved by the α-entropy $s_\alpha(x) := x^\alpha + (1-x)^\alpha - 1$.[2]

In the same vein, if the product (X,Y) is nondegenerate (see below), then the system of functional equations (3)-(4), with measurable *unknowns* $S_1[X]$, $S_1[Y]$, and $S_1[(X,Y)]$, is uniquely solved by the corresponding Shannon entropies (1), up to a multiplicative constant. This holds even for the situation pictured in Figure 1, that is evidently the simplest possible choice that can be broken down in two different ways (see Proposition 4.5.4).

More importantly, the chain-rule-like functional equations (3)-(4) accept a cohomological interpretation. Let us define, for any probabilistic functional $P \mapsto f(P)$, a new

---

[2]For a detailed historical introduction and a comprehensive treatment of the subject, up to 1975, see the book by Aczél and Daróczy [1975].

functional $X.f$ given by

$$(X.f)(P) := \sum_{x \in \mathscr{E}_X} X_* P(X) f(Y_* P|_{X=x}). \tag{9}$$

in order to rewrite (3) as

$$0 = X.S_1[Y] - S_1[(X, Y)] + S_1[X]. \tag{10}$$

The notation is meant to suggest an action of random variables on probabilistic functionals, and in fact the equality $Z.(X.f) = (Z, X).f$ holds. There is an strong resemblance between (10) and a cocycle equation in group cohomology. Baudot and Bennequin [2015] formalized this analogy introducing an adapted cohomology theory—information cohomology—through an explicit differential complex that recovered the equations (10) as 1-cocycle conditions. They used presheaves, exploiting a notion of *locality* specific to the problem: the entropy of a variable X only depends on the marginalized version $X_* P$ of any global law P.[3]

1.2. CATEGORIES OF OBSERVABLES. Information cohomology was introduced in [Baudot and Bennequin, 2015] considering presheaves on *information structures*, that were either categories of partitions of a given measurable space or categories of orthogonal decompositions of a Hilbert space. The partitions corresponded to atoms of the σ-algebras generated by measurable functions (classical observables) with finite range, and the orthogonal decompositions appeared as eigenspaces of self-adjoint operators (quantum observables) with finitely many different eigenvalues.

Inspired by [Gromov, 2012], we wanted to approach measurements from a categorical viewpoint, describing directly the relations between their outputs and without presupposing the existence of an underlying probability space or Hilbert space.[4] A probability space is only necessary to represent a collection of observables by measurable functions with a common domain, as customary in "classical" probability theory (as opposed to "quantum"). However, the existence of such representation is not trivial: some collections of variables are contextual (see Section 2.3) and therefore violate generalized Bell inequalities [Abramsky and Hardy, 2012, Prop. III.1], which make them incompatible with such *classical representations*. The sets of outputs can also be interpreted as the spectra

---

[3]Cathelineau [1988] was the first to find a cohomological interpretation for the fundamental equation (8): an analogue of it is involved in the computation of the homology of SL_2 over a field of characteristic zero, with coefficients in the adjoint action; however, this result was not explicitly connected to Shannon entropy or information theory. The first published work in this direction is a note by Kontsevich (reproduced as an appendix in [Elbaz-Vincent and Gangl, 2002]), that introduces $H_p(x) = \sum_{k=1}^{p-1} \frac{x^k}{k}$ as "a residue modulo p" of entropy, being the only continuous map $f : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ that verifies $f(x) = f(1 - x)$ and an equation equivalent to (8). He proves that a related function defines a cohomology class in $H^2(F, F)$, for $F = \mathbb{R}$ or $\mathbb{Z}/p\mathbb{Z}$. Several works connected to motives or polylogarithms have emphasized the role of the fundamental equation, for instance [Cathelineau, 1996; Elbaz-Vincent and Gangl, 2002, 2015; Bloch and Esnault, 2003].

[4]A similar approach is taken in [Matveev and Portegies, 2018].

of self-adjoint operators, in such a way that some contextual collections have *quantum representations*.

In view of the foregoing, we introduce here a more general definition of information structure, that covers the classical and quantum cases at the same time and extends without modification to continuous random variables (see [Vigneaux, 2019]). This allows us to introduce a category of information structures and to treat the algebraic aspects of the theory (Section 3) in a unified manner, once for all these cases. To attain this flexibility and generality, the definition decouples the combinatorial structure of joint measurements and the local models of the outputs of each individual measurement.

Let $\mathbf{Meas}_{\mathrm{surj}}$ be the category of measurable spaces and measurable surjections between them.

1.2.1. DEFINITION. *A* conditional meet semilattice *is a poset[5] that satisfies the following property:*

> *for any $X, Y, Z \in \mathrm{Ob}\,\mathbf{S}$, if $Z \to X$ and $Z \to Y$, then the categorical product $X \wedge Y$ exists.* $\qquad$ (11)

It is unital *whenever it has a terminal object, denoted $\top$.*

An information structure *is a pair $(\mathbf{S}, \mathscr{M})$, where $\mathbf{S}$ is a unital conditional meet semilattice and $\mathscr{M} : \mathbf{S} \to \mathbf{Meas}_{\mathrm{surj}}$ is a functor[6] (say $\mathscr{M}_X = (\mathscr{E}_X, \mathfrak{B}_X)$, for each $X \in \mathrm{Ob}\,\mathbf{S}$) that satisfies:*

1. *$\mathscr{E}_\top \cong \{*\}$, with the trivial σ-algebra;*

2. *for every $X \in \mathrm{Ob}\,\mathbf{S}$ and any $x \in \mathscr{E}_X$, the σ-algebra $\mathfrak{B}_X$ contains the singleton $\{x\}$;*

3. *for every diagram $X \xleftarrow{\;\pi\;} X \wedge Y \xrightarrow{\;\sigma\;} Y$ the measurable map*

$$\mathscr{M}_{X \wedge Y} \hookrightarrow \mathscr{M}_X \times \mathscr{M}_Y, z \mapsto (x(z), y(z)) := (\mathscr{M}\pi(z), \mathscr{M}\sigma(z))$$

   is an injection.

The objects of the conditional meet semilattice $\mathbf{S}$ stand for observables and the arrows encode the relation of *refinement* between them (think of refinements of σ-algebras or orthogonal decompositions). The terminal object is "certainty", the meet $X \wedge Y$ represents the joint measurement of X and Y, and condition (11) accommodates the impossibility of doing some joint measurements. For instance, in quantum mechanics, it is only possible to jointly measure X and Y if they commute, in which case the observable (X, Y) induces

---

[5]A partially ordered set (poset) is set C with a binary relation $\leq$ that is reflexive, antisymmetric and transitive. Equivalently, it is a small category $\mathbf{C}$ such that:

1. for any pair of objects A, B, there is at most one morphism from A to B, and

2. if there is a morphism from A to B and a morphism from B to A, then $A = B$.

We move freely between both descriptions.

[6]Given a functor $\mathscr{F} : \mathbf{S} \to \mathbf{Sets}$, we denote its value at $X \in \mathrm{Ob}\,\mathbf{S}$ by $\mathscr{F}(X)$ or $\mathscr{F}_X$.

an orthogonal decomposition of the Hilbert space that refines the decompositions induced by X and Y. In turn, the functor $\mathscr{M}$ represents the possible outputs of each observable. A refinement $\pi : X \to Y$ translates into a surjection $\pi_* \equiv \mathscr{M}\pi : \mathscr{M}_X \to \mathscr{M}_Y$ that induces an injection at the level of the algebras of events $\pi^* : \mathfrak{B}_Y \to \mathfrak{B}_X$ that maps A to $\mathscr{M}\pi^{-1}(A)$;[7]compare with the *extensions* of probability spaces discussed in [Tao, 2012, p. 3]. The set $\mathscr{E}_{X \wedge Y}$ represents the possible outputs of the joint measurement $X \wedge Y$, hence it can be identified with a subset of $\mathscr{E}_X \times \mathscr{E}_Y$ as in Section 1.1. When convenient, we use the notations common in probability theory: $\{X = x\}$ means "the element x contained in $\mathscr{E}_X$" and $\{X = x, Y = y\}$ should be interpreted as *the* element z of $\mathscr{E}_{X \wedge Y}$ mapped to x by $\mathscr{E}_{X \wedge Y} \to \mathscr{E}_X$ and to y by $\mathscr{E}_{X \wedge Y} \to \mathscr{E}_Y$ (if such z does not exist, $\{X = x, Y = y\} = \emptyset$).

In Section 2.1 we also define the morphisms between information structures and prove that the category **InfoStr** thus obtained has countable products and arbitrary coproducts (Proposition 2.2.2).

Probability laws come as a functor $\mathscr{P} : \mathbf{S} \to \mathbf{Sets}$ that associates to each $X \in \mathrm{Ob}\,\mathbf{S}$ the set $\mathscr{P}_X$ of measures P on $\mathscr{M}_X$ such that $P(\mathscr{E}_X) = 1$. Each arrow $\pi : X \to Y$ induces a measurable surjection $\mathscr{M} : \mathscr{M}_X \to \mathscr{M}_Y$, and $\mathscr{P}\pi : \mathscr{P}_X \to \mathscr{P}_Y$ is defined to be the push-forward of measures: for every $B \in \mathfrak{B}_Y$,

$$(\mathscr{P}\pi(P))(B) = P(\mathscr{M}\pi^{-1}(B)). \tag{12}$$

This operation is called *marginalization*. We write π_* or Y_* instead of $\mathscr{P}\pi$, if there is no risk of ambiguity; this notation is compatible with that of Section 1.1.

1.2.2. EXAMPLE. [Simplicial information structures] If I is any set, let $\mathbf{\Delta}(\mathbf{I})$ be the poset of its finite subsets, with an arrow $A \to B$ whenever $B \subset A$. A simplicial subcomplex of $\mathbf{\Delta}(\mathbf{I})$ is a full subcategory $\mathbf{K}$ such that, for any given object of $\mathbf{K}$ ("a cell"), all its subsets are also objects of $\mathbf{K}$ ("faces"). Given a collection $\{(\mathscr{E}_i, \mathfrak{B}_i)\}_{i \in I}$ of masurable spaces, let $\mathscr{M} : \mathbf{\Delta}(\mathbf{I}) \to \mathbf{Sets}$ be the functor that associates to each $A \subset I$ the set $\mathscr{E}_A := \prod_{i \in A} \mathscr{E}_i$ with the product σ-algebra $\mathfrak{B}_A := \bigotimes_{i \in A} \mathfrak{B}_i$, and to each arrow in $\mathbf{\Delta}(\mathbf{I})$ the corresponding canonical projector. The pair $(\mathbf{K}, \mathscr{M}|_{\mathbf{K}})$ is a *simplicial information structure*.[8]

A particular case of this construction—such that all the $\mathscr{E}_i$ are equal—appears in [Abramsky and Brandenburger, 2011], which introduces a sheaf-theoretic treatment of nonlocality and contextuality. There, the elements of I are called *measurements*, the minimal objects in the poset $\mathbf{K}$ are called *maximal measurement contexts* (without loss of generality, it is supposed that the 0-skeleton of $\mathbf{K}$ is I itself), and the sheaf $\mathscr{M} : \mathbf{\Delta}(\mathbf{I}) \to \mathbf{Sets}$ is called *sheaf of events*. The follow-up article [Abramsky et al., 2015] introduces a more general notion of possible events, allowing any subfunctor $\mathscr{N} : \mathbf{\Delta}(\mathbf{I}) \to \mathbf{Sets}$ of $\mathscr{M}$

---

[7]Proof: If $\pi^*(A) = \pi^*(B)$, then $\pi^{-1}(A \Delta B) = \emptyset$, where Δ is the symmetric difference. Since $\mathscr{M}\pi$ is surjective, this implies that $A \Delta B = \emptyset$ i.e. $A = B$. Of course, one could relax the surjectivity introducing ideals of negligible sets (e.g. through a reference measure) and asking the preimage under $\mathscr{M}\pi$ of any negligible set to be negligible, but—at least in the discrete case—there is no loss of generality in supposing directly that $\mathscr{M}\pi$ is a surjection.

[8]It is worth noting that *abelian* (co)presheaves on some complex $\mathbf{K}$ are cellular (co)sheaves in the sense of [Curry, 2013].

that is "flasque beneath the cover"—which means that $\mathcal{N}\pi$ is surjective for any arrow π in $\mathbf{K}$—and such that any section of $\mathcal{N}|_{\mathbf{K}}$ induces an element of $\mathcal{N}(I)$. Our definition of information structure further generalizes these ideas: it can be applied to nonsimplicial settings (see the examples in Section 2.1) and it does not make reference to global facts (which are treated in the context of representations, see Section 2.3).[9]

An information structure $(\mathbf{S}, \mathcal{M})$ is called finite if all the measurable sets $(\mathcal{E}_X, \mathfrak{B}_X)$ in the image of $\mathcal{M}$ are finite. Section 2.3 studies the conditions under which the observables of a finite information structure $(\mathbf{S}, \mathcal{M})$ can be represented as measurable functions on a unique sample space $(\Omega, \mathfrak{F})$. A necessary and sufficient condition is the existence of a global section $s(x)$ of $\mathcal{E}$ compatible with any given value $x \in \mathcal{E}_X$ assigned to any observable X. For the sake of completeness, we also define quantum representations.

1.3. COHOMOLOGICAL CHARACTERIZATION OF α-ENTROPIES. Let $\mathbf{S}$ be a conditional meet semilattice, and $\mathscr{S}$ be the presheaf of monoids that maps each $X \in \mathrm{Ob}\,\mathbf{S}$ to the set $\mathscr{S}_X \equiv \mathscr{S}(X) := \{ Y \in \mathrm{Ob}\,\mathbf{S} : X \to Y \}$ equipped with the product $(Y, Z) \mapsto YZ := Y \wedge Z$, and each arrow $X \to Y$ in $\mathbf{S}$ to the inclusion $\mathscr{S}_Y \hookrightarrow \mathscr{S}_X$. There is an associated presheaf of induced algebras $X \mapsto \mathscr{A}_X := \mathbb{R}[\mathscr{S}_X]$. The category of $\mathscr{A}$-modules—abelian presheaves with a natural action of $\mathscr{A}$—is abelian and has enough injective objects, hence it is possible to define right derived functors using standard tools in homological algebra [Grothendieck, 1957; Weibel, 1994].

A particular example of $\mathscr{A}$-module is the space of probabilistic functionals that appear in Section 1.1; the $\mathscr{A}$-action was defined by (9). Another is the trivial $\mathscr{A}$-module $\mathbb{R}_{\mathbf{S}}$, that associates to each object $X \in \mathrm{Ob}\,\mathbf{S}$ the set $\mathbb{R}$ with trivial $\mathscr{A}_X$ action and to every arrow in $\mathbf{S}$ the identity map.

---

[9]Let $\mathscr{P}_{\mathbf{K}}$ be the functor of probabilities on $\mathcal{M}|_{\mathbf{K}}$. The sections of $\mathscr{P}_{\mathbf{K}}$ are the elements of $\Gamma(\mathbf{K}, \mathscr{P}_{\mathbf{K}}) := \mathrm{Hom}_{[\mathbf{K},\mathbf{Sets}]}(*, \mathscr{P}_{\mathbf{K}})$, where $[\mathbf{K}, \mathbf{Sets}]$ is the category of $\mathbf{Sets}$-valued functors on $\mathbf{K}$ and $*$ is the functor that associates to each $X \in \mathrm{Ob}\,\mathbf{K}$ a singleton. An element $s \in \Gamma(\mathbf{K}, \mathscr{P}_{\mathbf{K}})$ is a collection of probabilities that are mutually compatible under marginalizations; this appears in the literature as *pseudo-marginals* [Vontobel, 2013] or *no-signaling empirical models* [Abramsky and Brandenburger, 2011], among other names. Remark that, in the simplicial case, $\mathcal{M}$ and the sheaf $\mathscr{P}$ of probabilities on it are naturally defined on the whole category $\mathbf{\Delta}(\mathbf{I})$, which can be seen as a larger geometrical space that contains $\mathbf{K}$. The *marginal problem* consists in determining when a pseudo-marginal s on $\mathbf{K}$ can be extended to a section $\tilde{s}$ of $\mathscr{P}$ on $\mathbf{\Delta}(\mathbf{I})$ that coincides with s over each $A \in \mathrm{Ob}\,\mathbf{K}$; this extension $\tilde{s}$ is meant to represent a joint state of the observables indexed by I, compatible with the known local interactions. It is well known that such extension does not always exist: this correspond to *frustration* in statistical mechanics [Pelizzola, 2005; Matsuda, 2001], *contextuality* in quantum mechanics, and paradoxes in logic [Abramsky et al., 2015; Abramsky and Hardy, 2012; Fritz and Chaves, 2013]. The problem also appears in the context of *graphical models*, where the belief propagation algorithm converges to a pseudo-marginal on the smallest simplicial subcomplex that contains all the factors of the model (we are identifying here a factor node in a graphical model with its boundary, which is a subset of the variable nodes), see [Pelizzola, 2005] and [Mézard and Montanari, 2009, Ch. 9]. A sheaf-theoretic treatment of the marginal problem was introduced in [Abramsky and Brandenburger, 2011]. The determination of eventual connections between information cohomology and their cohomology of contextuality is an exciting open problem.

1.3.1. DEFINITION. *The information cohomology of* **S** *with coefficients in the* $\mathscr{A}$*-module* $\mathscr{F}$ *is*

$$H^\bullet(\mathbf{S}, \mathscr{F}) := \operatorname{Ext}^\bullet(\mathbb{R}_\mathbf{S}, \mathscr{F}). \tag{13}$$

Recurring to the (relative) bar resolution [Mac Lane, 1994, Ch. IX], we obtain a computable version of this cohomology. The bar construction gives a resolution

$$0 \longleftarrow \mathbb{R}_\mathbf{S} \xleftarrow{\ \epsilon\ } \mathscr{B}_0 \xleftarrow{\ \partial_1\ } \mathscr{B}_1 \xleftarrow{\ \partial_2\ } \mathscr{B}_2 \xleftarrow{\ \partial_3\ } \dots \ , \tag{14}$$

where each $\mathscr{B}_i$ is a *relative* projective $\mathscr{A}$-module (see the appendix). More explicitly, for each $X \in \operatorname{Ob}\mathbf{S}$ and $n \in \mathbb{N}$, the module $\mathscr{B}_n(X)$ is freely generated over $\mathscr{A}_X$ by $\{\,[X_1|...|X_n] : X_1, ..., X_n \in \mathscr{S}_X\,\}$; in the case of $\mathscr{B}_0(X)$, simply by the "empty" symbol $[\,]$. Proposition 3.2.2 proves that, due to the conditional existence of products in the definition of $\mathbf{S}$, these $\{\mathscr{B}_i\}_i$ are projective objects in the category $\mathbf{Mod}(\mathscr{A})$, which in turn implies that $H^\bullet(\mathbf{S}, \mathscr{F})$ can be identified with the cohomology of the differential complex $(C^\bullet(\mathbf{S}, \mathscr{F}), \delta)$, where $C^n(\mathbf{S}, \mathscr{F}) := \operatorname{Hom}_\mathscr{A}(\mathscr{B}_n, \mathscr{F})$ and $\delta f := f\partial$. An element $f \in C^n(\mathbf{S}, \mathscr{F})$ consists of several components $\{f_X : \mathscr{B}_n(X) \to \mathscr{F}(X)\}_{X \in \operatorname{Ob}\mathbf{S}}$. Each map $\delta \equiv \delta^n : C^n(\mathbf{S}, \mathscr{F}) \to C^{n+1}(\mathbf{S}, \mathscr{F})$ is defined as follows: for each $X \in \operatorname{Ob}\mathbf{S}$, $X_1, .., X_{n+1} \in \mathscr{S}_X$, and $f \in C^n(\mathbf{S}, \mathscr{F})$,

$$(\delta f)_X[X_1|...|X_{n+1}] = X_1.f_X[X_2|...|X_{n+1}] + \sum_{k=1}^n (-1)^k f_X[X_1|...|X_k X_{k+1}|...|X_n]$$
$$+ (-1)^{n+1} f_X[X_1|...|X_n]. \tag{15}$$

To simplify notation, we write $f_X[X_2|...|X_{n+1}]$ instead of $f_X([X_2|...|X_{n+1}])$, etc.

We introduce a particular family of $\mathscr{A}$-modules, made of probabilistic functionals. Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure, and let $\mathscr{Q}$ denote any subfunctor of $\mathscr{P}$ stable under conditioning (*adapted probability functor*). We represent every probability law by its density P with respect to the counting measure, which can be seen as a vector in $\mathbb{R}^{\mathscr{E}_X}$, so that $\mathscr{P}_X$ is the standard (probability) simplex $\Delta(\mathscr{E}_X)$ in $\mathbb{R}^{\mathscr{E}_X}$. Let $\mathscr{F}(X)$ be the additive abelian group of measurable[10] real-valued functions on $\mathscr{Q}(X)$ and, for any arrow $\pi : X \to Y$, let $\mathscr{F}\pi : \mathscr{F}(Y) \to \mathscr{F}(X)$ be precomposition with marginalization: $\mathscr{F}\pi(f) = f \circ \pi_*$. We obtain in this way a contravariant functor $\mathscr{F}$ on $\mathbf{S}$.

For each $Y \in \mathscr{S}_X$, $f \in \mathscr{F}(X)$ and $P \in \mathscr{Q}(X)$, define

$$(Y.f)(P) = \sum_{\substack{y \in \mathscr{E}_Y \\ Y_* P(y) \neq 0}} (Y_* P(y))^\alpha f(P|_{Y=y}). \tag{16}$$

---

[10]If $(\Omega, \mathfrak{F})$ is a measurable space and A is a subset of Ω, then $\mathfrak{F}' = \{\,A \cap F : F \in \mathfrak{F}\,\}$ is a σ-algebra of subsets of A, *induced* by $\mathfrak{F}$. For each $X \in \operatorname{Ob}\mathbf{S}$, we identify $\mathscr{P}(X)$ with the standard simplex $\Delta(\mathscr{E}_X)$ equipped with the σ-algebra $\mathfrak{S}$ induced by the Borel σ-algebra on $\mathbb{R}^{|\mathscr{E}_X|}$ (with its standard topology); equivalently, $\mathfrak{S}$ is the Borel σ-algebra associated to the subspace topology on $\Delta(\mathscr{E}_X)$. Then $\mathscr{Q}(X) \subset \mathscr{P}(X)$ becomes a measurable space with the σ-algebra induced by $\mathfrak{S}$.

This turns each $\mathscr{F}(X)$ into an $\mathscr{A}_X$-module and this action is functorial, in such a way that $\mathscr{F}$ becomes an $\mathscr{A}$-module $\mathscr{F}_\alpha$ (for any $\alpha > 0$). We call *probabilistic* the information cohomology with coefficients in some $\mathscr{F}_\alpha$. With reference to the cochain complex $(C^\bullet(\mathbf{S}, \mathscr{F}_\alpha), \delta)$ introduced above, we can determine the following facts.

A 0-cochain is a collection $\{f_X[\,]\}_{X \in \mathrm{Ob}\, S}$ that is local: for any X, $f_X[\,](P) = f_\top[\,](1)$, so $f_X[\,]$ equals a constant $K \in \mathbb{R}$. The 0-cochain f is a 0-cocycle if for any $X \to Y$ in $\mathbf{S}$ $(\delta f)_X[Y] = Y.f_X[\,] - f_X[\,]$, which evaluated on a probability $P \in \mathscr{Q}(X)$ reads

$$(\delta f)_X[Y](P) = \sum_{y \in \mathscr{E}_Y} (Y_* P(y))^\alpha K - K = \begin{cases} 0 & \text{if } \alpha = 1 \\ K S_\alpha[Y](Y_* P) & \text{otherwise} \end{cases}.$$

In other words: every cochain is a 0-cocycle if $\alpha = 1$; there are no 0-cocycles if $\alpha \neq 1$, but Tsallis entropy appears as 1-coboundary multiplied by a global constant K.

The 1-cochains are characterized by collections of functionals

$$\{f[X] : \mathscr{Q}(X) \to \mathbb{R}\}_{X \in \mathrm{Ob}\, \mathbf{S}},$$

which takes into account the naturality of f: $f_Y[X](P) = f_X[X](X_* P) =: f[X](X_* P)$. The 1-cocycles additionally satisfy

$$0 = X.f[Y] - f[XY] + f[X] \tag{17}$$

as functions on $\mathscr{Q}(XY)$, where marginalizations are implicit. As explained in Section 1.1, this equation and its analogue with X and Y permuted imply that $f[\cdot] = K_{XY} S_\alpha[\cdot]$, for a constant $K_{XY} \in \mathbb{R}$. This holds as long as $\mathscr{Q}_{XY}$ contains enough probabilities: a precise sufficient condition is stated in the definition of nondegeneracy for the product of two observables (Definition 4.5.6). Despite being quite involved, this definition is one of the main contributions of this article. It is instrumental to write a more explicit proof of [Baudot and Bennequin, 2015, Thm. 1] and its generalization, Theorem 1.3.2. If an observable Z can be written as a nondegenerate product, we say that it is *nontrivially reducible*.

Remind that a category $\mathbf{J}$ is connected if any two objects $j, k \in \mathbf{J}$ can be joined by a finite sequence of arrows

$$j = j_0 \leftarrow j_1 \to j_2 \leftarrow \cdots \leftarrow j_{2n-1} \to j_{2n} = k. \tag{18}$$

Every category $\mathbf{J}$ is a disjoint union (coproduct in $\mathbf{Cat}$) of connected categories $\mathbf{J}_k$, called *connected components* [Mac Lane, 1998, p. 90]. We denote by $\pi_0(\mathbf{J})$ the set of connected components of a small category $\mathbf{J}$.

1.3.2. THEOREM. *Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure, and let $\mathscr{Q}$ be an adapted probability functor such that all marginalization maps are surjective. Denote by $\mathbf{S}^*$ the full subcategory of $\mathbf{S}$ generated by $\{ X \in \mathrm{Ob}\, \mathbf{S} : |\mathscr{E}_X| > 1 \}$. Suppose that for every $X \in \mathrm{Ob}\, \mathbf{S}^*$ there exists a nontrivially reducible object $Z \in \mathrm{Ob}\, \mathbf{S}$ such that $Z \to X$. Then, there is*

a linear isomorphism $\varphi : \mathbb{R}^{\pi_0(\mathbf{S}^*)} \to Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ *that maps* $\Lambda = (\lambda_\mathbf{C})_{\mathbf{C} \in \pi_0(\mathbf{S}^*)}$ *to the natural transformation* S_α^Λ *defined as follows: for every* $X \in \mathrm{Ob}\,\mathbf{S}$ *and* $Y \in \mathscr{S}_X$,

$$(S_\alpha^\Lambda)_X[Y] = \begin{cases} \lambda_\mathbf{C}(S_\alpha)_X[Y] & \text{if } Y \in \mathrm{Ob}\,\mathbf{C} \\ 0 & \text{if } Y \notin \mathrm{Ob}\,\mathbf{S}^* \end{cases}. \tag{19}$$

Under this isomorphism, $\delta C^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q}))$ *corresponds to* $\langle 0 \rangle \subset \mathbb{R}^{\pi_0(\mathbf{S})}$, *whereas for* $\alpha \neq 1$, $\delta C^0(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ *corresponds to the diagonal* $\Delta \subset \mathbb{R}^{\pi_0(\mathbf{S})}$. *Hence* $H^1(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) \cong \mathbb{R}^{\pi_0(\mathbf{S}^*)}$ *and* $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \cong \mathbb{R}^{\pi_0(\mathbf{S}^*)}/\Delta$ *when* $\alpha \neq 1$.

One may say that a pair $(\mathbf{S}, \mathscr{Q})$ is *strongly connected* if $\mathbf{S}^*$ is connected and every object can be refined by a nontrivially reducible one. Shannon entropy is to some extent analogous to the fundamental class of an orientable connected manifold: a generator of the one-dimensional $H^1(\mathbf{S}, \mathscr{F}(\mathscr{Q}))$ for a strongly connected pair $(\mathbf{S}, \mathscr{Q})$.

For a simplicial structure $(\mathbf{K}, \mathscr{M})$, the theorem says that

$$H^1(\mathbf{S}, \mathscr{F}_1(\mathscr{P})) \cong \mathbb{R}^{\beta_0(\mathbf{K})} \quad \text{and} \quad H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \cong \mathbb{R}^{\beta_0(\mathbf{K})-1} \text{ when } \alpha \neq 1, \tag{20}$$

where $\beta_0(\mathbf{K})$ is the 0-th Betti number of (the geometric realization of) $\mathbf{K}$. Higher cohomology groups might be linked to the higher Betti numbers.

When some of the minimal objects of $\mathbf{S}$ is irreducible (it cannot be written as a nontrivial product), the group $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ has infinite dimension; Proposition 4.6.3 makes this precise. In 4.6 we also study other pathological examples.

At the end, we give an interpretation of H^0, H^1 and H^2 in terms of invariant sections, crossed homomorphism, and extensions, respectively, following the classic arguments about Hochschild cohomology.

2. The category of information structures

2.1. **Terminology and examples.** An information structure $(\mathbf{S}, \mathscr{M})$—see Definition 1.2.1—is said to be *bounded* if the poset $\mathbf{S}$ has finite height. It is *finite* if all the sets $\mathscr{E}_X$ are finite, in which case $\mathscr{E}_X$ corresponds to the atoms of $\mathfrak{B}_X$ and the algebra can be omitted from the description. We denote a finite structure by $(\mathbf{S}, \mathscr{E})$, where $\mathscr{E}$ is a covariant functor from $\mathbf{S}$ to $\mathbf{Sets}$. The cohomological computations in Section 4 concern finite structures, but the general constructions in Section 3—among them the definition of information cohomology—do not require this hypothesis. In fact, they only depend on the combinatorial object $\mathbf{S}$.

2.1.1. **Example.** [Concrete structures] The motivating example for the theory is the original version of information structures introduced in [Baudot and Bennequin, 2015].

Given a set Ω, let $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ be the category *finite observables*; the objects of this category are finite partitions of Ω, and there is an arrow $X \to Y$ whenever X refines Y. In this case, X discriminates better between the *configurations* $w \in \Omega$. The category $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ has a terminal object: the trivial partition $\top := \{\Omega\}$. When Ω is finite, it

also has an initial object: the partition by points, that we denote by $\bot$. The categorical product $X \times Y$ of two partitions X and Y is the coarsest partition that refines both. This product is commutative, associative, idempotent and unitary ($\top \times X = X$).

A classical *information structure* in the sense of [Baudot and Bennequin, 2015] is a full subcategory $\mathbf{S}$ of $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ such that

- $\top \in \mathrm{Ob}\,\mathbf{S}$;

- for any X, Y, Z in $\mathrm{Ob}\,\mathbf{S}$, if $X \to Y$ and $X \to Z$, then $Y \times Z$ belongs to $\mathbf{S}$.

We call $\mathbf{S}$ a *concrete structure*. If $\square : \mathbf{Obs}_{\mathrm{fin}}(\Omega) \to \mathbf{Sets}$ denotes the "forgetful" functor that maps the partition $X = \{A_1, ..., A_n\}$ to the set $\{A_1, ..., A_n\}$ and each arrow $X \to Y$ in $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ to the unique surjective map $\square\pi : \mathscr{E}_X \to \mathscr{E}_Y$ such that $B = \cup_{A \in \mathscr{E}\pi^{-1}(B)} A$ for any $B \in \mathscr{E}_Y$, the pair $(\mathbf{S}, \square)$ is a finite information structure according to Definition 1.2.1.

Concrete structures turn out to be too restrictive. For instance, Baudot and Bennequin associate to any finite indexed collection $\Sigma = (S_1, ..., S_n)$ of partitions of Ω a "simplicial structure" $\mathbf{S}(\mathbf{K})$: a subcategory of $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ that contains $\prod_{i \in A} S_i$ for any object A of a simplicial subcomplex $\mathbf{K}$ of the abstract simplex $\Delta(\{1, ..., n\})$—see the notation introduced in Example 1.2.2; by convention, the empty product gives the trivial partition. Such construction does not necessarily give an information structure (in their sense). For example: if $n = 3$, $\Omega = \{0, 1\}^2$, S_i is the partition induced by the projection on the i-th component ($i = 1, 2$), $S_3 = \{\{(0,0)\}, \{(0,0)\}^c\}$, and the maximal cells of $\mathbf{K}$ are $\{1, 2\}$ and $\{3\}$, then $S_1 \times S_2$ is the atomic partition, that refines all the others, while some products (like $S_1 \times S_3$) are not in $\mathbf{S}(\mathbf{K})$. In our framework, the category $\mathbf{S}(\mathbf{K})$ appears as a classical representation (cf. Section 2.3) of the (generalized) information structure $(\mathbf{K}, \mathscr{E})$, where $\mathscr{E} : \mathbf{K} \to \mathbf{Sets}$ is given by $\mathscr{E}_{\{1\}} = \mathscr{E}_{\{2\}} = \mathscr{E}_{\{3\}} = \{0, 1\}$, $\mathscr{E}_{\{1,2\}} = \mathscr{E}_{\{1\}} \times \mathscr{E}_{\{2\}}$, the maps induced by the arrows in $\mathbf{K}$ being canonical projections.

2.1.2. EXAMPLE. [Homogeneous structures] Let G be a locally compact, Hausdorff topological group. Any collection $\mathfrak{C}$ of closed subgroups of G that contains G and is conditionally closed under intersections (i.e. for any $M, N, O \in \mathfrak{C}$, if $N \subset M$ and $N \subset O$, then $M \cap O \in \mathfrak{C}$) defines a conditional meet semilattice $\mathbf{S}$, whose arrows correspond to inclusions. Let $\mathscr{M}$ be the functor that associates to each subgroup N the (Hausdorff) quotient space G/N with the Borel σ-algebra induced by the quotient topology, and to each arrow $N \to M$ the canonical projection $\pi_{M,N} : G/N \to G/M$ that sends the coset gN to gM. The information structures $(\mathbf{S}, \mathscr{M})$ obtained in this way are called *homogeneous*, because each coset space G/M is a homogeneous space for G.[11]

A particular example of this construction was introduced by the author in [Vigneaux, 2019, Ch. 11], where $G = (\mathbb{R}^n, +)$. The resulting pairs $(\mathbf{S}, \mathscr{M})$ are called there *Grassmannian structures* and play a key role in the computation of the information cohomology associated to continuous observables with gaussian laws.

---

[11]Moreover, the diagrams $\mathscr{M}(\mathbf{S})$ obtained for finite G are exactly the *minimal homogeneous diagrams* in [Matveev and Portegies, 2018, Sec. 2.7], provided each quotient G/N is equipped with the uniform measure. It is explained there that homogeneous diagrams approximate asymptotically any diagram of probability spaces.

2.1.3. RELATION WITH IDEMPOTENT MONOIDS. Recall that a monoid $(M, \cdot, e)$ is *idempotent* if for all $m \in M$, $m \cdot m = m$. Any conditional meet semilattice **S** induces a presheaf of idempotent monoids on it: for each $X \in \mathrm{Ob}\,\mathbf{S}$, set $\mathscr{S}_X := \{Y \in \mathrm{Ob}\,\mathbf{S} \mid X \to Y\}$, with the monoid structure given by the product of observables in **S**: $(Z, Y) \mapsto ZY := Z \wedge Y$; an arrow $X \to Y$ in **S** induces an inclusion $\mathscr{S}_Y \hookrightarrow \mathscr{S}_X$.

Furthermore, there is a well-known equivalence between idempotent monoids and meet semilattices with a terminal object. For a proof, see e.g. [Connes and Consani, 2019, Prop. 2.1].

2.1.4. PROPOSITION. *If $(M, \cdot, e)$ is an idempotent monoid, then the condition*

$$x \leq y \Leftrightarrow x \cdot y = x \tag{21}$$

defines a partial order on M such that any two elements of M have a meet and e is the greatest element.

Conversely, if $(E, \leq)$ is a poset with a greatest element in which any two elements $x, y \in E$ have a meet $x \wedge y$, then E endowed with the addition $(x, y) \mapsto x \wedge y$ is an idempotent monoid.

The two functors just introduced are inverses of each other.

Is there a counterpart to conditional meet semilattices with a terminal object in the theory of idempotent monoids? The following result serves as a partial answer. It involves *upper sets* of an idempotent monoid: a subset H of an idempotent monoid M—equipped with the partial order in (21)—is called an *upper set* if $h \in H$ and $h \leq m$ implies that $m \in H$. For example, the simplicial subcomplex **K** in Example 1.2.2 defines an upper set of $\mathbf{\Delta}(\mathbf{I})$, seen as an idempotent monoid according to Proposition 2.1.4.

2.1.5. PROPOSITION. *Let $(M, \cdot, e)$ be an idempotent monoid, and **M** its associated poset (seen as a category). The full subcategory of **M** defined by any nonempty upper set H of M is a unital conditional meet semilattice.*

PROOF. First, $e \in H$, because e is greater than any element of H. Second, if x, y, z be elements of H such that $z \leq x$ and $z \leq y$, then $z \leq x \wedge y$ in virtue of the universal property of $\wedge$ in **M**, which in turn implies that $x \wedge y \in H$ (by definition of upper set). ∎

Based on this result, new examples of information structures may arise in connection with *idempotent mathematics* [Litvinov, 2007] and algebra "over $\mathbb{F}_1$" [Connes and Consani, 2019].

2.2. MORPHISMS AND (CO)PRODUCTS. Given categories **A**, **B** and **C**, and a functor $\xi : \mathbf{A} \to \mathbf{B}$, define the pullback $\xi^* : [\mathbf{B}, \mathbf{C}] \to [\mathbf{A}, \mathbf{C}]$ by $\mathscr{F} \mapsto \mathscr{F} \circ \phi$. It commutes with limits and colimits [Artin et al., 1972a, Sec. I.5].

2.2.1. DEFINITION. *A morphism $\xi : \mathbf{S}_1 \to \mathbf{S}_2$ of conditional meet semilattices is a functor (i.e. a monotone map) with the following property: if $X \wedge Y$ exists in $\mathbf{S}_1$, then $\xi(X \wedge Y) = \xi(X) \wedge \xi(Y)$.*

A morphism $\phi : (\mathbf{S}, \mathscr{M}) \to (\mathbf{S}', \mathscr{M}')$ *between information structures is a pair* $\phi = (\phi_0, \widehat{\phi})$ *such that* ϕ_0 *is a morphism of conditional meet semilattices that maps* $\top_{\mathbf{S}}$ *to* $\top_{\mathbf{S}'}$, *and* $\widehat{\phi} : \mathscr{M} \Rightarrow \phi_0^* M'$ *is a natural transformation. If there is no risk of ambiguity, we write* ϕ *instead of* ϕ_0.

Given $\phi : (\mathbf{S}, \mathscr{M}) \to (\mathbf{S}', \mathscr{M}')$ *and* $\psi : (\mathbf{S}', \mathscr{M}') \to (\mathbf{S}'', \mathscr{M}'')$, *their composition* $\psi \circ \phi$ *is defined as* $(\psi_0 \circ \phi_0, \widehat{\psi} \circ \widehat{\phi} : \mathscr{M} \Rightarrow \phi_0^* \psi_0^* M'')$.

We denote by **InfoStr** *the category of information structures obtained in this way.*

Note that, if $X \wedge Y$ exists, then $\xi(X \wedge Y) \to \xi(X)$ and $\xi(X \wedge Y) \to \xi(X)$, and thus the product $\xi(X) \wedge \xi(Y)$ exists too, in virtue of Definition 1.2.1.

A morphism of information structures is a particular case of morphism of $\mathbf{Meas}_{\mathrm{surj}}$-valued covariant diagrams [De Silva and Barbosa, 2019, Def. 3.1]. We want ϕ_0 to respect the unit and the products, so that it induces a morphism between the corresponding presheaves of idempotent monoids, see Proposition 3.2.3.

The preceding definition is one of the main motivations for our generalized setting. In fact, one could imagine a correspondence between the partitions of two concrete structures (Example 2.1.1) defined on different sample spaces, but in which category would that correspondence take place? Since we eliminated the explicit reference to the sample space in our definition of information structure, the introduction of morphisms becomes straightforward. This allows the computation of products and coproducts. The connection to the sample spaces is not completely lost, but reformulated in the language of representations: as a consequence of Proposition 2.3.9, if $\mathbf{S}_i$ is a concrete structure on Ω_i $(i = 1, 2)$, then the objects of $\mathbf{S}_1 \times \mathbf{S}_2$ can be identified with partitions of $\Omega_1 \times \Omega_2$, as one would expect.

2.2.2. PROPOSITION. *The category* **InfoStr** *has countable products and arbitrary coproducts.*

PROOF. Let $\mathbf{0}$ be the category that has $\top$ as the only object and $\mathrm{id}_\top$ as the only morphism, and let $\mathscr{M}_{\mathbf{0}}$ be the functor that associates to $\top$ the set $\{*\}$ equipped with the atomic σ-algebra. Clearly $(\mathbf{0}, \mathscr{M}_{\mathbf{0}})$ is initial and terminal in the category **InfoStr**, hence it corresponds to the empty product and coproduct respectively.

Nonempty products: Given information structures $(\mathbf{S}_i, \mathscr{M}_i)$ indexed by i in an arbitrary set I, we introduce first the ordinary categorical product $\mathbf{S} = \prod_{i \in I} \mathbf{S}_i$: its objects are I-tuples $\langle X_i \rangle_{i \in I}$ with $X_i \in \mathrm{Ob}\,\mathbf{S_i}$ for each $I \in I$; there is an arrow $\langle \pi_i \rangle_{i \in I} : \langle X_i \rangle_{i \in I} \to \langle Y_i \rangle_{i \in I}$ whenever $\pi_i : X_i \to Y_i$ in $\mathbf{S}_i$ for each $i \in I$. Then a functor $\mathscr{M} : \mathbf{S} \to \mathbf{Meas}_{\mathrm{surj}}$ is defined as follows: for each $X = \langle X_i \rangle_{i \in I} \in \mathrm{Ob}\,\mathbf{S}$ the measurable space $\mathscr{M}(X)$ is the set $\mathscr{E}(X) := \prod_{i \in I} \mathscr{E}_i(X_i)$ equipped with the product σ-algebra $\mathfrak{B}(X) := \bigotimes_{i \in I} \mathfrak{B}_i(X_i)$, which is the smallest σ-algebra that makes every canonical projection $\widehat{p^i}_{\langle X_i \rangle_{i \in I}} : \mathscr{E}(\langle X_i \rangle_{i \in I}) \to \mathscr{E}_i(X_i)$ measurable [Cohn, 2013, Sec. 5.1]; at the level of morphisms, $\mathscr{M}(\langle \pi_i \rangle_{i \in I}) := \prod_{i \in I} \mathscr{M}_i(\pi_i)$, which comes from the product in **Sets**.

The pair $(\mathbf{S}, \mathscr{M})$ is an information structure. It is easy to verify that $\mathbf{S}$ is a poset with terminal object $\langle \top_{\mathbf{S}_i} \rangle_{i \in I}$. The conditional existence of products also holds: if $\langle X_i \rangle_{i \in I}$, $\langle Y_i \rangle_{i \in I}$ and $\langle Z_i \rangle_{i \in I}$ are objects of $\mathbf{S}$ such that $\langle X_i \rangle_{i \in I} \to \langle Y_i \rangle_{i \in I}$ and $\langle X_i \rangle_{i \in I} \to \langle Z_i \rangle_{i \in I}$,

then for every $i \in I$, $Y_i \xleftarrow{\pi_{Y_i}} X_i \xrightarrow{\pi_{Z_i}} Z_i$ in $\mathbf{S}_i$, which in turn implies that $Y_i \wedge Z_i$ exists in $\mathbf{S}_i$ by definition of conditional meet semilattice; the reader can verify that

$$\langle Y_i \rangle_{i \in I} \wedge \langle Z_i \rangle_{i \in I} = \langle Y_i \wedge Z_i \rangle_{i \in I}.$$

The functor $\mathscr{M}$ also has the desired properties. It is clear that $\mathscr{E}(\langle \top_{\mathbf{S}_i} \rangle_{i \in I}) \cong \{*\}$. If I is countable, then for any $(x_i)_{i \in I} \in \mathscr{E}(\langle X_i \rangle_{i \in I})$ the singleton $\{(x_i)_{i \in I}\}$ belongs to $\mathfrak{B}(\langle X_i \rangle_{i \in I})$, because it can be written as a countable intersection $\bigcap_{i \in I} (\widehat{p^i}_{\langle X_i \rangle_{i \in I}})^{-1}(x_i)$. Finally, when $\mathscr{M}$ is applied to the product $\langle Y_i \rangle_{i \in I} \wedge \langle Z_i \rangle_{i \in I}$ and its projections, one gets

$$\mathscr{M} \langle Y_i \rangle_{i \in I} \xleftarrow{\mathscr{M} \langle \pi_{Y_i} \rangle_{i \in I}} \mathscr{M} \langle Y_i \wedge Z_i \rangle_{i \in I} \xrightarrow{\mathscr{M} \langle \pi_{Z_i} \rangle_{i \in I}} \mathscr{M} \langle Z_i \rangle_{i \in I}.$$

The map

$$\mathscr{M} \langle \pi_{Y_i} \rangle_{i \in I} \times \mathscr{M} \langle \pi_{Z_i} \rangle_{i \in I} : \mathscr{M} \langle Y_i \wedge Z_i \rangle_{i \in I} \to \mathscr{M} \langle Y_i \rangle_{i \in I} \times \mathscr{M} \langle Z_i \rangle_{i \in I}$$

is injective, because for any $(y_i)_{i \in I} \in \mathscr{E}(\langle Y_i \rangle_{i \in I})$ and $(z_i)_{i \in I} \in \mathscr{E}(\langle Z_i \rangle_{i \in I})$, the elementary properties of set operations imply that

$$\begin{aligned}
(\mathscr{M} \langle \pi_{Y_i} \rangle_{i \in I} &\times \mathscr{M} \langle \pi_{Z_i} \rangle_{i \in I})^{-1}((y_i)_{i \in I}, (z_i)_{i \in I}) \\
&= (\mathscr{M} \langle \pi_{Y_i} \rangle_{i \in I})^{-1}((y_i)_{i \in I}) \cap (\mathscr{M} \langle \pi_{Z_i} \rangle_{i \in I})^{-1}((z_i)_{i \in I}) \\
&= \left\{ \prod_{i \in I} \mathscr{M}_i \pi_{Y_i}^{-1}(y_i) \right\} \cap \left\{ \prod_{i \in I} \mathscr{M}_i \pi_{Z_i}^{-1}(z_i) \right\} \qquad \text{(by def. of } \mathscr{M}) \\
&= \prod_{i \in I} \left\{ \mathscr{M}_i \pi_{Y_i}^{-1}(y_i) \cap \mathscr{M}_i \pi_{Z_i}^{-1}(z_i) \right\},
\end{aligned}$$

and the cardinality of each factor in the last expression is at most 1.

For each $i \in I$, we introduce a morphism of information structures $p^i : (\mathbf{S}, \mathscr{M}) \to (\mathbf{S}_i, \mathscr{M}_i)$ such that $p^i{}_0$ maps each object or morphism $\langle A_i \rangle_{i \in I}$ to A_i, and

$$\widehat{p^i}_{\langle X_i \rangle_{i \in I}} : \prod_{i \in I} \mathscr{M}_i(X_i) \to \mathscr{M}_i(X_i)$$

is the canonical projection (which is measurable by definition of the product σ-algebra, see above). We claim that $\mathbf{S}$, with the projections p^i just introduced, is the product of $(\mathbf{S}_i, \mathscr{M}_i)_{i \in I}$ in $\mathbf{InfoStr}$, written $\prod_{i \in I}(\mathbf{S}_i, \mathscr{M}_i)$, unique up to unique isomorphism (we also use the symbol $\times$ for finite products). In fact, given an I-cone $\{f^i : (\mathbf{R}, \mathscr{F}) \to (\mathbf{S}_i, \mathscr{M}_i)\}_{i \in I}$ in $\mathbf{InfoStr}$ (where I is seen as a discrete category), define $\langle f^i \rangle_{i \in I} : (\mathbf{R}, \mathscr{F}) \to (\mathbf{S}, \mathscr{M})$ by

$$\begin{aligned}
(\langle f^i \rangle_{i \in I})_0 : &\mathbf{R} \to \mathbf{S} \\
&R \mapsto \langle f^i(R) \rangle_{i \in I}
\end{aligned}$$

for any object or morphism R; for any $X \in \mathrm{Ob}\,\mathbf{R}$, the surjection

$$\widehat{\langle f^i \rangle_{i \in I}}(X) : \mathscr{F}(X) \to \mathscr{M}(\langle f^i(X) \rangle_{i \in I}) = \prod_{i \in I} \mathscr{M}_i(f^i(X))$$

is the map $\langle \widehat{f^i}_X \rangle_{i \in I}$ induced by the I-cone $\{\widehat{f^i}_X : \mathscr{F}(X) \to \mathscr{M}_i(f^i(X))\}_{i \in I}$ in **Sets**, in such a way that $p^i \circ \langle f^i \rangle_{i \in I} = f^i$ for all $i \in I$.

Nonempty coproducts: Given information structures $\{(\mathbf{S}_i, \mathscr{M}_i)\}_{i \in I}$, define a category $\mathbf{S}$ such that

- $\mathrm{Ob}\,\mathbf{S} = \bigsqcup_{i \in I} \mathrm{Ob}\,\mathbf{S_i}/\sim$, where $\sim$ is the smallest equivalence relation such that $\top_{\mathbf{S}_i} \sim \top_{\mathbf{S}_j}$ for all $i, j \in I$;

- $A \to B$ in $\mathbf{S}$ if and only if $A \to B$ in $\mathbf{S}_i$ for some i.

Let $\mathscr{M} : \mathbf{S} \to \mathbf{Sets}$ be the functor that coincides with $\mathscr{M}_i$ on $\mathbf{S}_i$. The pair $(\mathbf{S}, \mathscr{M})$ is an information structure: the properties in Definition 1.2.1 are verified locally on each $\mathbf{S}_i$.

Injections $j^i : \mathbf{S}_i \to \mathbf{S}$ are defined in the obvious way: $j_0^i(A) = A$ for $A \in \mathrm{Ob}\,\mathbf{S_i}$ or $A \in \mathrm{Hom}(\mathbf{S}_i)$, and the mappings $\widehat{j^i_X}$ are identities. If $\{f^i : (\mathbf{S}_i, \mathscr{M}_i) \to (\mathbf{R}, \mathscr{F})\}_{i \in I}$ is an I-cocone, define

$$(\langle f^i \rangle_{i \in I})_0 : \mathbf{S} \to \mathbf{R}$$
$$A \mapsto f^i(A) \text{ if } A \in \mathrm{Ob}\,\mathbf{S_i} \text{ or } A \in \mathrm{Hom}(\mathbf{S}_i)$$

and, if $X \in \mathrm{Ob}\,\mathbf{S_i}$, set $(\widehat{\langle f^i \rangle_{i \in I}})_X = \widehat{f^i}_X$. By construction, $\langle f^i \rangle_{i \in I} \circ j^i = f^i$. Therefore, $(\mathbf{S}, \mathscr{M})$ is the coproduct of $\{(\mathbf{S}_i, \mathscr{M}_i)\}_{i \in I}$ in **InfoStr**, denoted $\coprod_{i \in I}(\mathbf{S}_i, \mathscr{M}_i)$ (we also use $\sqcup$ for finite coproducts), which is unique up to unique isomorphism. ∎

2.2.3. REMARK. If $(\mathbf{S}_1, \mathscr{M}_1)$ and $(\mathbf{S}_2, \mathscr{M}_2)$ are bounded structures, their product and coproduct are bounded too. In fact, if the height of the poset $\mathbf{S}_i$ is N_i ($i = 1, 2$), then the height of $\mathbf{S}_1 \times \mathbf{S}_2$ is $N_1 + N_2$ and that of $\mathbf{S}_1 \sqcup \mathbf{S}_2$ equals $\max(N_1, N_2)$. Similarly, if both structures are finite, their product and coproduct is finite too.

2.2.4. REMARK. If each measurable space $(\mathscr{E}(X), \mathfrak{B}(X))$ appearing in $\mathbf{S}_1$ and $\mathbf{S}_2$ verifies that $\mathscr{E}(X)$ is second countable topological space and $\mathfrak{B}(X)$ is its Borel σ-algebra, then each algebra $\mathfrak{B}(X_1) \otimes \mathfrak{B}(X_2)$ on $\mathscr{E}(X_1) \times \mathscr{E}(X_2)$ equals the Borel σ-algebra on this space [Vigneaux, 2019, Prop. 9.1].

One could remove the requirement of surjectivity in Definition 1.2.1: products, coproducts and even other (co)limits would make sense. However, the probabilistic interpretation of such structures would be different: they do not have classical or quantum representations in the sense of Definitions 2.3.1 and 2.3.12, nor the cohomological results in Section 4 apply to them (because a product $X \wedge Y$ is always degenerate if one of the "coordinate projections" $\mathscr{E}_{X \wedge Y} \to \mathscr{E}_X$ or $\mathscr{E}_{X \wedge Y} \to \mathscr{E}_Y$ is not surjective). The definition of information cohomology would still hold, since it only depends on the underlying conditional meet semilattice.

2.3. DIGRESSION: REPRESENTATIONS. We introduce here the notion of *representation* of an information structure in terms of classical observables (measurable functions) or quantum observables (self-adjoint operators), as a bridge between our categorical definitions and the more traditional models used in classical and quantum probability theory. The rest of the paper does not depend on this section. For simplicity, we restrict ourselves to finite information structures.

Recall that $\mathbf{Obs}_{\text{fin}}(\Omega)$ denotes the poset of finite partitions of a set Ω, ordered by the relation of refinement, and $\square$ is the "forgetful" functor from $\mathbf{Obs}_{\text{fin}}(\Omega)$ into $\mathbf{Sets}$ introduced in Example 2.1.1, that maps each partition $\{A_1, ..., A_n\}$ to the set $\{A_1, ..., A_n\}$ and each arrow of refinement to a surjection.

2.3.1. DEFINITION. *A* classical representation *of a finite information structure* $(\mathbf{S}, \mathscr{E})$ *is a pair* (Ω, ρ)*, where* Ω *is a set and* $\rho = (\rho_0, \widehat{\rho}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{Obs}_{\text{fin}}(\Omega), \square)$ *is a morphism of information structures such that* $\widehat{\rho} : \mathscr{E} \to \rho_0^* \square$ *is a natural isomorphism (i.e. the components* $\widehat{\rho}_X : \mathscr{E}(X) \to \square \rho_0(X)$ *are bijections, natural in* X*).*

If (Ω, ρ) is a classical representation of $\mathbf{S}$, each observable X in $\mathbf{S}$ can be associated with a unique function $\widetilde{X} : \Omega \to \mathscr{E}_X$, in such a way that $\rho_0(X)$ is the partition induced by $\widetilde{X}$ and $\widehat{\rho}_X(x) = \widetilde{X}^{-1}(x)$. Since ρ_0 is a morphism of conditional meet semilattices, for any $X, Y \in \mathrm{Ob}\,\mathbf{S}$ the joint $(\widetilde{X}, \widetilde{Y}) : \Omega \to \mathscr{E}_X \times \mathscr{E}_Y$ is equivalent to $\widetilde{X \wedge Y} : \Omega \to \mathscr{E}_{XY}$, in the sense that both induce the same partition of Ω.

The next proposition points to a close link between representations and $\lim \mathscr{E}$. Recall that the limit of the functor $\mathscr{E} : \mathbf{S} \to \mathbf{Sets}$ is defined as

$$\lim \mathscr{E} := \mathrm{Hom}_{[\mathbf{S}, \mathbf{Sets}]}(*, \mathscr{E}), \tag{22}$$

where $[\mathbf{S}, \mathbf{Sets}]$ is the category of covariant functors from $\mathbf{S}$ to $\mathbf{Sets}$, and $*$ is the functor that associates to each object a one-point set; equivalently

$$\lim \mathscr{E} \cong \left\{ (s_Z)_{Z \in \mathrm{Ob}\,\mathbf{S}} \in \prod_{Z \in \mathrm{Ob}\,\mathbf{S}} \mathscr{E}(Z) \,:\, \mathscr{E}\pi_{YX}(s_X) = s_Y \text{ for all } \pi_{YX} : X \to Y \right\}, \tag{23}$$

where s_Z denotes $\varphi(*)$ for any $\varphi \in \mathrm{Hom}_{[\mathbf{S}, \mathbf{Sets}]}(*, \mathscr{E})$. The requirements imposed on $(s_Z)_{Z \in \mathrm{Ob}\,\mathbf{S}}$ in (23) are referred hereafter as *compatibility conditions*. We denote the restriction of each projection $\pi_{\mathscr{E}(X)} : \prod_{Z \in \mathrm{Ob}\,\mathbf{S}} \mathscr{E}(Z) \to \mathscr{E}(X)$ to $\lim \mathscr{E}$ by the same symbol. We interpret the limit as all possible combinations of compatible outcomes.

2.3.2. PROPOSITION. *If* $(\mathbf{S}, \mathscr{E})$ *has a classical representation* $(\Omega, \rho, \widehat{\rho})$*, then for any* $X \in \mathrm{Ob}\,\mathbf{S}$ *and any* $x \in \mathscr{E}(X)$*, there exists an element* $s(x) \in \lim \mathscr{E}$ *such that* $\pi_{\mathscr{E}(X)}(s(x)) = x$.

PROOF. For each X in $\mathrm{Ob}\,\mathbf{S}$, there is a surjection $\pi_X : \Omega \to \mathscr{E}(X)$ obtained as the composition of $\xi_X : \Omega \to \square \rho_0(X)$, which maps $\omega \in \Omega$ to the part that contains it, and $\widehat{\rho}_X^{-1} : \square \rho_0(X) \to \mathscr{E}(X)$. The maps $\{\pi_X\}$ define a cone over $\mathscr{E}$ i.e. given $f : Y \to Z$ in $\mathbf{S}$,

one has a commutative diagram

$$\begin{array}{ccccc}
\Omega & \xrightarrow{\xi_Y} & \Box\rho_0(Y) & \xrightarrow{\widehat{\rho_Y}^{-1}} & \mathscr{E}(Y) \\
\Big\| & & \Big\downarrow{\scriptstyle\Box\rho_0 f} & & \Big\downarrow{\scriptstyle\mathscr{E}f} \\
\Omega & \xrightarrow{\xi_Z} & \Box\rho_0(Z) & \xrightarrow{\widehat{\rho_Z}^{-1}} & \mathscr{E}(Z)
\end{array} \;;$$

the commutation of the left square comes from the definition of $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ and $\Box$, and the right square commutes because $\widehat{\rho}$ is a natural isomorphism. Therefore, there is a map $\pi : \Omega \to \lim\mathscr{E}$ and the desired section is obtained as the image under π of any $\omega \in \widehat{\rho}_X(x)$. $\blacksquare$

2.3.3. DEFINITION. *An information structure is* noncontextual *if, for all $X \in \mathrm{Ob}\,\mathbf{S}$ and all $x \in \mathscr{E}(X)$, there exists an element $s(x) \in \lim\mathscr{E}$ such that $\pi_{\mathscr{E}(X)}(s(x)) = x$.*

Thus information structures are sufficiently flexible to model *contextual* situations, which arise when data is locally consistent but globally inconsistent. This happens in different domains, notably in quantum mechanics and in database theory. In the terminology of [Abramsky et al., 2015, Sec. 3], a structure is said to be *logically contextual at a value $x \in \mathscr{E}_X$* if x belongs to no compatible family of measurements, i.e. there is no section $s(x) \in \lim\mathscr{E}$ such that $\pi_{\mathscr{E}(X)}(s(x)) = x$, and *strongly contextual* if $\mathscr{E}$ does not accept any global section, i.e. $\lim\mathscr{E} = \emptyset$.

2.3.4. THEOREM. *A finite information structure $(\mathbf{S}, \mathscr{E})$ is noncontextual if and only if it has a classical representation.*

PROOF. Proposition 2.3.2 already showed that existence of a classical representation implies noncontextuality.

In the other direction, let us suppose that a finite structure $(\mathbf{S}, \mathscr{E})$ is noncontextual. We are going to show that it can be represented by partitions of $\lim\mathscr{E}$. Define first $\tau_0 : \mathbf{S} \to \mathbf{Obs}_{\mathrm{fin}}(\lim\mathscr{E})$ as follows: associate to $X \in \mathrm{Ob}\,\mathbf{S}$ the set $\tau_0(X) := \{\pi_{\mathscr{E}(X)}^{-1}(x)\}_{x\in\mathscr{E}(X)}$, which is clearly a partition of $\lim\mathscr{E}$; the associated map $\widehat{\tau}_X : \mathscr{E}_X \to \tau(X)$, $x \mapsto \pi_{\mathscr{E}(X)}^{-1}(x)$ is a bijection, because noncontextuality means that each $\pi_{\mathscr{E}(X)}^{-1}(x)$ is nonempty. Given $\pi_{YX} : X \to Y$, there is a corresponding arrow $\tau_0(X) \to \tau_0(Y)$ in $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$ in virtue of the identity

$$\pi_{\mathscr{E}(Y)}^{-1}(y) = \bigcup_{x\in\mathscr{E}\pi_{YX}^{-1}(y)} \pi_{\mathscr{E}(X)}^{-1}(x), \tag{24}$$

which is proved as follows: if $x \in \mathscr{E}\pi_{YX}^{-1}(y)$ and $s \in \pi_{\mathscr{E}(X)}^{-1}(x)$, then

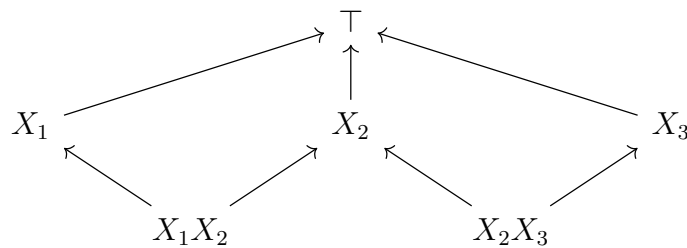$$\pi_{\mathscr{E}(Y)}(s) = \mathscr{E}\pi_{YX}(\pi_{\mathscr{E}(X)}(s)) = \mathscr{E}\pi_{YX}(x) = y,$$

which means $\cup_{x\in\mathscr{E}\pi_{YX}^{-1}(y)}\pi_{\mathscr{E}(X)}^{-1}(x) \subset \pi_{\mathscr{E}(Y)}^{-1}(y)$; to prove the other inclusion, take $s = (s_Z)_{Z\in\mathrm{Ob}\,\mathbf{S}} \in \pi_{\mathscr{E}(Y)}^{-1}(y)$ and note that s_X must satisfy–by definition–the compatibility condition $\mathscr{E}\pi_{YX}(s_X) = s_Y = y$, thus $s_X \in \mathscr{E}\pi_{YX}^{-1}(y)$ and s itself belong to $\cup_{x\in\mathscr{E}\pi_{YX}^{-1}(y)}\pi_{\mathscr{E}(X)}^{-1}(x)$.

Finally, to prove that τ_0 defines a morphism of conditional meet semilattices, consider a diagram $X \leftarrow X \wedge Y \rightarrow Y$, and an arbitrary partition W of $\lim \mathscr{E}$ that refines $\tau_0(X)$ and $\tau_0(Y)$. We have to show that W also refines $\tau_0(X \wedge Y)$. If W refines $\tau_0(X)$, each $w \in W$ (w is a subset of $\lim \mathscr{E}$) is mapped to certain x_w by $\pi_{\mathscr{E}(X)}$; analogously, $\pi_{\mathscr{E}(Y)}(w) = \{y_w\}$. This means that $\pi_{\mathscr{E}(X \wedge Y)}(w) = \{z_w\}$, where z_w is the only point of $\mathscr{E}(X \wedge Y)$ that satisfies $\mathscr{E}(\pi_{X(X \wedge Y)})(z_w) = x_w$, $\mathscr{E}(\pi_{X(X \wedge Y)})(z_w) = y_w$, which means that $w \subset \pi_{\mathscr{E}(X \wedge Y)}^{-1}(z_w)$. Thus, W refines $\tau_0(X \wedge Y)$. ∎

2.3.5. EXAMPLE. In the notation of Example 1.2.2, let $\mathbf{K}$ be simplicial subcomplex of $\Delta(\{1, 2, 3\})$ with minimal objects $\{1, 2\}$, $\{1, 3\}$, and $\{2, 3\}$, which can be pictured as a triangle. We consider three possible functors $\mathscr{E}$. All of them associate to $\{1\}$, $\{2\}$, and $\{3\}$ the set $\{0, 1\}$. We give $\mathscr{E}_{\{i,j\}}$ as a subset of $\mathscr{E}_{\{i\}} \times \mathscr{E}_{\{j\}}$ so that the maps $\mathscr{E}_{\{i,j\}} \rightarrow \mathscr{E}_{\{i\}}$ are the restrictions of the canonical projectors.

1. If $\mathscr{E}_{\{i,j\}} = \{(0, 1), (1, 0)\}$ (for any $i, j \in \{1, 2, 3\}$ such that $i \neq j$), then $\lim \mathscr{E} = \emptyset$ i.e. the structure is strongly contextual, and it has no classical representation. This means that there is no assignment of values for the three observables associated to the vertices of the triangle that is consistent with the constraints associated to the edges.

2. If $\mathscr{E}_{\{i,j\}} = \{(0, 0), (1, 1)\}$ (for any $i, j \in \{1, 2, 3\}$ such that $i \neq j$), them $\lim \mathscr{E} \cong \{0, 1\}$. The observables associated to the vertices must take all the same value.

3. If only $\mathscr{E}_{\{1,2\}} = \{(0, 0), (1, 1)\}$, and the other two $\mathscr{E}_{\{i,j\}}$ equal $\mathscr{E}_{\{i\}} \times \mathscr{E}_{\{j\}}$, then $\lim \mathscr{E}$ has just six different elements. This corresponds to imposing a constraint that makes the observables associated to $\{1\}$ and $\{2\}$ identical.

2.3.6. EXAMPLE. Given a concrete structure $(\mathbf{S}, \square)$ on a set Ω (see Example 2.1.1), the set $\lim \square$ may differ from the original Ω, as the following example shows. Set $\Omega = \{1, 2, 3, 4\}$ and $X_i = \{\{i\}, \Omega \setminus \{i\}\}$, for $i = 1, ..., 4$. Let $\mathbf{S}$ be the concrete structure that includes only the partitions X_1, X_2, X_3, $X_1 X_2$, and $X_2 X_3$. The corresponding generalized information structure is given by a conditional meet semilattice represented by the graph

and the functor $\square$ can pictured as

$$\{\{1,2,3,4\}\}$$

$$\{\{1\},\{2,3,4\}\} \qquad \{\{2\},\{1,3,4\}\} \qquad \{\{3\},\{1,2,4\}\}$$
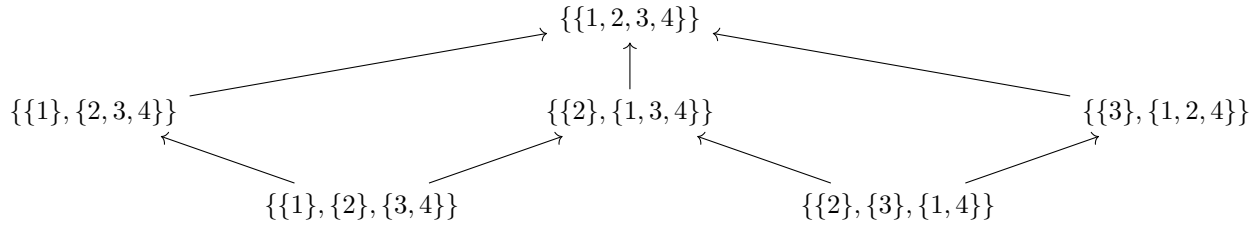
$$\{\{1\},\{2\},\{3,4\}\} \qquad \{\{2\},\{3\},\{1,4\}\}$$

where each arrow corresponds to a surjection of finite sets that sends I to J when $I \subset J$. In this case, $\lim \mathscr{E} \subset \{*\} \times \mathscr{E}(X_1) \times \mathscr{E}(X_2) \times \mathscr{E}(X_3) \times \mathscr{E}(X_1 X_2) \times \mathscr{E}(X_2 X_3)$ corresponds to the set

$$
\begin{aligned}
\lim \mathscr{E} = \{&(\{1,2,3,4\},\{1\},\{1,3,4\},\{3\},\{1\},\{3\}),(\{1,2,3,4\},\{1\},\{1,3,4\},\{1,2,4\},\{1\},\{1,4\}),\\
&(\{1,2,3,4\},\{2,3,4\},\{2\},\{1,2,4\},\{2\},\{2\}),(\{1,2,3,4\},\{2,3,4\},\{1,3,4\},\{3\},\{3,4\},\{3\}),\\
&(\{1,2,3,4\},\{2,3,4\},\{1,3,4\},\{1,2,4\},\{3,4\},\{1,4\})\}.
\end{aligned}
$$

The difference between Ω and $\lim \mathscr{E}$ is explained by the presence of

$$(\{1,2,3,4\},\{1\},\{1,3,4\},\{3\},\{1\},\{3\});$$

this measurement is impossible in the concrete structure $\mathbf{S} \subset \mathbf{Obs}_{\mathrm{fin}}(\Omega)$ (because the underlying sample $\omega \in \Omega$ should belong to $\{1\}$ and $\{3\}$), but the observables in $(\mathbf{S},\square)$ cannot distinguish between the points 1 and 3, a sort of nonseparability. In fact, if we also include $X_1 X_3$ at the beginning, we obtain $\Omega \cong \lim \mathscr{E}$.

2.3.7. PROPOSITION. *The product and the coproduct of two noncontextual structures is noncontextual.*

PROOF. Let $\mathbf{S}_1$ and $\mathbf{S}_2$ be noncontextual structures. We use the notations in the proof of Proposition 2.2.2.

Products: Consider a point $(x_1, x_2) \in \mathscr{E}(\langle X_1, X_2 \rangle)$. There exist sections

$$s^i(x_i) = (s^i_Z(x_i))_{Z \in \mathrm{Ob}\, \mathbf{S_i}} \in \lim \mathscr{E}_i \subset \prod_{Z \in \mathrm{Ob}\, \mathbf{S_i}} \mathscr{E}_i(Z),$$

such $\pi_{\mathscr{E}_i(X_i)}(s(x_i)) = x_i$ (for $i = 1, 2$). Note that the vector

$$s(x_1, x_2) := (s^1_{Z_1}(x_1), s^2_{Z_2}(x_2))_{\langle Z_1, Z_2 \rangle \in \mathrm{Ob}\, \mathbf{S}} \in \prod_{\langle Z_1, Z_2 \rangle \in \mathrm{Ob}\, \mathbf{S}} \mathscr{E}(\langle Z_1, Z_2 \rangle)$$

satisfies all the compatibility conditions and is therefore in $\lim \mathscr{E}$. By definition, $\pi_{\mathscr{E}(\langle X_1, X_2 \rangle)}(s(x_1, x_2)) = (x_1, x_2)$.

Coproducts: given $X \in \mathbf{S}_1$, $x \in \mathscr{E}(X)$, there exists $s^1(x) = (s_Z(x))_{Z \in \mathrm{Ob}\, \mathbf{S_1}} \in \lim \mathscr{E}_1$ satisfying $\pi_{\mathscr{E}(X)}(s^1(x)) = x$, and similarly for $\mathscr{E}_2$; we can build a new vector $(s_Z(x))_{Z \in \mathrm{Ob}\, \mathbf{S}} \in \mathrm{Ob}\, \mathbf{S_1} \sqcup \mathrm{Ob}\, \mathbf{S_2}$ such that $s_Z = s^i_Z$ if $Z \in \mathrm{Ob}\, \mathbf{S_i}$; luckily, for $\top$ there is no choice. ∎

2.3.8. REPRESENTATIONS OF FINITE PRODUCTS AND COPRODUCTS. There is a general construction of a representation for a product or coproduct of two information structures whenever each of the factors is already represented.

Let Ω_1 and Ω_2 be nonempty sets. Given collections $\mathfrak{A} = \{A_i\}_i$ of subsets of Ω_1 and $\mathfrak{B} = \{B_j\}_j$ of subsets of Ω_2, denote by $\mathfrak{A} \times \mathfrak{B}$ the collection $\{ A_i \times B_j : A_i \in \mathfrak{A} \text{ and } B_j \in \mathfrak{B} \}$ of subsets of $\Omega_1 \times \Omega_2$. If $\mathfrak{A}$ and $\mathfrak{B}$ are partitions, then $\mathfrak{A} \times \mathfrak{B}$ is a partition too.

Let (Ω_i, ρ^i), with $\rho^i = (\rho_i, \widehat{\rho^i})$, be a classical representation of $(\mathbf{S}_i, \mathscr{E}_i)$, for $i = 1, 2$. Associate to each observable $\langle X_1, X_2 \rangle \in \mathrm{Ob}\,\mathbf{S_1} \times \mathbf{S_2}$ the partition of $\Omega_1 \times \Omega_2$ given by

$$\rho^\times(\langle X_1, X_2 \rangle) := \rho_0^1(X_1) \times \rho_0^2(X_2). \tag{25}$$

There is a natural transformation $\widehat{\rho^\times}_{\langle X_1, X_2 \rangle} : \mathscr{E}(\langle X_1, X_2 \rangle) \to \Box\rho^\times(\langle X_1, X_2 \rangle)$ that maps (x_1, x_2) to $\widehat{\rho^1}_{X_1}(x_1) \times \widehat{\rho^2}_{X_2}(x_2)$.

Analogously, for each $X \neq \top$ in $\mathrm{Ob}\,\mathbf{S_1} \sqcup \mathbf{S_2}$, let us define the partition of $\Omega_1 \times \Omega_2$ given by

$$\rho^\sqcup(X) = \begin{cases} \rho^1(X) \times \{\Omega_2\} & \text{if } X \in \mathrm{Ob}\,\mathbf{S_1} \\ \{\Omega_1\} \times \rho^2(X) & \text{if } X \in \mathrm{Ob}\,\mathbf{S_2} \end{cases}. \tag{26}$$

In particular, $\rho^\sqcup(\top) = \{\Omega_1 \times \Omega_2\}$. The maps $\widehat{\rho^\sqcup}_X$ are $x \mapsto \widehat{\rho^1}(x) \times \{\Omega_2\}$ or $x \mapsto \{\Omega_1\} \times \widehat{\rho^2}(x)$ accordingly.

2.3.9. PROPOSITION. *Let (Ω_i, ρ^i) be a classical representation of $(\mathbf{S}_i, \mathscr{E}_i)$, for $i = 1, 2$. Then*

1. *$(\Omega_1 \times \Omega_2, \rho^\times)$ is a classical representation of $(\mathbf{S}_1, \mathscr{E}_1) \times (\mathbf{S}_2, \mathscr{E}_2)$;*

2. *$(\Omega_1 \times \Omega_2, \rho^\sqcup)$ is a classical representation of $(\mathbf{S}_1, \mathscr{E}_1) \sqcup (\mathbf{S}_2, \mathscr{E}_2)$.*

The proof depends on the following lemma.

2.3.10. LEMMA.

1. *If $\mathfrak{A} = \{A_i\}_i$ and $\mathfrak{A}' = \{A_j'\}_j$ are finite partitions of a set Ω, then $\sigma(\mathfrak{A}, \mathfrak{A}') = \sigma(\{A_i \cap A_j'\}_{i,j})$, and the nonempty elements of $\{A_i \cap A_j'\}_{i,j}$ are the atoms of $\sigma(\mathfrak{A}, \mathfrak{A}')$.*

2. *If $\mathfrak{A} = \{A_i\}_i$, $\mathfrak{A}' = \{A_j'\}_j$ are two finite partitions of Ω_1 and $\mathfrak{B} = \{B_l\}_l$, $\mathfrak{B}' = \{B_m\}_m$ two finite partitions of Ω_2, then $(\mathfrak{A} \times \mathfrak{B})(\mathfrak{A}' \times \mathfrak{B}') = \mathfrak{A}\mathfrak{A}' \times \mathfrak{B}\mathfrak{B}'$, where juxtaposition of partitions denotes their product in $\mathbf{Obs}_{\mathrm{fin}}(\Omega)$.*

PROOF.

1. On the one hand, note that each set $A_i \cap A_j'$ is contained in $\sigma(\mathfrak{A}, \mathfrak{A}')$, therefore $\sigma(\{A_i \cap A_j'\}_{i,j}) \subset \sigma(\mathfrak{A}, \mathfrak{A}')$. On the other, each generator $A_i \in \mathfrak{A}$ of $\sigma(\mathfrak{A}, \mathfrak{A}')$ can be written as

$$A_i = A_i \cap \Omega = A_i \cap \left(\bigcup_j A_j' \right) = \bigcup_j (A_i \cap A_j'),$$

and similarly for the generators $A_j' \in \mathfrak{A}'$, which implies that $\sigma(\mathfrak{A}, \mathfrak{A}') \subset \sigma(\{A_i \cap A_j'\}_{i,j})$. The reader can verify that the nonempty elements of $\{A_i \cap A_j'\}_{i,j}$ are atoms.

2. The previous result can be read as $\mathfrak{A}\mathfrak{A}' = \{A_i \cap A'_j\}_{i,j}$. The set-theoretical identity

$$(A_i \times B_l) \cap (A'_j \times B'_m) = (A_i \cap A'_j) \times (B_l \cap B'_m), \tag{27}$$

implies that the atoms of $(\mathfrak{A} \times \mathfrak{B})(\mathfrak{A}' \times \mathfrak{B}')$ and $\mathfrak{A}\mathfrak{A}' \times \mathfrak{B}\mathfrak{B}'$ coincide.

∎

PROOF OF PROPOSITION 2.3.9. Most verifications are almost immediate from the definitions. We simply prove that $\rho^\times(\langle X_1, X_2 \rangle \wedge \langle Y_1, Y_2 \rangle) = \rho^\times(\langle X_1, X_2 \rangle)\rho^\times(\langle Y_1, Y_2 \rangle)$. Note that

$$\begin{aligned}
\rho^\times(\langle X_1, X_2 \rangle \wedge \langle Y_1, Y_2 \rangle) &= \rho^\times(\langle X_1 \wedge Y_1, X_2 \wedge Y_2 \rangle) \\
&= \rho^1(X_1 \wedge Y_1) \times \rho^2(X_2 \wedge Y_2) \\
&= \rho^1(X_1)\rho^1(Y_1) \times \rho^2(X_2)\rho^2(Y_2) \\
&= (\rho^1(X_1) \times \rho^2(X_2))(\rho^1(Y_1) \times \rho^2(Y_2)) \\
&= \rho^\times(\langle X_1, X_2 \rangle)\rho^\times(\langle Y_1, Y_2 \rangle)
\end{aligned}$$

The first equality comes from the construction of $\mathbf{S}_1 \times \mathbf{S}_2$; the second, from the definition of $\rho^\times$; the third, from the fact that $\rho^1 = \rho^1_0$ and $\rho^2 = \rho^2_0$ are morphisms of conditional meet semilattices; the fourth equality is just a consequence of Lemma 2.3.10, and the fifth is a rewriting of the previous one. ∎

The partitions of $\Omega_1 \times \Omega_2$ in the image of $\rho^\sqcup$ are also in the image of $\rho^\times$. This is consistent with the existence of a canonical morphism of structures $\phi : (\mathbf{S}_1, \mathscr{E}_1) \sqcup (\mathbf{S}_2, \mathscr{E}_2) \to (\mathbf{S}_1, \mathscr{E}_1) \times (\mathbf{S}_2, \mathscr{E}_2)$, with ϕ_0 given at the level of objects by the injection

$$X \mapsto \begin{cases} \top_{\mathbf{S}_1 \times \mathbf{S}_2} & \text{if } X = \top_{\mathbf{S}_1 \sqcup \mathbf{S}_2} \\ \langle X, \top_{\mathbf{S}_2} \rangle & \text{if } X \in \mathrm{Ob}\,\mathbf{S}_1 \\ \langle \top_{\mathbf{S}_1}, X \rangle & \text{if } X \in \mathrm{Ob}\,\mathbf{S}_2 \end{cases}, \tag{28}$$

and the corresponding components $\widehat{\phi}_X$ being the obvious bijections: $\mathscr{E}_1(X) \to \mathscr{E}_1(X) \times \{*\}$ when $X \in \mathrm{Ob}\,\mathbf{S}_1$ or $\mathscr{E}_2(X) \to \{*\} \times \mathscr{E}_2(X)$ when $X \in \mathrm{Ob}\,\mathbf{S}_2$. The representation $(\Omega_1 \times \Omega_2, \rho^\times)$ on $(\mathbf{S}_1, \mathscr{E}_1) \times (\mathbf{S}_2, \mathscr{E}_2)$ may be pulled back under ϕ, and the resulting representation $(\Omega_1 \times \Omega_2, \rho^\times \circ \phi)$ on $(\mathbf{S}_1, \mathscr{E}_1) \sqcup (\mathbf{S}_2, \mathscr{E}_2)$ coincides with $(\Omega_1 \times \Omega_2, \rho^\sqcup)$. This is a particular case of a more general procedure to pull back representations, valid for any morphism of structures $\phi = (\phi_0, \widehat{\phi})$ such that $\widehat{\phi}$ is a natural isomorphism.

2.3.11. QUANTUM OBSERVABLES. For the sake of completeness, we also indicate how an information structure may be represented by noncommutative observables adapted to quantum mechanics.

Let V be a finite dimensional Hilbert space: a complex vector space with a positive definite hermitian form $\langle \cdot, \cdot \rangle$. In the quantum setting, random variables are generalized by endomorphisms of V (operators). An operator H is called hermitian if for all $u, v \in V$,

one has $\langle u, Hv \rangle = \langle Hu, v \rangle$. A quantum observable is a hermitian operator: the result of a quantum experiment is supposed to be an eigenvalue of such operator, that is always a real number.

A fundamental result of linear algebra, the Spectral Theorem [Halmos, 1958, Sec. 79], says that each hermitian operator Z can be decomposed as a weighted sum of positive hermitian projectors $Z = \sum_{j=1}^{K} z_j V_j$ where $z_1, ..., z_K$ are the (pairwise distinct) *real* eigenvalues of Z. Each V_j is the projector on the eigenspace spanned by the eigenvectors of z_j; the dimension of this subspace equals the multiplicity of z_j as eigenvalue. As hermitian projectors, they satisfy the equation $V_j^2 = V_j$ and $V_j^* = V_j$. They are also mutually orthogonal ($V_j V_k = 0$ for integers j, k), and their sum equals the identity, $\sum_{1 \le j \le K} V_j = \mathrm{id}_V$.

In analogy to the classical case, we consider equivalent two hermitian operators that define the same direct sum decomposition $\{V_j\}_j$ of V by means of the Spectral Theorem, ignoring the particular eigenvalues. For us, observable and direct sum decomposition are then interchangeable terms. In what follows, we denote by V_α both the subspace of V and the orthogonal projector on it. A decomposition $\{V_\alpha\}_{\alpha \in A}$ is said to refine $\{V'_\beta\}_{\beta \in B}$ if each V'_β can be expressed as sum of subspaces $\{V_\alpha\}_{\alpha \in A_\beta}$, for certain $A_\beta \subseteq A$. In that case we say also that $\{V_\alpha\}_{\alpha \in A}$ divides $\{V'_\beta\}_{\beta \in B}$, and we write $\{V_\alpha\}_{\alpha \in A} \to \{V'_\beta\}_{\beta \in B}$. With this arrows, direct sum decompositions form a category denoted $\mathbf{DSD}(V)$; the opposite category is the poset of quantum measurement contexts studied in [Constantin and Döring, 2012]. The reader may verify that it is a conditional meet semilattice with terminal object $\{V\}$, but not a meet semilattice. We also introduce in this case a functor $\square$ that maps each decomposition $\{V_\alpha\}_{\alpha \in A}$ to the set $\{V_\alpha\}_{\alpha \in A}$, and each arrow of refinement to a surjection.

2.3.12. DEFINITION. *A* quantum representation *of a finite information structure* $(\mathbf{S}, \mathscr{E})$ *is a pair* (V, ρ)*, where* V *is a Hilbert space and* $\rho = (\rho_0, \widehat{\rho}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{DSD}(V), \square)$ *is a morphism of information structures such that* $\widehat{\rho}$ *is a natural isomorphism (i.e. for each* $X \in \mathrm{Ob}\,\mathbf{S}$*, the component* $\widehat{\rho}_X : \mathscr{E}(X) \to \square\rho(X)$ *is a bijection).*

The cohomological computations in Section 4 concern classical probabilities, but the general constructions in Section 3 only depend on the conditional meet semilattice with terminal object, hence they are equally valid in the quantum case. Using those results, one may recover the explicit cochain complex that defined quantum information cohomology in [Baudot and Bennequin, 2015].

3. Information cohomology via derived functors

In this section, we define information cohomology as a derived functor, following a remark in [Baudot and Bennequin, 2015], and then reobtain the explicit cochain complex used there. The main new ingredient is Proposition 3.2.2, which proves that the relative bar resolution is a projective resolution. We suppose that the reader is familiar with abelian categories and derived functors: we use the definitions and notations in [Weibel, 1994, Ch. 1 & 2].

3.1. DEFINITION. Let $\mathbf{S}$ be a conditional meet semilattice with terminal object $\top$. We view it as a site with the trivial topology, such that every presheaf is a sheaf. For each $X \in \mathrm{Ob}\,\mathbf{S}$, set $\mathscr{S}_X := \{Y \in \mathrm{Ob}\,\mathbf{S} \mid X \to Y\}$, with the monoid structure given by the product of in $\mathbf{S}$: $(Z, Y) \mapsto ZY := Z \wedge Y$. Let $\mathscr{A}_X := \mathbb{R}[\mathscr{S}_X]$ be the corresponding monoid algebra. The contravariant functor $X \mapsto \mathscr{A}_X$ is a sheaf of rings; we denote it by $\mathscr{A}$. The pair $(\mathbf{S}, \mathscr{A})$ is a ringed site.

The category $\mathbf{Mod}(\mathscr{A})$ is abelian [Stacks Project Authors, 2018, Lemma 03DA] and has enough injective objects [Stacks Project Authors, 2018, Theorem 01DU]. For a fixed object $\mathscr{O}$ of $\mathbf{Mod}(\mathscr{A})$, the covariant functor $\mathrm{Hom}(\mathscr{O}, -)$ is always additive and left exact: the associated right derived functors are $R^n \mathrm{Hom}(\mathscr{O}, -) =: \mathrm{Ext}^n(\mathscr{O}, -)$, for $n \geq 0$.

Let $\mathbb{R}_{\mathbf{S}}(X)$ be the $\mathscr{A}_X$-module defined by the trivial action of $\mathscr{A}_X$ on the abelian group $(\mathbb{R}, +)$ (for $s \in \mathscr{S}_X$ and $r \in \mathbb{R}$, take $s \cdot r = r$). The presheaf that associates to each $X \in \mathrm{Ob}\,\mathbf{S}$ the module $\mathbb{R}_{\mathbf{S}}(X)$, and to each arrow the identity map is denoted $\mathbb{R}_{\mathbf{S}}$.

In Section 1.3, we have defined the *information cohomology* associated to the conditional meet semilattice $\mathbf{S}$, with coefficients in $\mathscr{F} \in \mathbf{Mod}(\mathscr{A})$, as

$$H^{\bullet}(\mathbf{S}, \mathscr{F}) := \mathrm{Ext}^{\bullet}(\mathbb{R}_{\mathbf{S}}, \mathscr{F}). \tag{29}$$

In all the examples contained in this article, [Baudot and Bennequin, 2015] or [Vigneaux, 2019], the sheaf $\mathscr{F}$ is obtained composing the functor $\mathscr{M}$ of an information structure $(\mathbf{S}, \mathscr{M})$ with other functors.

Information cohomology is formally analogous to group cohomology. In this case, one begins with a multiplicative group G and constructs the free abelian group $\mathbb{Z}[G]$, whose elements are finite sums $\sum m_g g$, with $g \in G$ and $m_g \in \mathbb{Z}$. The product of G induces a product between two such elements, and makes $\mathbb{Z}[G]$ a ring, called the integral group ring of G. The category of $\mathbb{Z}[G]$-modules is abelian and has enough injective objects. The cohomology groups of G with coefficients in a $\mathbb{Z}[G]$-module A are defined by

$$H^n(G, A) = \mathrm{Ext}^n(\mathbb{Z}, A), \tag{30}$$

where $\mathbb{Z}$ is the trivial module.

Let $\mathbf{C}$ be an abelian category with enough injectives, and suppose that we are interested in computing the groups $\{\mathrm{Ext}^n(A, B)\}_{n \geq 0}$ for certain fixed objects A and B. In addition, we assume that A has a *projective* resolution $0 \leftarrow A \leftarrow P_0 \leftarrow P_1 \leftarrow \ldots$. Then, Theorem 4.6.10 in [Schapira, 2008] implies that, for all $n \geq 0$,

$$(R^n \mathrm{Hom}_{\mathbf{C}}(A, -))(B) \simeq (R^n \mathrm{Hom}_{\mathbf{C}}(-, B))(A). \tag{31}$$

We denote $(R^n \mathrm{Hom}_{\mathbf{C}}(-, B))(A)$ by $\underline{\mathrm{Ext}}^n(A, B)$. They are given by the formulas

$$\underline{\mathrm{Ext}}^0(A, B) = \ker(\mathrm{Hom}(P_0, B) \to \mathrm{Hom}(P_1, B)), \tag{32}$$

$$\underline{\mathrm{Ext}}^i(A, B) = \frac{\ker(\mathrm{Hom}(P_i, B) \to \mathrm{Hom}(P_{i+1}, B))}{\mathrm{im}(\mathrm{Hom}(P_{i-1}, B) \to \mathrm{Hom}(P_i, B))}, \quad \text{for } i \geq 1. \tag{33}$$

3.2. NONHOMOGENEOUS BAR RESOLUTION. In this section, we introduce a projective resolution of the sheaf of $\mathscr{A}$-modules $\mathbb{R}_{\mathbf{S}}$: a long exact sequence

$$0 \longleftarrow \mathbb{R}_{\mathbf{S}} \xleftarrow{\ \epsilon\ } \mathscr{B}_0 \xleftarrow{\ \partial_1\ } \mathscr{B}_1 \xleftarrow{\ \partial_2\ } \mathscr{B}_2 \xleftarrow{\ \partial_3\ } \ldots \qquad (34)$$

that will allow us to determine the information cohomology, in accordance with (32) and (33).

For any $n \geq 0$, let $\mathscr{B}_n(X)$ be the tensor product over $\mathbb{R}$ of $n+1$ copies of $\mathscr{A}_X$, i.e. $\mathscr{B}_n(X) = \mathscr{A}_X^{\otimes(n+1)}$, equipped with an action of $\mathscr{A}_X$ given by

$$(a, b_0 \otimes b_1 \otimes \cdots \otimes b_n) \mapsto a b_0 \otimes b_1 \otimes \cdots \otimes b_n.$$

Equivalently, $\mathscr{B}_n(X)$ is the free $\mathscr{A}_X$ module generated by the symbols $[X_1|...|X_n] := 1 \otimes X_1 \otimes \cdots \otimes X_n$, where $\{X_1, ..., X_n\} \subset \mathscr{S}_X$. Remark that $\mathscr{B}_0(X)$ is the free module on one generator $[]$. An arrow $X \to Y$ in $\mathbf{S}$ induces an inclusion $\mathscr{S}_Y \hookrightarrow \mathscr{S}_X$, hence an inclusion $\mathscr{B}_n(Y) \hookrightarrow \mathscr{B}_n(X)$, implying that $\mathscr{B}_n = \mathscr{A}^{\otimes n+1}$ is a presheaf of $\mathscr{A}$-modules for each $n \geq 0$.

We introduce now $\mathscr{A}_X$-module morphisms: an augmentation $\epsilon_X : \mathscr{B}_0(X) \to \mathbb{R}_{\mathbf{S}}(X)$ given by the equation $\epsilon([]) = 1$, and boundary morphisms $\partial : \mathscr{B}_n(X) \to \mathscr{B}_{n-1}(X)$ given by

$$\partial([X_1|...|X_n]) = X_1[X_2|...|X_n] + \sum_{k=1}^{n-1}(-1)^k[X_1|...|X_k X_{k+1}|...|X_n] + (-1)^n[X_1|...|X_{n-1}].$$
$$(35)$$

These morphisms are natural in X.

3.2.1. PROPOSITION. *The complex* (34) *is a resolution of the sheaf* $\mathbb{R}_{\mathbf{S}}$.

PROOF. The construction corresponds to the relatively projective bar resolution [Mac Lane, 1994, Ch. IX], more specifically to the example developed at the end of Appendix A, setting $\mathscr{R}$ and $\mathscr{T}$ there equal to $\mathscr{S}$ and $\mathbb{R}_{\mathbf{S}}$, respectively. The resolution $\mathscr{B}_\bullet$ introduced above is $B_\bullet\mathscr{C}$, for $\mathscr{C} = \mathbb{R}_{\mathbf{S}}$. The notation can be simplified, because $\mathscr{C}(X)$ is generated freely by 1 as an $\mathbb{R}$-module. Therefore, $B_0\mathscr{C}$ is generated over $\mathscr{A}_X$ by the symbol $[1]$, written simply as $[]$. In general, $B_n\mathscr{C}(X)$ is generated over $\mathscr{A}_X$ by the symbols $[X_1|...|X_n|1]$, or simply $[X_1|...|X_n]$ if we omit the 1. ∎

Thus far we have a resolution with relatively free objects, that in general need not be projective. However, the special properties of $\mathbf{S}$ allow us to improve the result.

3.2.2. PROPOSITION. *For each* $n \geq 0$, *the sheaf* $\mathscr{B}_n$ *is a projective object in* $\mathbf{Mod}(\mathscr{A})$.

PROOF. Let $\mathscr{T}$ be the presheaf of sets defined by $\mathscr{T}(X) = \{ [X_1|...|X_n] : X_i \in \mathscr{S}_X \}$, for $X \in \mathrm{Ob}\, S$. We have $\mathscr{B}_n = \mathscr{A}[\mathscr{T}]$ i.e. the free presheaf of $\mathscr{A}$-modules generated by $\mathscr{T}$. Like in the case of groups or modules, there is a free-forgetful adjunction [Stacks Project Authors, 2018, Lemma 03A8]

$$\sim \, : \mathrm{Hom}_\mathscr{A}(\mathscr{A}[\mathscr{T}], \mathscr{G}) \xrightarrow{\sim} \mathrm{Hom}_{\mathbf{PSh(S)}}(\mathscr{T}, \mathscr{G}). \qquad (36)$$

To show that $\mathscr{B}_n$ is projective, one should establish the existence of an arrow $\eta :$ $\mathscr{A}[\mathscr{T}] \to \mathscr{C}$ that makes the diagram

$$
\begin{array}{ccc}
 & & \mathscr{A}[\mathscr{T}] \\
 & \overset{\eta}{\swarrow} & \downarrow{\scriptstyle \epsilon} \\
\mathscr{C} & \underset{\sigma}{\twoheadrightarrow} & \mathscr{D}
\end{array}
$$

in $\mathbf{Mod}(\mathscr{A})$ commute, for any epimorphism σ and morphism ϵ. By the adjunction, it suffices to show the existence of a morphism of presheaves $\tilde{\eta} : \mathscr{T} \to \mathscr{C}$ such that the diagram

$$
\begin{array}{ccc}
 & & \mathscr{T} \\
 & \overset{\tilde{\eta}}{\swarrow} & \downarrow{\scriptstyle \tilde{\epsilon}} \\
\mathscr{C} & \underset{\sigma}{\twoheadrightarrow} & \mathscr{D}
\end{array}
$$

in $\mathbf{PSh}(\mathbf{S})$ commutes.

To define $\tilde{\eta}$, one has to determine the image of every symbol $[X_1|...|X_n]$, each time it appears in a set $\mathscr{T}(X)$. Remark that

$$
[X_1|...|X_n] \in \mathscr{T}(X) \Leftrightarrow (\forall i)(X \to X_i) \Leftrightarrow X \to X_1 \cdots X_n = \prod_{i=1}^{n} X_i
$$

The last equivalence is true due to the definition of $\mathbf{S}$. To solve the lifting problem, it is enough to pick $m \in \sigma^{-1}_{\prod_{i=1}^{n} X_i}(\tilde{\epsilon}([X_1|...|X_n]))$, and define $\tilde{\eta}_{\prod_{i=1}^{n} X_i}([X_1|...|X_n]) := m$. This choice gives, by funtoriality, a well defined value $\tilde{\eta}_X([X_1|...|X_n]) = \mathscr{C}\pi(m)$ over each X such that $\pi : X \to \prod_{i=1}^{n} X_i$ in $\mathbf{S}$. ∎

The existence of this projective resolution just depends on the definition of a conditional meet semilattice (Definition 1.2.1). It appears in the computation of classical and quantum information cohomology (see next section and [Baudot and Bennequin, 2015]): the difference between these cases lies in the coefficients.

3.2.3. PROPOSITION. *Given a conditional meet semilattice $\mathbf{S}$ (resp. $\mathbf{S}'$), let $\mathscr{S}$ (resp. $\mathscr{S}'$) denote the associated presheaf of monoids and $\mathscr{A}$ (resp. $\mathscr{A}'$) the presheaf of algebras induced by $\mathscr{S}$ (resp. $\mathscr{S}'$).*

For every morphism $\phi : \mathbf{S} \to \mathbf{S}'$ between conditional meet semilattices such that $\phi(\top) = \top$, there are is a morphism of presheaf of monoids $\phi_* : \mathscr{S} \to \phi^*\mathscr{S}'$ given by $\phi_*^X : \mathscr{S}_X \to \mathscr{S}'_{\phi(X)}$, $Y \mapsto \phi(Y)$, for each $X \in \mathrm{Ob}\,\mathbf{S}$. It can be extended linearly to $\phi_* : \mathscr{A} \to \mathscr{A}'$.

The transformation $\phi_* : \mathscr{A} \to \mathscr{A}'$ induces a map $\phi^* : \mathbf{Mod}(\mathscr{A}') \to \mathbf{Mod}(\mathscr{A})$ as follows: given an $\mathscr{A}'$-module $\mathscr{M}'$, $\mathscr{A}$ acts on $\phi^*\mathscr{M}'$ by the formula

$$
\mathscr{A}_X \times \mathscr{M}'(\phi(X)) \to \mathscr{M}'(\phi(X)), \ (a, m) \to \phi(a)m. \tag{37}
$$

Let $\mathscr{B}_n$ (resp. $\mathscr{B}_n'$) denote $(\mathscr{A})^{\otimes n+1}$ (resp. $(\mathscr{A}')^{\otimes n+1}$), with an action of $\mathscr{A}$ (resp. $\mathscr{A}'$) by left multiplication on the first factor. The maps of $\mathscr{A}$-modules $\Phi_*^n : \mathscr{B}_n \to \phi^* B_n'$, whose

components are

$$\Phi^n_*(X) : \mathscr{B}_n(X) \to \mathscr{B}'_n(\phi(X)), \quad [Y_1|...|Y_n] \mapsto [\phi(Y_1)|...|\phi(Y_n)], \tag{38}$$

define a morphism in $\mathbf{Ch}(\mathbf{Mod}(\mathscr{A}))$, *the category of chain complexes of* $\mathscr{A}$*-modules.*

PROOF. The verifications are straightforward. They are left to the reader. ∎

3.3. DESCRIPTION OF COCYCLES. We have built the projective resolution (34) of $\mathbb{R}_{\mathbf{S}}$ in $\mathbf{Mod}(\mathscr{A})$. For every $\mathscr{A}$-module $\mathscr{F}$, the information cohomology $H^\bullet(\mathbf{S}, \mathscr{F})$ can be computed as $\underline{\mathrm{Ext}}^n(\mathbb{R}_{\mathbf{S}}, \mathscr{F})$, defined in formulas (32) and (33) i.e. we deal with the cohomology of the differential complex $(C^n(\mathbf{S}, \mathscr{F}), \delta)$, where

$$C^n(\mathbf{S}, \mathscr{F}) := \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_n, \mathscr{F})$$

and δ is given by (15) bellow. A morphism f in $C^n(\mathbf{S}, \mathscr{F})$ is called n-cochain. More explicitly, an n-cochain f consists of a collection of morphisms $f_X \in \mathrm{Hom}_{\mathscr{A}_X}(\mathscr{B}_n(X), \mathscr{F}_X)$ that satisfies the following conditions:

1. f is a natural transformation (a functor of presheaves): given $\pi : X \to Y$, the diagram

$$\begin{array}{ccc} \mathscr{B}_n(Y) & \xrightarrow{f_Y} & \mathscr{F}_Y \\ \downarrow & & \downarrow{\scriptstyle \mathscr{F}(\pi)} \\ \mathscr{B}_n(X) & \xrightarrow{f_X} & \mathscr{F}_X \end{array}$$

   commutes. We refer to this property as *(joint) locality*, for reasons that become evident in the following section.

2. f is compatible with the action of $\mathscr{A}$: for every $X \in \mathrm{Ob}\,\mathbf{S}$, the diagram

$$\begin{array}{ccc} \mathscr{A}_X \times \mathscr{B}_n(X) & \longrightarrow & \mathscr{B}_n(X) \\ \downarrow{\scriptstyle 1 \times f_X} & & \downarrow{\scriptstyle f_X} \\ \mathscr{A}_X \times \mathscr{F}_X & \longrightarrow & \mathscr{F}_X \end{array}$$

   commutes. This means that f_X is *equivariant*; in particular, $f_X(Y[Z]) = Y.f_X[Z]$ whenever $Y \in \mathscr{S}_X$.

Since $\mathscr{B}_n(X)$ is a free module, f_X is determined by the values on the generators $[X_1|...|X_n]$. Just to simplify notation, we write $f_X[X_1|...|X_n]$ instead of $f_X([X_1|...|X_n])$.

The coboundary of $f \in C^n(\mathbf{S}, \mathscr{F})$ is the $(n+1)$-cochain $\delta f = f\partial : \mathscr{B}^{n+1} \to \mathscr{F}$; (15) gives a more explicit description. As customary, a cochain $f \in C^n(\mathbf{S}, \mathscr{F})$ is called an n-*cocycle* when $\delta f = 0$; the submodule of all n-cocycles is denoted by $Z^n(\mathbf{S}, \mathscr{F})$. The image under δ of C^{n-1} is another submodule of $C^n(\mathbf{S}, \mathscr{F})$, denoted $\delta C^{n-1}(\mathbf{S}, \mathscr{F})$; its elements are called n-*coboundaries*. By definition, $\delta C^{-1}(\mathbf{S}, \mathscr{F}) = \langle 0 \rangle$, the trivial module. Since $\delta^2 = 0$, δC^{n-1} is a submodule of Z^n. With this notation, $H^n(\mathbf{S}, \mathscr{F}) = Z^n(\mathbf{S}, \mathscr{F})/\delta C^{n-1}(\mathbf{S}, \mathscr{F})$, for every $n \geq 0$.

4. Probabilistic information cohomology

In this section, all information structures are supposed to be finite (see Section 2.1). The treatment of continuous classical random variables presents many technical complications; the particular case of gaussian laws is the subject of [Vigneaux, 2019, Part IV].

We introduce probabilities as a covariant functor on an information structure; the measurable, real-valued probabilistic functions of probabilities (*probabilistic functionals*) form a presheaf $\mathscr{F}_\alpha$ of $\mathscr{A}$-modules; the action of $\mathscr{A}$ depends on a positive parameter α. We compute information cohomology with coefficients in $\mathscr{F}_\alpha$: the corresponding α-entropy appears as the unique 1-cocycle on each "connected component" of the structure.

4.1. PROBABILITIES. Given a finite information structure $(\mathbf{S}, \mathscr{E})$, let $\mathscr{P} : \mathbf{S} \to \mathbf{Sets}$ be the functor that associates to each $X \in \mathrm{Ob}\,\mathbf{S}$ the set

$$\mathscr{P}(X) := \left\{ p : \mathscr{E}_X \to [0,1] \ : \ \sum_{x \in \mathscr{E}_X} p(x) = 1 \right\}, \tag{39}$$

of probability laws for X, and to each arrow $\pi : X \to Y$ the *marginalization* map $\mathscr{P}\pi : \mathscr{P}(X) \to \mathscr{P}(Y)$, also denoted π_*, given by

$$\forall P \in \mathscr{P}_X, \ \forall y \in \mathscr{E}_Y \quad \mathscr{P}\pi(P)(y) = \sum_{x \in \mathscr{E}\pi^{-1}(y)} P(x). \tag{40}$$

We adopt the probabilistic notation, in the following sense: given an arrow $\pi_{YX} : X \to Y$ in $\mathbf{S}$, a law $P \in \mathscr{P}_X$, and $y \in \mathscr{E}(Y)$, the notation $P(Y = y)$ means $P(\mathscr{E}\pi_{YX}^{-1}(y)) = \pi_{YX*}P(y)$; similarly, if $Y \xleftarrow{\pi_{YX}} X \xrightarrow{\pi_{ZX}} Z$ is a diagram in $\mathbf{S}$, the notation $P(Y = y, Z = z) \equiv P(\{Y = y\} \cap \{Z = z\})$ means $P(\mathscr{E}\pi_{YX}^{-1}(y) \cap \mathscr{E}\pi_{ZX}^{-1}(z))$, which equals $P(\mathscr{E}\langle \pi_{YX}, \pi_{ZX}\rangle^{-1}(w(y,z)))$ for the unique $w(y,z) \in \mathscr{E}_{YX}$ sent to $(y,z) \in \mathscr{E}_Y \times \mathscr{E}_Z$ by the injection in Definition 1.2.1-3.

Given an arrow $\pi_{ZX} : X \to Z$, a law $P \in \mathscr{P}(X)$, and $z \in \mathscr{E}_Z$ such that $P(Z = z) > 0$, the conditional law $P|_{Z=z}$ is defined by

$$P|_{Z=z}(x) := \frac{P(\{X = x\} \cap \{Z = z\})}{P(Z = z)} = \frac{P(x \in \mathscr{E}\pi_{ZX}^{-1}(z))}{P(Z = z)}. \tag{41}$$

Conditioning commutes with marginalizations: given arrows $\pi_{YX} : X \to Y$ and $\pi_{ZY} : Y \to Z$,

$$\pi_*^{YX}(P|_{Z=z})(y) = \sum_{x \in \mathscr{E}\pi_{YX}^{-1}(y)} \frac{P(\{x\} \cap \mathscr{E}\pi_{ZX}^{-1}(z))}{P(Z = z)} = \frac{\sum_{x \in \mathscr{E}\pi_{YX}^{-1}(y)} P(\{x\} \cap \mathscr{E}\pi_{ZX}^{-1}(z))}{P(Z = z)}$$

$$= \frac{P(\mathscr{E}\pi_{YX}^{-1}(y) \cap \mathscr{E}\pi_{YX}^{-1}(\mathscr{E}\pi_{YZ}^{-1}(z)))}{P(Z = z)} = \frac{\pi_*^{YX}P(y \cap \mathscr{E}\pi_{YZ}^{-1}(z))}{\pi_*^{YX}P(Z = z)}$$

$$= (\pi_*^{YX}P)|_{Z=z}(y).$$

More generally, an *adapted probability functor* $\mathscr{Q} : \mathbf{S} \to \mathbf{Sets}$ on an information structure $(\mathbf{S}, \mathscr{E})$ is a subfunctor of $\mathscr{P}$ that is stable under conditioning: for every arrow $X \to Z$ in $\mathbf{S}$, every law $P \in \mathscr{Q}_X$, and every $z \in \mathscr{E}_Z$ such that $P(Z = z) > 0$, the law $P|_{Z=z}$ belongs to $\mathscr{Q}_X$.

For instance, a functor $\mathscr{Q}$ that associates to each X a (geometric) simplicial subcomplex $\mathscr{Q}_X$ of the probability simplex $\mathscr{P}_X \subset \mathbb{R}^{E_X}$ is adapted (see [Baudot and Bennequin, 2015, Sec. 2.1]); the restricted marginalizations are simplicial maps. In [Baudot and Bennequin, 2015, Sec. 2.4], it is argued that such functors model exclusion rules between possible events.

4.2. FUNCTIONAL MODULE. Let $(\mathbf{S}, \mathscr{E})$ be an information structure, and $\mathscr{Q}$ an adapted probability functor. For each $X \in \mathrm{Ob}\,\mathbf{S}$, let $\mathscr{F}_X = \mathscr{F}_X(\mathscr{Q})$ be the real vector space of measurable functions on $\mathscr{Q}_X$; we call it functional space. For each arrow $\pi : X \to Y$ in $\mathbf{S}$, there is a morphism $\pi^* : \mathscr{F}_Y \to \mathscr{F}_X$ defined by $\pi^* f(P_X) = f(\pi_* P_X)$. Therefore, $\mathscr{F}$ is a contravariant functor from $\mathbf{S}$ to the category of real vector spaces.

The functional space $\mathscr{F}_X$ admits an action of the monoid $\mathscr{S}_X$ (parameterized by $\alpha > 0$): for $Y \in \mathscr{S}_X$, and $f \in \mathscr{F}_X$, the new function $Y.f$ is given by

$$\forall P \in \mathscr{Q}_X, \quad (Y.f)(P) = \sum_{\substack{y \in \mathscr{E}_Y \\ Y_* P(y) \neq 0}} (Y_* P(y))^\alpha f(P|_{Y=y}). \tag{42}$$

By Proposition 4.2.1, there is a morphism of monoids $\mathscr{S}_X \to \mathrm{End}(\mathscr{F}_X)$, given by Equation (42), that extends by linearity to a morphism of rings $\Lambda_\alpha(X) : \mathscr{A}_X \to \mathrm{End}(\mathscr{F}_X)$. This means that, for each $\alpha > 0$, $\mathscr{F}_X$ has the structure of a $\mathscr{A}_X$-module, denoted $\mathscr{F}_\alpha(X)$.[12]
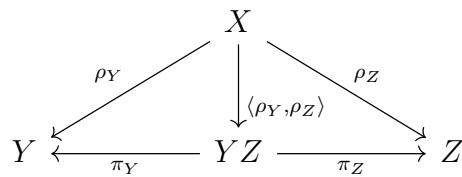
4.2.1. PROPOSITION. *Given any* $X \in \mathrm{Ob}\,\mathbf{S}$, $Y, Z \in \mathscr{S}_X$, *and* $f \in \mathscr{F}(\mathscr{Q}_X)$, *the identities*

$$\top.f = f \quad and \quad (ZY).f = Z.(Y.f)$$

hold.

PROOF. Since it is obvious that $\top.f = f$, we only prove the other.

The universal property of products gives the commutative diagram:

Equation (42) directly implies that, for any $P \in \mathscr{Q}_X$,

$$Z.(Y.f)(P) = \sum_{\substack{z \in \mathscr{E}_Z \\ Z_* P(z) \neq 0}} P(Z = z)^\alpha \sum_{\substack{y \in \mathscr{E}_Y \\ Y_* P|_{Z=z}(y) \neq 0}} (P|_{Z=z}(Y = y))^\alpha f((P|_{Z=z})|_{Y=y}) \tag{43}$$

---

[12] As $\mathscr{A}_X$ is a $\mathbb{R}$-algebra, it comes with an inclusion $f_X : \mathbb{R} \to \mathscr{A}_X, r \mapsto r\top_{\mathbf{S}}$. The composite $\Lambda_\alpha(X) \circ f_X$ gives an action of $\mathbb{R}$ over $\mathscr{F}_X$, that coincides with the usual multiplication of functions by scalars.

By definition of conditional probabilities,

$$P(Z = z)P|_{Z=z}(Y = y) = P(\{Y = y\} \cap \{Z = z\}).$$

The pairs (y, z) that appear in the sum are such that $P(\{Y = y\} \cap \{Z = z\}) \neq 0$, so $P(Y = y)$ and $P(Z = z)$ are different from zero; in this case, for any $B \subset X$,

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

Finally, Definition 1.2.1 guarantees that the nonempty sets

$$\{Y = y\} \cap \{Z = z\} = \mathscr{E}\rho_Y^{-1}(y) \cap \mathscr{E}\rho_Z^{-1}(z) \subset \mathscr{E}_X$$

are the preimage under $\mathscr{E}\langle\rho_Y, \rho_Z\rangle$ of a *unique* element $w_{y,z} \in \mathscr{E}_{YZ}$; moreover, for every element $w \in \mathscr{E}(YZ)$ we find such set. Therefore,

$$Z.(Y.f)(P) = \sum_{\substack{w_{y,z} \in \mathscr{E}_{YZ} \\ YZ_*P(w) \neq 0}} (YZ_*P(w_{y,z}))^\alpha f(P|_{Z=z,Y=y}) = (ZY).f(P).$$

∎

The next proposition shows that this action is compatible with the morphisms between functional modules. Hence, the sheaf $\mathscr{F}_\alpha(\mathcal{Q})$ belongs to $\mathbf{Mod}(\mathscr{A})$, and can be used as coefficients in information cohomology.

4.2.2. PROPOSITION. *Given $\pi_{YX} : X \to Y$ and $\pi_{ZY} : Y \to Z$, the action of Z makes the following diagram commute*

$$\begin{array}{ccc} \mathscr{F}(\mathcal{Q}_Y) & \xrightarrow{Z} & \mathscr{F}(\mathcal{Q}_Y) \\ \downarrow{\scriptstyle \pi_{YX}^*} & & \downarrow{\scriptstyle \pi_{YX}^*} \\ \mathscr{F}(\mathcal{Q}_X) & \xrightarrow{Z} & \mathscr{F}(\mathcal{Q}_X) \end{array}$$

PROOF. We must prove that, for all $f_Y \in \mathscr{F}(\mathcal{Q}_Y), P \in \mathcal{Q}_X$, the equality $(Z.f_Y)(\pi_*^{YX}P) = Z.(f_Y \circ \pi_*^{YX})(P)$. On the one hand,

$$(Z.f_Y)(\pi_*^{YX}P) = \sum_{\substack{z \in \mathscr{E}_Z \\ \pi_*^{ZY}\pi_*^{YX}P(z) \neq 0}} \pi_*^{ZY}\pi_*^{YX}P(z)f_Y((\pi_*^{YX}P)|_{Z=z}), \tag{44}$$

and on the other,

$$Z.(f_Y \circ \pi_*^{YX})(P) = \sum_{\substack{z \in \mathscr{E}_Z \\ \pi_*^{ZX}P(z) \neq 0}} \pi_*^{ZX}P(z)f_Y(\pi_*^{YX}(P|_{Z=z})). \tag{45}$$

The two expressions coincide since marginalizations are functorial, $\pi_*^{ZY}\pi_*^{YX} = \pi_*^{ZX}$, and commute with conditioning (cf. Section 4.1). ∎

4.3. FUNCTORIALITY. Let $\phi = (\phi_0, \widehat{\phi}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{S}', \mathscr{E}')$ be a morphism between finite information structures, and let $\mathscr{Q}$ be a probability functor on $\mathbf{S}$ and $\mathscr{Q}'$, a probability functor on $\mathbf{S}'$ (not necessarily adapted). Given a $X \in \mathrm{Ob}\,\mathbf{S}$ and a law $P \in \mathscr{Q}_X$, define a law $\phi_*^X P$ on $\mathscr{E}'_{\phi_0(X)}$ by the equation

$$\forall x' \in \mathscr{E}_{\phi_0(X)}, \quad \phi_*^X P(x') = \sum_{x \in \widehat{\phi}_X^{-1}(x')} P(x). \tag{46}$$

We call this operation *external marginalization*.

4.3.1. LEMMA. *Provided that for all $X \in \mathrm{Ob}\,\mathbf{S}$ and all $P \in \mathscr{Q}_X$ the law $\phi_*^X P$ belongs to $\mathscr{Q}'_{\phi_0(X)}$,* (46) *defines a natural transformation $\phi_* : \mathscr{Q} \to \phi_0^* \mathscr{Q}'$ i.e. (internal) marginalization and external marginalization commute.*

PROOF. For every arrow $\pi : X \to Y$ and $y' \in \mathscr{E}_{\phi(Y)}$,

$$(\mathscr{Q}'(\phi_0 \pi)(\phi_*^X P))(y') \overset{(\mathrm{def.}\ \mathscr{Q}')}{=} \sum_{x' \in (\mathscr{E}'\phi_0\pi)^{-1}(y')} (\phi_*^X(P))(x')$$

$$\overset{(\mathrm{def.}\ \phi_*^X)}{=} \sum_{x' \in (\mathscr{E}'\phi_0\pi)^{-1}(y')} \sum_{x \in \widehat{\phi}_X^{-1}(x')} P(x)$$

$$= \sum_{x \in (\mathscr{E}\phi_0\pi \circ \widehat{\phi}_X)^{-1}(y')} P(x)$$

$$= \sum_{x \in (\widehat{\phi}_Y \circ \mathscr{E}\pi)^{-1}(y')} P(x)$$

$$= \sum_{y \in \widehat{\phi}_Y^{-1}(y')} \sum_{x \in \mathscr{E}\pi^{-1}(y)} P(x)$$

$$\overset{(\mathrm{def.}\ \phi_*^Y)}{=} \phi_*^Y(\mathscr{Q}(\pi)(P)).$$

The fourth equality comes from the naturality of $\widehat{\phi}$, cf. Definition 2.2.1. ∎

Using the previous lemma, it is easy to verify that ϕ_* induces a natural transformation $\phi^* : \phi_0^* \mathscr{F}' \to \mathscr{F}$—where $\mathscr{F}_\alpha = \mathscr{F}(\mathscr{Q})$ and $\mathscr{F}' = \mathscr{F}_\alpha(\mathscr{Q}')$—that maps $f' \in \mathscr{F}'_{\phi_0(X)}$ to $f' \circ \phi_*^X$.

Some morphisms of information structures induce morphisms at the level of cohomology.

4.3.2. PROPOSITION. *Let $\phi = (\phi_0, \widehat{\phi}) : (\mathbf{S}, \mathscr{E}) \to (\mathbf{S}', \mathscr{E}')$ be a morphism of information structures; let $\mathscr{Q}$ (resp. $\mathscr{Q}'$) be an adapted probability functor on $\mathbf{S}$ (resp. $\mathbf{S}'$). Suppose that*

  1. for all $X \in \mathrm{Ob}\,\mathbf{S}$, the map $\widehat{\phi}_X$ is a bijection, and

  2. for all $X \in \mathrm{Ob}\,\mathbf{S}$ and all $P \in \mathscr{Q}_X$, the law $\phi_*^X P$ belongs to $\mathscr{Q}'_{\phi(X)}$.

Then, there exist a cochain map

$$\Phi_\bullet^* : (C^\bullet(\mathbf{S}', \mathscr{F}_\alpha'), \delta) \to (C^\bullet(\mathbf{S}, \mathscr{F}_\alpha), \delta), \tag{47}$$

where $\mathscr{F}_\alpha = \mathscr{F}_\alpha(\mathscr{Q})$ and $\mathscr{F}_\alpha' = \mathscr{F}_\alpha(\mathscr{Q}')$, given by the formula

$$\forall f' \in C^n(\mathbf{S}', \mathscr{F}_\alpha'), \quad (\Phi_n^* f')_Y[X_1|...|X_n](P) := f'_{\phi_0(Y)}[\phi_0(X_1)|...|\phi_0(X_n)](\phi_*^Y(P)). \tag{48}$$

The chain map induces a morphism of graded vector spaces in cohomology

$$\Phi_\bullet^* : H^\bullet(\mathbf{S}', \mathscr{F}_\alpha') \to H^\bullet(\mathbf{S}, \mathscr{F}_\alpha). \tag{49}$$

PROOF. From Proposition 3.2.3, we know that ϕ induces a morphism

$$\{\Phi_*^n : \mathscr{B}_n \to \mathscr{B}_n' \circ \phi_0\}_n$$

of chain complexes of $\mathscr{A}$-modules.

Let f' be an element of $\mathrm{Hom}_{\mathscr{A}'}(\mathscr{B}_n', \mathscr{F}_\alpha')$.

First, we claim that $\Phi_n^* f'$ is a natural transformation. Given an arrow $\pi : X \to Y$ in **S**, the diagram

$$
\begin{array}{ccccccc}
\mathscr{B}_n(Y) & \xrightarrow{\Phi_*^n(Y)} & \mathscr{B}_n'(\phi_0(Y)) & \xrightarrow{f'_{\phi_0(Y)}} & \mathscr{F}'(\phi_0(Y)) & \xrightarrow{\phi_Y^*} & \mathscr{F}(Y) \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle \mathscr{F}'\phi_0\pi} & & \downarrow{\scriptstyle \mathscr{F}\pi} \\
\mathscr{B}_n(X) & \xrightarrow{\Phi_*^n(X)} & \mathscr{B}_n'(\phi_0(X)) & \xrightarrow{f'_{\phi_0(X)}} & \mathscr{F}'(\phi_0(X)) & \xrightarrow{\phi_X^*} & \mathscr{F}(X)
\end{array}
$$

in $\mathbf{PSh}(\mathbf{S})$ commutes, because Φ_*^n, f', and ϕ^* are natural transformations. Since the composition of three successive horizontal arrows give the components $(\Phi_n^* f)_X$ and $(\Phi_n^* f)_Y$, this implies the claim.

Second, $\Phi_n^* f'$ is a morphism of $\mathscr{A}$-modules, which means that for every $X \in \mathrm{Ob}\,\mathbf{S}$ the diagram

$$
\begin{array}{ccc}
\mathscr{A}_X \times \mathscr{B}_n(X) & \longrightarrow & \mathscr{B}_n(X) \\
\downarrow{\scriptstyle 1 \times (\Phi_n^* f')_X} & & \downarrow{\scriptstyle (\Phi_n^* f')_X} \\
\mathscr{A} \times \mathscr{F}(X) & \longrightarrow & \mathscr{F}(X)
\end{array}
$$

commutes i.e. for all $X_0, ..., X_n \in \mathscr{S}_X$,

$$\phi_X^*(f'_{\phi_0(X)}(\phi_0(X_0)[\phi(X_1)|...|\phi(X_n)])) = X_0.(\phi_X^* f'_{\phi_0(X)}([\phi(X_1)|...|\phi(X_n)])) \tag{50}$$

as functions on $\mathscr{Q}_X$. For any P in $\mathscr{Q}_X$, one has

$$
\begin{aligned}
(\phi_X^*(&f'_{\phi_0(X)}(\phi_0(X_0)[\phi(X_1)|...|\phi(X_n)])))(P) \\
&= (f'_{\phi_0(X)}(\phi_0(X_0)[\phi(X_1)|...|\phi(X_n)]))(\phi_*^X P) \\
&= (\phi_0(X_0).f'_{\phi_0(X)}([\phi(X_1)|...|\phi(X_n)]))(\phi_*^X P) \\
&= \sum_{x_0' \in \mathscr{E}'_{\phi_0(X_0)}} \phi_*^X P(\phi_0(X_0) = x_0) f'_{\phi_0(X)}([\phi(X_1)|...|\phi(X_n)])((\phi_*^X P)|_{\phi_0(x_0)=x_0'}), \tag{51}
\end{aligned}
$$

where the first equality is implied by the definition of ϕ^*, the second by the $\mathscr{A}'$-equivariance of f', and the third by the definition of the action of $\mathscr{S}'_X$ on $\mathscr{F}'$. Since $\widehat{\phi}_X$ is a bijection, each $x'_0 \in \mathscr{E}'_{\phi_0(X_0)}$ is in correspondence with an element $x_0 = x_0(x'_0)$ of $\mathscr{E}_{X_0}$, and

$$\phi^X_* P(\phi_0(X_0) = x'_0) = P(\widehat{\phi}^{-1}_X((\mathscr{E}'\phi_0\pi_{X_0X})^{-1}(x'_0)))$$

equals

$$P(X_0 = x_0) = P((\mathscr{E}\pi_{X_0X})^{-1}(\widehat{\phi}^{-1}_{X_0}(x'_0)))$$

as a simple consequence of the naturality of $\widehat{\phi}$, which applied to $\pi_{X_0X} : X \to X_0$ gives the commutative diagram

$$
\begin{array}{ccc}
\mathscr{E}_X & \xrightarrow{\widehat{\phi}_X} & \mathscr{E}_{\phi_0(X)} \\
{\scriptstyle \mathscr{E}\pi_{X_0X}}\downarrow & & \downarrow{\scriptstyle \mathscr{E}'\phi_0\pi_{X_0X}} \\
\mathscr{E}_{X_0} & \xrightarrow{\widehat{\phi}_{X_0}} & \mathscr{E}_{\phi_0(X_0)}
\end{array} \ .
$$

So, provided that conditioning commutes with external marginalization, i.e.

$$(\phi^X_* P)|_{\phi_0(x_0)=x'_0} = \phi^X_*(P|_{X_0=x_0}), \tag{52}$$

the sum in (51) can be rewritten as

$$\sum_{x_0 \in \mathscr{E}_{X_0}} P(X_0 = x_0) f'_{\phi_0(X)}([\phi(X_1)|...|\phi(X_n)])(\phi^X_*(P|_{X_0=x_0})), \tag{53}$$

which is precisely the right-hand side of (50) evaluated at P. Hence $\Phi^*_n f'$ is a morphism of $\mathscr{A}$-modules.

To establish (52), remark that for any $z' \in \mathscr{E}_{\phi_0(X)}$ there is a unique $z \in \mathscr{E}_X$ such that $\widehat{\phi}_X(z) = z'$, and

$$
\begin{aligned}
(\phi^X_* P)|_{\phi_0(X_0)=x'_0}(z') &= \frac{\phi^X_* P(z' \in (\mathscr{E}'\phi_0\pi_{X_0X})^{-1}(x'_0))}{\phi^X_* P(\phi_0(X_0) = x'_0)} \\
&= \frac{P(z \in \widehat{\phi}^{-1}_X((\mathscr{E}'\phi_0\pi_{X_0X})^{-1}(x'_0)))}{P(X_0 = x_0)} \\
&= \frac{P(z \in (\mathscr{E}\pi_{X_0X})^{-1}(\widehat{\phi}^{-1}_{X_0}(x'_0)))}{P(X_0 = x_0)},
\end{aligned}
$$

where the first equality follows from the definition of conditioning, the second from the definition of external marginalization, and the third from the naturality of $\widehat{\phi}$. Since $\widehat{\phi}^{-1}_{X_0}(x'_0) = x_0$, the last expression equals $P|_{X_0=x_0}(z) = (\phi^X_*(P|_{X_0=x_0}))(z')$.

Finally, the commutativity of the diagram

$$
\begin{array}{ccccc}
\mathscr{B}_{n+1}(X) & \xrightarrow{(\partial_{n+1})_X} & \mathscr{B}_n(X) & \xrightarrow{(\Phi_n^* f')_X} & \mathscr{F}(X) \\
{\scriptstyle (\Phi_*^{n+1})_X} \downarrow & & {\scriptstyle (\Phi_*^n)_X} \downarrow & & \uparrow {\scriptstyle \phi_X^*} \\
\mathscr{B}'_{n+1}(\phi_0(X)) & \xrightarrow{(\partial_{n+1})_X} & \mathscr{B}'_n(\phi_0(X)) & \xrightarrow{f'_{\phi_0(X)}} & \mathscr{F}'(\phi_0(X)),
\end{array}
$$

with the upper arc labelled $(\delta^n \Phi_n^* f')_X$ and the lower arc labelled $(\delta^n f')_{\phi_0(X)}$

—which follows from the definitions and the naturality of Φ_*^n—entails

$$(\delta^n f')_X = \phi_X^* \circ (\delta^n f')_{\phi_0(X)} \circ (\Psi_*^{n+1})_X =: (\Phi_{n+1}^*(\delta^n f))_X.$$

In other words, $\Phi_\bullet^*$ and δ commute, thus $\Phi_\bullet^*$ is a cochain map and as such it induces a morphism at the level of cohomology. ∎

4.3.3. REMARK. TAC's referee has suggested that the pullback functor $\phi^* : \mathbf{Mod}(\mathscr{A}') \to \mathbf{Mod}(\mathscr{A})$ in Proposition 4.2.2 might be exact and induce a transformation $H^\bullet(\mathbf{S}', \mathscr{F}') \to H^\bullet(\mathbf{S}, \phi^*\mathscr{F})$ for every $\mathscr{A}'$-module $\mathscr{F}'$. They conjecture that the composition of such transformation with another map coming from the functoriality of $H^\bullet(\mathbf{S}, -)$ may then reproduce the morphism $\Phi_\bullet^*$ in (49), avoiding explicit computations with cocycles.

4.3.4. COROLLARY. *If $\phi : \mathbf{S} \to \mathbf{S}'$ is an isomorphism of information structures, and $\mathscr{Q}'(\phi(X)) = \phi_*^X(\mathscr{Q}_X)$ for every $X \in \mathrm{Ob}\,\mathbf{S}$, then $\Phi^* : H^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')) \to H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ is an isomorphism too.*

PROOF. For every $X \in \mathrm{Ob}\,\mathbf{S}$, one has $\widehat{\phi^{-1}}_{\phi(X)} \circ \widehat{\phi}_X = \mathrm{id}_X$ and $\widehat{\phi}_X \circ \widehat{\phi^{-1}}_{\phi(X)} = \mathrm{id}_{\phi(X)}$, thus $\widehat{\phi}_X$ and $\widehat{\phi^{-1}}_{\phi(X)}$ are bijections. Proposition 4.3.2 ensures the existence of $\Phi_\bullet^* : H^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')) \to H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$ and $(\Phi^{-1})_\bullet^* : H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \to H^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}'))$ that are inverse to each other, which can be easily verified using (48). ∎

We recover from Proposition 4.3.2 a functorial property for concrete information structures.

4.3.5. PROPOSITION. *Consider concrete information structures $\mathbf{S} \subset \mathbf{Obs}_{\mathrm{fin}}(\Omega)$ and $\mathbf{S}' \subset \mathbf{Obs}_{\mathrm{fin}}(\Omega')$. Let $\mathscr{Q}$ (resp. $\mathscr{Q}'$) be an adapted probability functor defined on $\mathbf{S}$ (resp. $\mathbf{S}'$). Let $\sigma : \Omega \to \Omega'$ be a surjective function. Suppose that for all $X \in \mathrm{Ob}\,\mathbf{S}$, there exists $X' \in \mathrm{Ob}\,\mathbf{S}'$ such that σ descends to a bijection $\sigma_X : \Omega/X \xrightarrow{\sim} \Omega'/X'$, [13] and that for every $P \in \mathscr{Q}_X$, the marginalization $\sigma_{X*}P$ is in $\mathscr{Q}'_{X'}$.*

Then the observable X' is uniquely determined by X, the correspondence $\phi_0 : X \mapsto X'$ defines a morphism of information structures $\phi : (\mathbf{S}, \square) \to (\mathbf{S}', \square)$—see the notation in Example 2.1.1—, and there exists a morphism of graded vector spaces

$$\sigma^* : H^\bullet(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')) \to H^\bullet(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})), \tag{54}$$

defined at the level of cochains by (48), mutatis mutandis.

---

[13]Every partition X defines an equivalence relation and Ω/X denotes the corresponding quotient.

PROOF. Since σ is surjective, $\sigma^{-1}(A) \neq \sigma^{-1}(B)$ unless $A = B$, cf. footnote 7. If X', X'' of Ω' satisfying $F : \Omega/X \xrightarrow{\sim} \Omega'/X'$ and $G : \Omega/X \xrightarrow{\sim} \Omega'/X''$, then for every $A \in X$, $A = \sigma^{-1}(F(X)) = \sigma^{-1}(G(X))$, therefore $X' = X''$. The correspondence $X \mapsto X'$ defines a functor $\phi_0 : \mathbf{S} \to \mathbf{S}'$: given an arrow $\pi : X \to Y$ in $\mathbf{S}$, there is a corresponding surjection $\Box\pi : \Omega/X \to \Omega/Y$—using an obvious identification between each part and the corresponding class in the quotient—and $\sigma_Y \circ \Box\pi \circ \sigma_X^{-1} : \Omega'/\phi_0(X) \to \Omega'/\phi_0(Y)$ is also a surjection, that corresponds to an arrow $\phi(X) \to \phi(Y)$ in $\mathbf{S}'$. We take as $\widehat{\phi}_X : X \to \phi(X)$ the bijection of partitions induced by σ_X. The condition on the probabilities implies that $\phi_* : \mathscr{Q} \to \mathscr{Q}' \circ \phi_0$ is a natural transformation. Proposition 4.3.2 gives the desired result. $\blacksquare$

4.3.6. PROPOSITION. *Consider concrete information structures* $\mathbf{S} \subset \mathbf{Obs}_{\mathrm{fin}}(\Omega)$ *and* $\mathbf{S}' \subset \mathbf{Obs}_{\mathrm{fin}}(\Omega')$. *Let* $\mathscr{Q}$ *(resp.* $\mathscr{Q}'$*) be an adapted probability functor defined on* $\mathbf{S}$ *(resp.* $\mathbf{S}'$*). Let* $\eta : \Omega \to \Omega'$ *be a function. Suppose that for all* $X' \in \mathrm{Ob}\,\mathbf{S}$, *there exists* $X \in \mathrm{Ob}\,\mathbf{S}$ *such that* η *descends to a bijection* $\eta_X : \Omega/X \xrightarrow{\sim} \Omega'/X'$, *and that for every* $P' \in \mathscr{Q}'_{X'}$, *there exists* $P \in \mathscr{Q}_X$ *with* $P' = \eta_{X*}P$.

Then the observable X *is uniquely determined by* X', *the correspondence* $\phi_0 : X' \mapsto X$ *defines a morphism of information structures* $\phi : (\mathbf{S}', \Box) \to (\mathbf{S}, \Box)$, *and there exists a morphism of graded vector spaces*

$$\eta_* : H^m(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \to H^m(\mathbf{S}', \mathscr{F}_\alpha(\mathscr{Q}')), \tag{55}$$

defined at the level of cochains by (48)*, mutatis mutandis.*

PROOF. In view of the component-wise bijection, X is the partition $\{\eta_X^{-1}(B) : B \in X'\}$. The rest of the proof is analogous to the last one. $\blacksquare$

4.4. DETERMINATION OF H^0. Each 0-cochain $f[] \equiv f$ corresponds to a collection of functions $f_X(P_X) \in \mathscr{F}_\alpha(\mathscr{Q}_X)$, for each $X \in \mathrm{Ob}\,\mathbf{S}$, that satisfy $f_Y(Y_*P_X) = f_X(P_X)$ for any arrow $X \to Y$ in $\mathbf{S}$. Since $\top \in \mathrm{Ob}\,\mathbf{S}$, this means that f is constant. Given an arrow $X \to Y$, and a 0-cochain f such that $f_X(P) = K$,

$$(\delta f)_X[Y](P) = Y.f_X(P) - f_X(P) = \sum_{y \in \mathscr{E}_Y} P(Y = y)^\alpha f(P|_{Y=y}) - f(P)$$

$$= K \left(\sum_{y \in \mathscr{E}_Y} P(Y = y)^\alpha - 1 \right) = 0.$$

Thus $Z^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) = C^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) \cong \mathbb{R}$ and $Z^0(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) = \langle 0 \rangle$ when $\alpha \neq 1$ (as long as some $\mathscr{Q}_Y$ contains a nonatomic probability). Therefore, $H^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) \cong \mathbb{R}$, and $H^0(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) \cong \langle 0 \rangle$ when $\alpha \neq 1$.

4.5. LOCAL STRUCTURE OF 1-COCYCLES. Now we turn to $C^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$. The 1-cochains are families $\{f_X[Y] \mid X \in \mathrm{Ob}\,\mathbf{S}\}$ such that for all $Z \to X \to Y$, the equality $f_X[Y](X_*P_Z) = f_Z[Y](P_Z)$ holds (this is the *locality* in Section 3.3). This means that it is sufficient to know $f_Y[Y](Y_*P)$ to recover $f_X[Y](P)$, for any $X \to Y$; in this

sense, we usually omit the subindex and just write $f[Y]$. It follows that there is a bijective correspondence between 1-cochains f and collections of measurable functions $\{f[X] : \mathcal{Q}_X \to \mathbb{R}\}_{X \in \mathrm{Ob}\,\mathbf{S}}$.

The computations in the previous section imply that $\delta C^0(\mathbf{S}, \mathscr{F}_1(\mathcal{Q})) = \langle 0 \rangle$, whereas $\delta C^0(\mathbf{S}, \mathscr{F}_\alpha(\mathcal{Q})) \cong \mathbb{R}$ when $\alpha \neq 1$: in this case 1-coboundaries are multiples of the 1-cochain defined by $S_\alpha[X]$ in (6). We write $\delta C^0(\mathbf{S}, \mathscr{F}_\alpha(\mathcal{Q})) \cong \mathbb{R} \cdot S_\alpha$.

We refer to elements of $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathcal{Q}))$ as 1-cocycles of type α. By equation (15) and commutativity of the product, every 1-cocycle must satisfy the following symmetric equation

$$f[XY] = f[Y] + Y.f[X] = f[X] + X.f[Y]. \tag{56}$$

4.5.1. PROPOSITION. *Let f be a 1-cocycle.*

1. *For every $X \in \mathrm{Ob}\,\mathbf{S}$, if $|\mathscr{E}_X| = 1$, then $f[X] \equiv 0$. In particular, $f[\top] \equiv 0$.*

2. *For every $X \in \mathrm{Ob}\,\mathbf{S}$ and $x \in \mathscr{E}_X$, the equality $f[X](\delta_x) = 0$ holds.*

PROOF. Statement (1) is a particular case of (2); we prove the later. From $f[XX] = f[X] + X.f[X]$, we conclude that $X.f[X](P) = \sum_{x \in \mathscr{E}_X | P(x) \neq 0} P(x)^\alpha f[X](P|_{X=x}) = 0$, for any $P \in \mathcal{Q}_X$. Setting $P = \delta_x$, one obtains $f[X](\delta_x) = 0$. ∎

4.5.2. EXAMPLE. We compute $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$, taking $\mathbf{S}$ equal to $\bot \to \top$, and $\mathscr{E}(\bot) = \{a, b\}$. Proposition 4.5.1 implies that $f[\bot](1, 0) = f[\bot](0, 1) = 0$, as a consequence of $f[\bot] = f[\bot] + \bot.f[\bot]$. All the other relations derived from the cocycle condition (56) become tautological. Therefore, 1-cocycles are in correspondence with measurable functions f on arguments (p_a, p_b) such that $f(1, 0) = f(0, 1) = 0$. We conclude that $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$ has infinite dimension. For a more general condition under which $\dim H^1$ diverges, see Proposition 4.6.3.

The functions $S_\alpha[X]$ introduced in (1) and (6) define a 1-cochain, that is a 1-cocycle according to the following proposition.

4.5.3. PROPOSITION. [Chain rule] *Let $(\mathbf{S}, \mathscr{E})$ be a finite information structure, $\mathcal{Q}$ an adapted probability functor, X an element of $\mathrm{Ob}\,\mathbf{S}$, and Y, Z elements of $\mathscr{S}_X$. Then, for all $\alpha > 0$, the 1-cochain $S_\alpha \in \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_1, \mathscr{F}_\alpha(\mathcal{Q}))$ satisfies*

$$S_\alpha[YZ] = S_\alpha[Y] + Y.S_\alpha[Z]. \tag{57}$$

This means that S_α belongs to $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathcal{Q}))$.

PROOF. Let P be a probability in $\mathcal{Q}_X$. We further simplify the notation, writing $P(y)$ instead of $P(Y = y) = Y_*P(y)$, and $P(z|y)$ in place of $P(Z = z|Y = y)$. We label the points in $\mathscr{E}(YZ)$ by their image under the injection $\iota : \mathscr{E}(YZ) \to \mathscr{E}(Y) \times \mathscr{E}(Z)$, writing $w(y, z) \in \mathscr{E}(YZ)$.

1. Case $\alpha = 1$: by definition

$$-S_1[YZ](P) = \sum_{w(y,z)\in\mathscr{E}(YZ)} P(y,z)\log P(y,z)$$

and in fact we can extend this to a sum over the whole set $\mathscr{E}(X) \times \mathscr{E}(Y)$, setting $P(y,z) = 0$ whenever $(y,z) \notin \text{im}\,\iota$ (following the convention $0\log 0 = 0$). We rewrite the previous expression using the conditional probabilities

$$-S_1[YZ](P) = \sum_{\substack{y\in\mathscr{E}_Y \\ P(y)>0}} \sum_{z\in\mathscr{E}_Z} P(z|y)P(y)(\log P(y) + \log P(z|y))$$

$$= \sum_{\substack{y\in\mathscr{E}_Y \\ P(y)>0}} P(y)\log P(y) \sum_{z\in\mathscr{E}_Z} P(z|y) + \sum_{\substack{y\in\mathscr{E}_Y \\ P(y)>0}} P(y) \sum_{z\in\mathscr{E}_Z} P(z|y)\log P(z|y).$$

This gives the result, because $\sum_{z\in\mathscr{E}_Z} P(z|y) = 1$, and $\sum_{z\in\mathscr{E}_Z} P(z|y)\log P(z|y) = S_1[Z](P|_{Y=y})$. Cf. [Khinchin, 1957].

2. Case $\alpha \neq 1$: The result is a consequence of $\delta^2 = 0$, but can be proved by a direct computation.

$$(1-\alpha)(S[Y] + Y.S[Z]) = \left(\sum_{y\in\mathscr{E}_Y} P(y)^\alpha - 1\right) + \sum_{\substack{y\in\mathscr{E}_Y \\ P(y)>0}} P(y)^\alpha \left(\sum_{z\in\mathscr{E}_Z} P(z|y)^\alpha - 1\right)$$

$$= \sum_{\substack{y\in\mathscr{E}_Y \\ P(y)>0}} \sum_{z\in Z} P(z|y)^\alpha P(y)^\alpha - 1$$

$$= (1-\alpha)S[XY].$$

The last equality comes from $P(z|y)P(y) = P(z,y)$, and the fact that we can restrict the sum to $\mathscr{E}(YZ)$, neglecting terms that vanish.

$\blacksquare$

We shall see that any nontrivial 1-cocycle of type α is locally a multiple of S_α. Proposition 4.5.4 presents the solution to a functional equation that comes from the cocycle condition. Then, Proposition 4.5.7 determines the local form of a cocycle. Finally, Theorem 1.3.2 determine H^1 under appropriate nondegeneracy hypotheses on the information structure **S** and the probability functor $\mathscr{Q}$.

For convenience, we introduce the functions

$$s_1(p) := -p\log p - (1-p)\log(1-p); \tag{58}$$

$$s_\alpha(p) := \frac{1}{1-\alpha}(p^\alpha + (1-p)^\alpha - 1) \qquad (\text{for } \alpha \neq 1), \tag{59}$$

both defined for $p \in [0,1]$.

4.5.4. THEOREM. [Generalized FEITH] *Let $f_1, f_2 : \Delta^2 \to \mathbb{R}$ be two unknown measurable functions satisfying*

   1. $f_i(0,1) = f_i(1,0) = 0$ for $i = 1, 2$.

   2. for all $(p_0, p_1, p_2) \in \Delta^2$ such that $p_1 > 0$ and $p_2 > 0$,

$$(1 - p_2)^\alpha f_1 \left(\frac{p_0}{1 - p_2}, \frac{p_1}{1 - p_2} \right) - f_1(1 - p_1, p_1) \tag{60}$$

$$= (1 - p_1)^\alpha f_2 \left(\frac{p_0}{1 - p_1}, \frac{p_2}{1 - p_1} \right) - f_2(1 - p_2, p_2).$$

Then, $f_1 = f_2$ and there exists $\lambda \in \mathbb{R}$ such that $f_1(p) = \lambda s_\alpha(p)$.

PROOF. The restriction to $p_0 = 0$ (with $p_1 = x, p_2 = 1 - x$) implies $f_2(x, 1 - x) = f_1(1 - x, x)$. This can be used to rewrite (60) in terms of $u(x) := f_1(x, 1 - x)$. Setting $p_1 = x$, $p_2 = y$ and $p_0 = 1 - x - y$, we obtain the functional equation

$$u(1 - x) + (1 - x)^\alpha u \left(\frac{y}{1 - x} \right) = u(y) + (1 - y)^\alpha u \left(\frac{1 - x - y}{1 - y} \right). \tag{61}$$

This functional equation is related to the so-called "fundamental equation of information theory" (FEITH), which first appeared in [Tverberg, 1958]. Every measurable solution of (61) with $\alpha = 1$ has the form $u(x) = \lambda s_1(x)$, with $\lambda \in \mathbb{R}$ [Kannappan and Ng, 1973]. Analogously, [Bennequin and Vigneaux, 2020] shows that the general solution in the case $\alpha \neq 1$ is $u(x) = \lambda s_\alpha(x)$, with $\lambda \in \mathbb{R}$; this is directly connected to a generalization of the fundamental equation introduced in [Daróczy, 1970]. ∎

4.5.5. EXAMPLE. Let **S** be the poset represented by

$$\begin{array}{ccc} & \top & \\ \nearrow & & \nwarrow \\ X_1 & & X_2 \\ \nwarrow & & \nearrow \\ & X_1 X_2 & \end{array} \tag{62}$$
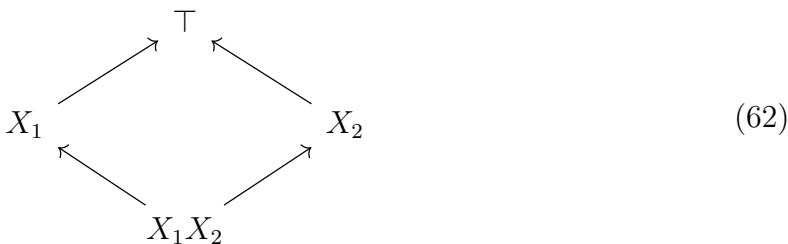
and $\mathscr{E}$ be the functor defined at the level of objects by $\mathscr{E}(X_1) = \{\{1\}, \{0, 2\}\}$, $\mathscr{E}(X_2) = \{\{2\}, \{0, 1\}\}$, and $\mathscr{E}(X_1 X_2) = \{\{0\}, \{1\}, \{2\}\}$; for each arrow $\pi : X \to Y$, the map $\pi_* : \mathscr{E}(X) \to \mathscr{E}(Y)$ sends $I \to J$ iff $I \subset J$. The pair $(\mathbf{S}, \mathscr{E})$ is an information structure (it comes from a concrete one). Consider $f \in Z^1(\mathscr{F}_\alpha(\mathscr{P}))$: the 1-cocycle condition means that, as functions on $\mathscr{P}(X_1 X_2)$,

$$f[X_1 X_2] = X_1.f[X_2] + f[X_1] \quad \text{and} \quad f[X_1 X_2] = X_2.f[X_1] + f[X_2]. \tag{63}$$

We write f_1, f_2 and f_{12} instead of $f[X_1]$, $f[X_2]$ and $f[X_1X_2]$, respectively. The determination of f_1 and f_2 such that $X_1.f_2 + f_1 = X_2.f_1 + f_2$ fix f_{12} completely. In terms of a probability (p_0, p_1, p_2) in $\mathscr{P}(X_1X_2)$ this equation is exactly (60). We conclude that every cocycle is a multiple of the corresponding α-entropy: there exists a unique constant $\lambda \in \mathbb{R}$ such that $f[Z](P) = \lambda S_\alpha[Z](P)$, for every observable $Z \in \mathrm{Ob}\,\mathbf{S}$ and every probability law $P \in \mathscr{P}(X_1X_2)$. This establishes that $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \cong \mathbb{R}$. Hence $H^1(\mathbf{S}, \mathscr{F}_1(\mathscr{P})) \cong \mathbb{R}$, and $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) \cong \langle 0 \rangle$. The hypotheses are minimal: on the one hand, if we remove X_1 or X_2, Proposition 4.6.3 shows that $\dim H^1 = \infty$; on the other, if $\mathscr{Q}_{X_1X_2}$ does not contain the interior of Δ^2, the equations (63) translate into a "degenerate" system of functional equations

$$f_{12}(p_0, 1 - p_0, 0) = f_1(1 - p_0, p_0),$$
$$f_{12}(p_0, 0, 1 - p_0) = f_2(1 - p_0, p_0),$$
$$f_{12}(0, p_1, 1 - p_1) = f_1(p_1, 1 - p_1) = f_2(1 - p_1, p_1),$$

thus one function remains arbitrary.

We want to generalize the previous computations to any product XY. It is already clear that $\mathscr{E}_{XY} = \mathscr{E}_X \times \mathscr{E}_Y$ is not necessary; at the same time, if one is "too far" from $\mathscr{E}_{XY} = \mathscr{E}_X \times \mathscr{E}_Y$ and $\mathscr{Q} = \mathscr{P}$, then the system of functional equations implied by the 1-cocycle condition degenerates. A precise sufficient condition is introduced in the following definition of *nondegenerate product*, which is better understood looking at the proof of Proposition 4.5.7. To determine the function $f[XY]$, for given observables X and Y, one first obtains the recursive formulas (67) and (68) for the functions $f[X]$ and $f[Y]$: given a total order of the sets $\mathscr{E}_X$ and $\mathscr{E}_Y$, the different steps of the recursion are coded by a path in $\mathbb{Z}^2$. Both formulas are a simplification of the symmetric equation (66) for particular laws $\tilde{P}$ given by Definition 4.5.6-1 that make one of the terms trivial. The recursive formulas involve a term where $f[X]$ and $f[Y]$ have only two nonzero arguments, and both are related by the FEITH-like functional equation (60) in Proposition 4.5.4; that this equation holds for any $(p_0, p_1, p_2) \in \Delta^2$ such that $p_1 > 0$ and $p_2 > 0$ is ensured by Definition 4.5.6-2.

4.5.6. DEFINITION. *Let X and Y be two objects of $\mathbf{S}$, such that $|\mathscr{E}_X| = k$ and $|\mathscr{E}_Y| = l$. Let ι be the inclusion $\mathscr{E}_{XY} \hookrightarrow \mathscr{E}_X \times \mathscr{E}_Y$. We call the product XY nondegenerate if $k, l \geq 2$ and there exist enumerations $\{x_1, ..., x_k\}$ of $\mathscr{E}_X$ and $\{y_1, ..., y_l\}$ of $\mathscr{E}_Y$, together with a North-East (NE) lattice path[14] $(\gamma_i)_{i=1}^m$ on $\mathbb{Z}^2$ going from $(1, 1)$ to (k, l), such that:*

1. *If $\gamma_i = (a, b)$ and $\gamma_{i+1} - \gamma_i = (1, 0)$, then for every $P \in \mathscr{Q}_X$ such that $\mathrm{supp}\,P \subset \{x_j : a \leq j \leq k\}$, there exists $\tilde{P} \in \mathscr{Q}_{XY}$ whose support is contained in*

$$\iota^{-1}(\{(x_a, y_{b+1})\} \cup \{(x_j, y_b) : a + 1 \leq j \leq k\})$$

---

[14]A North-East (NE) lattice path on $\mathbb{Z}^2$ is a sequence of points $(\gamma_i)_{i=1}^m \subset \mathbb{Z}^2$ such that $\gamma_{i+1} - \gamma_i \in \{(1, 0), (0, 1)\}$ for every $i \in \{1, ..., m - 1\}$.

or in

$$\iota^{-1}(\{(x_a, y_b)\} \cup \{ (x_j, y_{b+1}) \ : \ a + 1 \leq j \leq k \})$$

and such that $P = X_* \tilde{P}$. *(Remark that, for such values of* XY, *the value of the* X-*projection completely determine the* Y-*projection.)*

Analogously, if $\gamma_{i+1} - \gamma_i = (0, 1)$, *then for every* $P \in \mathcal{2}_Y$ *such that* $\operatorname{supp} P \subset \{ y_j : b \leq j \leq l \}$, *there exists a counting function* $\tilde{P} \in \mathcal{2}_{XY}$ *whose support is contained in*

$$\iota^{-1}(\{(x_{a+1}, y_b)\} \cup \{ (x_a, y_j) \ : \ b + 1 \leq j \leq l \})$$

or in

$$\iota^{-1}(\{(x_a, y_b)\} \cup \{ (x_{a+1}, y_j) \ : \ b + 1 \leq j \leq k \})$$

and such that $P = Y_* \tilde{P}$.

2. *For each* $\gamma_i = (a, b)$ *such that* $a < k$ *and* $b < l$, *there are elements* z_1, z_2, z_3 *in*

$$\iota^{-1}\{ (x_m, y_n) \ : \ a \leq m \leq a + 1 \ and \ b \leq n \leq b + 1 \}$$

such that $[\delta_{z_1}, \delta_{z_2}, \delta_{z_3}]$ *(the convex hull of the corresponding Dirac measures) is a subset of* $\mathcal{2}_{XY}$ *(which in turn is a subset of the simplex* $\mathscr{P}(X)$ *with vertices* $\{\delta_z\}_{z \in \mathscr{E}_{XY}}$).

In particular, the first condition implies (when $a = 1$ or $b = 1$) that $\mathcal{2}_{XY} \to \mathcal{2}_X$ is surjective, and similarly for Y. The product of a observable with itself is always degenerate.

4.5.7. PROPOSITION. *Let* $(\mathbf{S}, \mathscr{E})$ *be a finite information structure,* $\mathcal{2}$ *an adapted probability functor, and* X, Y *two different observables in* $\operatorname{Ob} \mathbf{S}$ *such that* $XY \in \operatorname{Ob} \mathbf{S}$. *Let* f *be a 1-cocycle of type* α, *i.e. an element of* $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathcal{2}))$. *If* XY *is nondegenerate, there exists* $\lambda \in \mathbb{R}$ *such that*

$$f[X] = \lambda S_\alpha[X], \quad f[Y] = \lambda S_\alpha[Y], \quad f[XY] = \lambda S_\alpha[XY].$$

PROOF. As f is a 1-cocycle, it satisfies the two equations derived from (15)

$$Y.f[X] = f[XY] - f[Y] \tag{64}$$
$$X.f[Y] = f[XY] - f[X] \tag{65}$$

and therefore the symmetric equation

$$X.f[Y] - f[Y] = Y.f[X] - f[X]. \tag{66}$$

For a law P, we write

$$\begin{pmatrix} s & t & u & \dots \\ p & q & r & \dots \end{pmatrix}$$

if $P(s) = p$, $P(t) = q$, $P(u) = r$, etc. and the probabilities of the unwritten parts are zero.

Fix enumerations $(x_1, ..., x_k)$ and $(y_1, ..., y_l)$ that satisfy the definition of nondegenerate product, and let $\{\gamma_i\}_{i=1}^m$ be the corresponding NE path. Write $\gamma_i = (a, b)$. If $\gamma_{i+1} - \gamma_i = (1, 0)$, we shall show that the following recursive formula holds:

$$f[X] \begin{pmatrix} x_a & \cdots & x_k \\ \mu_a & \cdots & \mu_k \end{pmatrix} = (1 - \mu_a)^\alpha f[X] \begin{pmatrix} x_{a+1} & \cdots & x_k \\ \mu_{a+1}/(1 - \mu_a) & \cdots & \mu_k/(1 - \mu_a) \end{pmatrix}$$
$$+ f[X] \begin{pmatrix} x_a & x_{a+1} \\ \mu_a & 1 - \mu_a \end{pmatrix}. \quad (67)$$

Analogously, if $\gamma_{i+1} - \gamma_i = (0, 1)$,

$$f[Y] \begin{pmatrix} y_b & \cdots & y_l \\ \nu_b & \cdots & \nu_l \end{pmatrix} = (1 - \nu_b)^\alpha f[Y] \begin{pmatrix} y_{b+1} & \cdots & y_l \\ \nu_{b+1}/(1 - \nu_c) & \cdots & \nu_l/(1 - \nu_b) \end{pmatrix}$$
$$+ f[Y] \begin{pmatrix} y_b & y_{b+1} \\ \nu_b & 1 - \nu_b \end{pmatrix}. \quad (68)$$

Suppose that $\gamma_{i+1} - \gamma_i = (1, 0)$. Let

$$p = \begin{pmatrix} x_a & \cdots & x_k \\ \mu_a & \cdots & \mu_k \end{pmatrix}$$

be a probability in $\mathcal{Q}_X$ supported on $\{ x_i : a \leq i \leq k \}$. We know it has a preimage $\tilde{p}$ under marginalization X_* as in Definition 4.5.6-(1). For such law, knowledge of X implies knowledge of Y with certainty, therefore $X.f[Y](\tilde{p}) = 0$; by equation (65), $f[XY](\tilde{p}) = f[X](X_*\tilde{p}) = f[X](p)$. Equation (64) then reads

$$(1 - \mu_a)^\alpha f[X] \begin{pmatrix} x_{a+1} & \cdots & x_k \\ \mu_{a+1}/(1 - \mu_a) & \cdots & \mu_k/(1 - \mu_a) \end{pmatrix} =$$
$$f[X] \begin{pmatrix} x_a & \cdots & x_k \\ \mu_a & \cdots & \mu_k \end{pmatrix} - f[Y] \circ \tau \begin{pmatrix} y_b & y_{b+1} \\ 1 - \mu_a & \mu_a \end{pmatrix}, \quad (69)$$

where τ is the identity or the transposition of the arguments. In any case, we can set $\mu_{a+1} = 1 - \mu_a$ and $\mu_{a+2} = \ldots = \mu_k = 0$ (because this probability appears as the image of some of the probabilities on XY given by Definition 4.5.6-(2)), to conclude that

$$f[X] \begin{pmatrix} x_a & x_{a+1} \\ \mu_a & 1 - \mu_a \end{pmatrix} = f[Y] \circ \tau \begin{pmatrix} y_b & y_{b+1} \\ 1 - \mu_a & \mu_a \end{pmatrix}, \quad (70)$$

which combined with (69) implies (67). The identity (68) can be obtained analogously.

We proceed to the determination of

$$\phi_a(z) := f[X] \begin{pmatrix} x_a & x_{a+1} \\ z & 1 - z \end{pmatrix} \quad \text{and} \quad \psi_b(z) := f[Y] \begin{pmatrix} y_b & y_{b+1} \\ z & 1 - z \end{pmatrix}, \quad \text{for } z \in [0, 1].$$

Let z_1, z_2, z_3 be the three elements of $\mathscr{E}_{XY} \subset \mathscr{E}_X \times \mathscr{E}_Y$ such that $[\delta_{z_1}, \delta_{z_2}, \delta_{z_3}]$ is the subset of $\mathscr{Q}_{XY}$ given by Definition 4.5.6-(2) when $\gamma_i = (a, b)$. For any $\mu = (\mu_0, \mu_X, \mu_Y) \in \Delta^2$, set $P(z_i) := \mu_Y$, where z_i is the component that differs from the others w.r.t. the Y-projection, in such a way that $Y_*\mu$ is $(\mu_Y, 1 - \mu_Y)$. Similarly, set $P(z_j) := \mu_X$, where z_j is the component that differs from the others w.r.t. the X-projection. With this assignment, (66) reads

$$(1 - \mu_X)^\alpha f[Y] \circ \sigma \begin{pmatrix} y_b & y_{b+1} \\ \mu_0/(1 - \mu_X) & \mu_Y/(1 - \mu_X) \end{pmatrix} - f[Y] \circ \sigma \begin{pmatrix} y_b & y_{b+1} \\ 1 - \mu_Y & \mu_Y \end{pmatrix}$$
$$= (1 - \mu_Y)^\alpha f[X] \circ \tau \begin{pmatrix} x_a & x_{a+1} \\ \mu_0/(1 - \mu_Y) & \mu_Y/(1 - \mu_Y) \end{pmatrix} - f[X] \circ \tau \begin{pmatrix} x_a & x_{a+1} \\ 1 - \mu_X & \mu_X \end{pmatrix},$$
$$\tag{71}$$

where σ, τ are the identity or the transposition of both nontrivial arguments. In any case, this leads to the functional equation in Proposition 4.5.4, hence $\phi_a(z) = \psi_b(z) = \lambda s_\alpha(z)$ for certain $\lambda \in \mathbb{R}$ (the solution is symmetric in the arguments).

When considering γ_{i+1}, one finds the functions ϕ_a and ψ_{b+1}, or the functions ϕ_{a+1} and ψ_b, since the difference $\gamma_{i+1} - \gamma_i$ is either $(0, 1)$ or $(1, 0)$. This ensures that the constant λ that appears for each γ_i is always the same.

Repeat the process above with every γ_i $(1 \leq i \leq m)$. The system of equations (67) obtained in this way, together with the functions already determined

$$f[X] \begin{pmatrix} x_a & x_{a+1} \\ z & 1 - z \end{pmatrix} = \lambda s_\alpha(z, 1 - z) \quad \text{for } 1 \leq a \leq k - 1,$$

entails that

$$f[X](\mu_1, \dots, \mu_k) = \lambda \sum_{i=0}^{k-1} \left(1 - \sum_{j=1}^{i} \mu_j\right)^\alpha s_\alpha \left(\frac{\mu_{i+1}}{\left(1 - \sum_{j=1}^{i} \mu_j\right)}\right). \tag{72}$$

Set $T_i := 1 - \sum_{j=1}^{i} \mu_j$. An elementary computation shows that, when $\alpha = 1$,

$$\sum_{i=0}^{k-1} (1 - T_i) s_1 \left(\frac{\mu_{i+1}}{(1 - T_i)}\right) = \sum_{i=1}^{k} \mu_i \log \mu_i, \tag{73}$$

and when $\alpha \neq 1$,

$$\sum_{i=0}^{k-1} (1 - T_i)^\alpha s_\alpha \left(\frac{\mu_{i+1}}{(1 - T_i)}\right) = \sum_{i=1}^{k} \mu_i^\alpha - 1. \tag{74}$$

Therefore, for any $\alpha > 0$, we have $f[X] = \lambda S_\alpha[X]$. Analogously, $f[Y] = \lambda S_\alpha[Y]$. $\blacksquare$

4.6. DETERMINATION OF H^1. In this section, we prove Theorem 1.3.2, that determines completely H^1 in the probabilistic case under suitable hypotheses. We also discuss some pathological cases.

We call a observable Z *reducible* if there exist $X, Y \in \mathrm{Ob}\,\mathbf{S} \setminus \{1, Z\}$ such that $Z = XY$, and *irreducible* otherwise. It is *nontrivialy reducible* if the product XY is nondegenerate.

PROOF OF THEOREM 1.3.2. Let f be an element of $Z^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q}))$. We are going to determine $f_X[Y]$ for every $X \in \mathrm{Ob}\,\mathbf{S}$ and $Y \in \mathscr{S}_X$.

If X satisfies $|\mathscr{E}_X| = 1$, then $|\mathscr{E}_Y| = 1$ for any $Y \in \mathscr{S}_X$. We conclude that $f_X[Y] = 0$ in virtue of Proposition 4.5.1.

If X satisfies $|\mathscr{E}_X| > 1$, then it belongs to $\mathbf{S}^*$. Let $\mathbf{C}(X)$ denote the connected component of $\mathbf{S}^*$ that contains X. By hypothesis, there exists a nontrivially reducible object Z such that $Z \to X \to Y$. Then,

$$f[Y](Z_*P) = f_Z[Y](P) = f_Z[Z](P) - Y.f_Z[Z](P) = \begin{cases} 0 & \text{if } |\mathscr{E}_Y| = 1 \\ \lambda_Z S_\alpha[Y](Z_*P) & \text{if } |\mathscr{E}_Y| > 1 \end{cases}, \quad (75)$$

because $f_Z[Z]$ equals $\lambda_Z S_\alpha[Z]$, for some $\lambda_Z \in \mathbb{R}$, according to Proposition 4.5.7. Since Z_* is supposed to be surjective, this fully determines $f[Y]$. By naturality, $f_X[Y]$ is the map $P \mapsto f[Y](X_*P)$. Remark that Y also belongs to $\mathbf{C}(X)$; if Z' is another nontrivially reducible observable such that $Z' \to Y$ then $Z \in \mathbf{C}(X)$ too. Repeating the previous argument for Z', we conclude that $f[Y] = \lambda_Z S_\alpha[Y] = \lambda_{Z'} S_\alpha[Y]$, therefore $\lambda_Z = \lambda_{Z'} =: \lambda_{\mathbf{C}(X)}$.

More generally, if Z, Z' are any two nontrivially reducible objects in the same component $\mathbf{C}$ of $\mathbf{S}^*$, then there is a zig-zag diagram in $\mathbf{S}^*$ of the form

$$Z \to Y_1 \leftarrow X_1 \to Y_2 \leftarrow X_2 \to \cdots \leftarrow X_k \to Y_{k+1} \leftarrow Z'$$

for certain $k \in \mathbb{N}$, where $Y_i, X_i \in \mathrm{Ob}\,\mathbf{S}^*$. By hypothesis, each X_i is refined by a nontrivially reducible object Z_i, in such a way that we also have a diagram

$$Z \to Y_1 \leftarrow Z_1 \to Y_2 \leftarrow Z_2 \to \cdots \leftarrow Z_k \to Y_{k+1} \leftarrow Z'$$

in $\mathbf{S}^*$. The repeated application of the argument in the previous paragraph implies that $\lambda_Z = \lambda_{Z_1} = \cdots = \lambda_{Z'} =: \lambda_{\mathbf{C}}$.

Summarizing, f is such that $f_X[Y] = 0$ if $|\mathscr{E}_Y| = 1$ and $f_X[Y](P) = \lambda_{\mathbf{C}} S_\alpha[Y](X_*P)$ if Y belongs to a connected component $\mathbf{C}$ of $\mathbf{S}^*$ (hence X too). Therefore, there is a linear bijective correspondence between cocycles f and sequences of $(\lambda_{\mathbf{C}})_{\mathbf{C}} \in \mathbb{R}^{\pi_0(\mathbf{S}^*)}$. (The set $\mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_1, \mathscr{F}_\alpha)$ has a vector space structure given by "object-wise" addition: for every $\phi, \psi \in \mathrm{Hom}_{\mathscr{A}}(\mathscr{B}_1, \mathscr{F}_\alpha)$, $\alpha, \beta \in \mathbb{R}$, $X \in \mathrm{Ob}\,\mathbf{S}$, and $Y \in \mathscr{S}_X$, one has $(\alpha\phi + \beta\psi)_X[Y] = \alpha\phi_X[Y] + \beta\psi_X[Y]$.)

We saw in Section 4.5 that $\delta C^0(\mathbf{S}, \mathscr{F}_1(\mathscr{Q})) \cong \langle 0 \rangle$ and $\delta C^0(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{Q})) = \{\lambda S_\alpha : \lambda \in \mathbb{R}\}$. ∎

As a byproduct of the previous proof, we also obtain the following proposition.

4.6.1. PROPOSITION. *Let $\{(\mathbf{S}_i, \mathscr{E}_i, \mathscr{Q}_i)\}_{i \in I}$ be a collection of triples that satisfy separately the hypotheses stated in Theorem 1.3.2. Let $\bigsqcup_{i \in I} \mathscr{Q}$ denote the adapted functor of probabilities on $\bigsqcup_{i \in I} \mathbf{S}$ that coincides with $\mathscr{Q}_i$ on $\mathbf{S}_i$. Then,*

$$Z^1 \left(\bigsqcup_{i \in I} \mathbf{S}, \mathscr{F}(\bigsqcup_{i \in I} \mathscr{Q}) \right) \cong \prod_{i \in I} Z^1(\mathbf{S}_i, \mathscr{F}(\mathscr{Q}_i)).$$

PROOF. The category $(\bigsqcup_{i=1}^n \mathbf{S})^*$ is the disjoint union of the categories $\mathbf{S}_i^*$, for $i \in I$.   ∎

4.6.2. DEGENERATE CASES. Let us focus now on the particular case of bounded, finite information structures $(\mathbf{S}, \mathscr{E})$, with $\mathscr{E}$ conservative (e.g. obtained from a concrete structure) and $\mathscr{Q} = \mathscr{P}$. The hypothesis of Theorem 1.3.2 is satisfied if and only if every minimal object of the poset $\mathbf{S}$ is nontrivially reducible. Hence there are two kind of cases not covered by it:

1. there is an irreducible minimal object;

2. all minimal objects are reducible, but some of them cannot be written as nondegenerate products.

In the latter, all kinds of behaviors are possible, as the examples at the end this section show.

In the Example 4.5.2, we proved that $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$ has infinite dimension when $\mathbf{S} \cong \mathbf{Obs}_{\mathrm{fin}}(\{0, 1\})$; in this case, there is only one nontrivial observable and it is obviously irreducible. Now we proceed to the generalization of this result.

4.6.3. PROPOSITION. *Let $(\mathbf{S}, \mathscr{E})$ be a bounded, finite information structure such that $\mathscr{E}$ is conservative. Suppose $\mathbf{S}$ has an irreducible minimal object. Then, $\dim H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P})) = \infty$.*

PROOF. Let M be an irreducible minimal object. Remark that

$$T := \{ X \ : \ M \to X \text{ and } X \neq M \}$$

has a unique minimal element $\tilde{X}$ for the partial order $\mathbf{S}$ (if X_1 and X_2 were two different minimal elements, the diagram $X_1 \leftarrow M \to X_2$ would imply that $M = X_1 X_2$).

Recall that a 1-cochain f is uniquely determined by a collection of measurable functions $\{f[X] : \mathscr{Q}_X \to \mathbb{R}\}_{X \in \mathrm{Ob}\,\mathbf{S}}$. Let us set $f[X] = 0$ for every $X \in \mathrm{Ob}\,\mathbf{S} \setminus \{M\}$ and show that $f[M]$ can be chosen arbitrarily so that f is a 1-cocycle.

The 1-cocycle condition implies in particular that

$$f[M](P) = \tilde{X}.f[M](P) := \sum_{x \in \mathscr{E}_{\tilde{X}}} \tilde{X}_* P(x)^\alpha f[M](P|_{\tilde{X}=x}). \tag{76}$$

Given this functional equation, the others become redundant, since for any X coarser than M (hence coarser than $\tilde{X}$),

$$X.f[M] = X.(\tilde{X}.f[M]) = (X\tilde{X}).f[M] = \tilde{X}.f[M].$$

Since $\mathscr{E}$ is conservative, the induced surjection $\mathscr{E}\pi : \mathscr{E}_M \to \mathscr{E}_{\tilde{X}}$ is not a bijection. Therefore, the set $\mathscr{E}^* := \{ x \in \mathscr{E}_{\tilde{X}} : |(\mathscr{E}\pi_{\tilde{X}M})^{-1}(x)| \geq 2 \}$ is nonempty. For each element of $x \in \mathscr{E}^*$, one can introduce an arbitrary function g_x on $\Delta^{|(\mathscr{E}\pi_{\tilde{X}M})^{-1}(x)|-1}$ that vanishes on the vertices, so that $f[M](P|_{X=x}) = g_x(P|_{X=x})$ and the value of M on an arbitrary $P \in \mathscr{P}_X$ is given by (76). ∎

To close this section, we make some remarks about the case of reducible minimal objects that cannot be written as nondegenerate products. If the product is degenerate, multiple constants can appear or the dimension of $H^1(\mathbf{S}, \mathscr{F}_\alpha(\mathscr{P}))$ can become infinite, as the following examples show.

4.6.4. EXAMPLE. Consider the information structure $(\mathbf{S}, \mathscr{E})$ given by the poset $\mathbf{S}$ represented by
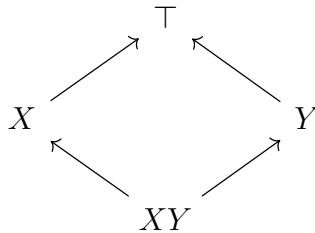
and the assignment $\mathscr{E}(X) = \{x_1, x_2, x_3, x_4\}$, $\mathscr{E}(Y) = \{y_1, y_2, y_3, y_4\}$, and

$$\mathscr{E}(XY) = (\{x_1, x_2\} \times \{y_1, y_2\}) \cup (\{x_3, x_4\} \times \{y_3, y_4\});$$

the surjections are the terminal maps and the restrictions of the canonical projectors. To determine a 1-cochain $f \in C^1(\mathbf{S}, \mathscr{F}(\mathscr{P}))$ it suffices to specify

$$f[X] \begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix}, f[X] \begin{pmatrix} x_3 & x_4 \\ p & 1-p \end{pmatrix}, f[Y] \begin{pmatrix} y_1 & y_2 \\ p & 1-p \end{pmatrix} \text{ and } f[Y] \begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix},$$

for arbitrary $p \in [0, 1]$. Proposition 4.5.4 allows us to conclude that

$$f[X] \begin{pmatrix} x_1 & x_2 \\ p & 1-p \end{pmatrix} = \lambda_1 s_\alpha(p), \quad f[Y] \begin{pmatrix} y_1 & y_2 \\ p & 1-p \end{pmatrix} = \lambda_1 s_\alpha(p). \tag{77}$$

We can use this proposition a second time to show that

$$f[X] \begin{pmatrix} x_3 & x_4 \\ p & 1-p \end{pmatrix} = \lambda_2 s_\alpha(p), \quad f[Y] \begin{pmatrix} y_3 & y_4 \\ p & 1-p \end{pmatrix} = \lambda_2 s_\alpha(p). \tag{78}$$

However, it is impossible to find a relation between λ_1 and λ_2 using (64) and (65). So $Z^1(\mathscr{F}_\alpha(\mathscr{Q})) \cong \mathbb{R}^2$.

4.6.5. EXAMPLE. If in the previous example

$$\mathscr{E}_{XY} = (\{x_1, x_2\} \times \{y_1, y_2\}) \cup \{(x_3, y_3), (x_4, y_4)\}, \tag{79}$$

it is still true that

$$f[X]\left(\begin{array}{cc} x_1 & x_2 \\ p & 1-p \end{array}\right) = \lambda_1 s_\alpha(p), \quad f[Y]\left(\begin{array}{cc} y_1 & y_2 \\ p & 1-p \end{array}\right) = \lambda_1 s_\alpha(p), \tag{80}$$

whereas (64) and (65) imply that

$$f[XY]\left(\begin{array}{cc} (x_3, y_3) & (x_4, y_4) \\ p & 1-p \end{array}\right) = f[X]\left(\begin{array}{cc} x_3 & y_3 \\ p & 1-p \end{array}\right) = f[Y]\left(\begin{array}{cc} y_3 & y_4 \\ p & 1-p \end{array}\right). \tag{81}$$

We conclude that any measurable $g : \Delta^1 \to \mathbb{R}$ such that $g(0, 1) = g(1, 0) = 0$ defines a solution, therefore $\dim Z^1(\mathscr{F}_\alpha(\mathscr{Q})) = \infty$.

4.7. INTERPRETATION: CROSSED HOMOMORPHISMS. The extension of the results in [Mac Lane, 1994, Ch. X] concerning the Hochschild cohomology of algebras to the case of presheaves of algebras is straightforward, taking into account that the relative bar construction in the appendix gives the explicit description of both.

For instance, $H^0(\mathbf{S}, \mathscr{N})$ corresponds to global sections of $\mathscr{N}$ that are invariant under the action of $\mathscr{A}$. When $\mathscr{N} = \mathscr{F}_\alpha(\mathscr{Q})$ the only sections are constant functionals; they are all invariant when $\alpha = 1$, and none of them is invariant when $\alpha \neq 1$.

A *crossed homomorphism* of $\mathscr{A}$ to $\mathscr{N}$ is a morphism of presheaves of $\mathbb{R}$-modules $f : \mathscr{A} \to \mathscr{N}$ satisfying the identity

$$f[X_1 X_2] = X_1 f[X_2] + f[X_1]. \tag{82}$$

The principal crossed homomorphisms have the form $f_n[X] = Xn - nX = Xn - n$ for some section n of $\mathscr{N}$. Therefore $H^1(\mathbf{S}, \mathscr{N})$ is the $\mathbb{R}$-module of crossed homomorphisms modulo the principal ones.

When $\mathscr{N} = \mathscr{F}_\alpha(\mathscr{Q})$ and the hypotheses in Theorem 1.3.2 are satisfied, the *only* crossed homomorphism is the corresponding α-entropy (up to multiplicative constants). Thus each entropy represents the unique nontrivial way of transforming a multiplicative operation involving partitions or σ-algebras into an additive operation on probabilistic functionals, introducing an appropriate "twist". In the degenerate case of Proposition 4.6.3, the algebra $\mathscr{A}_M$ is "too poor" to determine a unique crossed homomorphism. Moreover, the naturality of S_α implies that $S_\alpha[X](P)$ depends only on X_*P. This turns out to be the appropriate notion of locality and justifies the introduction of presheaves, and in this sense one could consider information cohomology as a "localized" version of Hochschild cohomology.

Finally, $H^2(\mathbf{S}, \mathscr{N})$ classifies $\mathbb{R}$-split singular algebra extensions of $\mathscr{N}$ by $\mathscr{A}$, see [Mac Lane, 1994, Thm. 3.1].

5. Final remarks

We have seen that Shannon entropy can be identified with a cohomology class in information cohomology, a topological invariant associated to a finite statistical system. It determines the "natural" way to turn the product of partitions/σ-algebras into an additive operation of probabilistic functionals that is compatible with variations of the probability laws.In this framework, each entropy S_α spontaneously appears indexed by the observables—via the bar construction—, and it is "local" in the sense that $S_\alpha[X](P)$ only depends on X_*P, the marginalization of the probability P on $\mathscr{E}_X$. This good notion of locality— encoded by the naturality of the involved functors—and the 1-cocycle condition replace completely the axioms invoked by the usual algebraic approaches (one does not need to suppose symmetry, convexity/concavity or a given asymptotic behavior).

Whereas the fact that Shannon or Tsallis entropies are 1-cocycles is related to well-known facts (because the 1-cocycle condition is the corresponding chain rule), the cohomological restatement gives new insights: it is possible to change the coefficients of the cohomology and interpret the same 1-cocycle condition for completely different objects. In [Vigneaux, 2019], it is showed that when the coefficients are certain "combinatorial" functionals ($\mathbb{R}_{>0}$-valued functions of statistical frequencies), the 1-cocycles are generalized multinomial coefficients and the chain rule corresponds to the multiplicative relations between them. The combinatorial and probabilistic cocycles are asymptotically related. The same reference introduces information structures of continuous observables with gaussian laws (cf. Example 2.1.2): in this setting, the dimension of the support appears as a 1-cocycle, the 1-cocycle condition being the nullity-rank theorem; similarly, the log of the determinant of the covariance matrix is a 1-cocycle, and the 1-cocycle condition is Schur's determinantal formula. It would be difficult to recognize all these identities as manifestations of the same thing without the general framework. The isolation of the common combinatorial structure in all these examples (the conditional meet semilattices) is instrumental to develop a definition of information cohomology that works in general.

It seems to us that the *connections* between the different points of view on entropy—algebraic, probabilistic, combinatorial, dynamical—are not yet systematically understood. We are convinced that the language of (pre)sheaves and their homological invariants opens many new directions of research, being rich enough to integrate constructions coming from these diverse domains. Further applications of our approach are related to several open problems, for instance:

1. the computation of cocycles of higher degrees, conjectured in [Baudot and Bennequin, 2015] to be new measures of mutual information of all orders;

2. the possible reformulation of Shannon's coding theorems as cohomological obstruction problems;

3. the computation of cohomology for categories of symplectic manifolds and reductions, giving adapted measures of information;

4. a functorial relation between classical and quantum information cohomology (and the corresponding concentration theorems) through geometric quantization, and

5. a categorical formulation of *Ruzsa's dictionary* [Ruzsa, 2009], that relates inequalities for cardinalities and entropies.

Only the first steps are made: we hope that other authors, with new ideas, will be able to go further.

A. Relative bar resolution

A.1. GENERAL RESULTS. In this subsection, we summarize the construction of a relative bar resolution from [Mac Lane, 1994, Ch. IX]. The purpose is to find the analogue of a free resolution of modules, but in the general context of abelian categories. Capital Latin letters $A, B, C...$ denote objects and Greek letters $\alpha, \beta...$ morphisms.

A *relative abelian category* is a pair of abelian categories **A** and **M** and a covariant functor $\Box : \mathbf{A} \to \mathbf{M}$ which is additive, exact and faithful (we write $\Box(X) = X_\Box$, for objects and morphisms).

A.1.1. EXAMPLE. The simple example to have in mind are R-modules and S-modules, when S is a subring of R with the same unit (write $\iota : S \to R$ for the injection). Every R-module A can be seen as an S-module $_\iota A$ by *restriction of scalars*; every R-module morphism $\alpha : A \to B$ is also a S-module morphism $_\iota \alpha : {_\iota A} \to {_\iota B}$. The assignment $\Box A := {_\iota A}$ and $\Box \alpha := {_\iota \alpha}$ defines a relative abelian category.

A short exact sequence $\chi \| \sigma$ in **A** is relatively split ($\Box$-split) if $\chi_\Box \| \sigma_\Box$ splits in **M**. A monomorphism χ is called allowable if $\chi \| \sigma$ is $\Box$-split for some σ; this is the case if and only if $\chi \| (\operatorname{coker} \chi)$ is $\Box$-split. Dually, an epimorphism is called allowable if $(\ker \sigma) \| \sigma$ is $\Box$-split. The following conditions on a morphism α are equivalent:

1. $\operatorname{im} \alpha$ is an allowable monomorphism and $\operatorname{coim} \alpha$ is an allowable epimorphism;

2. $\ker \alpha$ is an allowable monomorphism and $\operatorname{coker} \alpha$ is an allowable epimorphism;

A morphism is called allowable when it satisfies any of these conditions (see [Mac Lane, 1994, p. 264]).

A *relative projective object* P is any object of **A** such that, for every *allowable* epimorphism $\sigma : B \to C$, each morphism $\epsilon : P \to C$ of **A** can be factored through σ as $\epsilon = \sigma \epsilon'$ for some $\epsilon' : P \to A$.

A *resolvent pair* is a relative abelian category $\Box : \mathbf{A} \to \mathbf{M}$ together with a covariant functor $F : \mathbf{M} \to \mathbf{A}$ left adjoint to $\Box$.

A.1.2. PROPOSITION. *Let* $\square : \mathbf{A} \to \mathbf{M}$ *be a relative abelian category. The following conditions are equivalent:*

1. *there exists a covariant functor* $F : \mathbf{M} \to \mathbf{A}$ *left adjoint to* $\square$;

2. *there exist a covariant functor* $F : \mathbf{M} \to \mathbf{A}$, *and a natural transformation* $e : 1_{\mathbf{M}} \to \square F$ *(where* $1_{\mathbf{M}}$ *is the identity functor), such that every* $u : M \to A_{\square}$ *in* $\mathbf{M}$ *has a factorization* $u = \alpha_{\square} e_M$, *with* $\alpha : F(M) \to A$ *unique.*

Remark that e is the unit of the adjunction.

A.1.3. EXAMPLE. [continuation of A.1.1] Take $F(M) = R \otimes_S M$ and $e_M = 1 \otimes m \in F(M)$. Given a map of S-modules $u : M \to A_{\square}$, define $\alpha : FM \to A$ by $\alpha(1 \otimes m) = u(m)$.

A complex $\epsilon : X \to A$ over A (in $\mathbf{A}$) is a sequence of $\mathbf{A}$-objects and $\mathbf{A}$-morphisms $...X_n \to X_{n-1} \to ... \to X_1 \to X_0 \xrightarrow{\epsilon} A \to 0$, such that the composite of any two successive morphisms is zero. This complex is called a *resolution* of C if the sequence is exact, *relatively free* if each X_n has the form $F(M_n)$ for certain M_n in $\mathbf{M}$, and *allowable* if all its morphisms are allowable.

Each object C of $\mathbf{A}$ has a canonical relatively free resolution. Writing $\tilde{F}C$ for $F\square C$, and $\tilde{F}^n$ for its n-fold iteration, construct the objects $B_n(C) = \tilde{F}^{n+1}C$, for each $n \in \mathbb{N}$. Define $\mathbf{M}$-morphisms $s_{\bullet}$ between the corresponding objects

$$\square C \xrightarrow{s_{-1}} \square B_0(C) \xrightarrow{s_0} \square B_1(C) \xrightarrow{s_1} \square B_2(C) \xrightarrow{s_2} \ldots \tag{83}$$

as $s_{-1} := e(\square C)$ and $s_n := e(\square B_n(C))$ (here e is the natural transformation in Proposition A.1.2).

A.1.4. PROPOSITION. [see [Mac Lane, 1994, p. 268]] *There are unique* $\mathbf{A}$-morphisms

$$\epsilon : B_0(C) \to C, \quad \partial_{n+1} : B_{n+1}(C) \to B_n(C) \quad \text{for } n \in \mathbb{N},$$

which make $B(C) := \{B_n(C)\}_n$ *a relatively free allowable resolution of* C *with* s *as contracting homotopy in* $\mathbf{M}$. *This resolution, with its contracting homotopy, is a covariant functor of* C.

PROOF. We simply quote here the construction of ϵ and ∂_n. They form the following diagram (solid arrows belong to $\mathbf{A}$, and dashed arrows belong to $\mathbf{M}$):

$$0 \longleftarrow C \underset{s_{-1}}{\overset{\epsilon}{\longleftrightarrow}} B_0(C) \underset{s_0}{\overset{\partial_1}{\longleftrightarrow}} B_1(C) \underset{s_1}{\overset{\partial_2}{\longleftrightarrow}} B_2(C) \underset{s_2}{\overset{\partial}{\longleftrightarrow}} \ldots \tag{84}$$

By Proposition A.1.2, $1_{\square C}$ factors through a unique $\epsilon : B_0(C) \to C$; the formula $1_{\square C} = \epsilon_{\square} e_C$ shows that ϵ is allowable (note that ϵ is an epimorphism). Boundary operators are defined by recursion so that s will be a contracting homotopy. Given ϵ, the morphism $1_{\square B_0} - s_{-1}\epsilon_{\square}$ factors uniquely as $\partial_{1\square} s_0$, for some $\partial_1 : B_1(C) \to B_0(C)$. Similarly, $1_{\square B_n} -$

$s_{n-1}\partial_{n\square} : \square B_n(C) \to \square B_n(C)$ determines ∂_{n+1} given ∂_n, as the unique **A**-morphism such that $\partial_{n+1\square} s_n = 1_{\square B_n} - s_{n-1}\partial_{n\square}$.

$$\begin{array}{ccc} \square B_{n+1} & \xrightarrow{\ \partial_{1\square}\ } & \square B_n \\ {\scriptstyle s_n}\big\uparrow & \nearrow{\scriptstyle 1_{\square B_n}} & \\ \square B_n & & \end{array} \qquad (85)$$

∎

The resolution $B(C)$ is called the (unnormalized) bar resolution.

A.2. EXAMPLE: PRESHEAVES OF MODULES. We develop now the particular case relevant to our theory. Let **S** be a category, and $\mathscr{R}, \mathscr{T} : \mathbf{S}^{\mathrm{op}} \to \mathbf{Rings}$ presheaves, such that $\mathscr{T}_X$ is a subring of $\mathscr{R}_X$ with the same unit, for every $X \in \mathrm{Ob}\,T$. Take $\mathbf{A} = \mathbf{Mod}(\mathscr{R})$, the category of presheaves of $\mathscr{R}$-modules, and $\mathbf{M} = \mathbf{Mod}(\mathscr{T})$, the category of presheaves of $\mathscr{T}$-modules. A relative abelian category is obtained when $\square : \mathbf{A} \to \mathbf{M}$ is the forgetful functor over each X, as defined in Example A.1.1. The functor $F : \mathbf{M} \to \mathbf{A}$ sends a presheaf $\mathscr{P}$ to the new presheaf $X \mapsto \mathscr{R}_X \otimes_{\mathscr{T}_X} \mathscr{P}_X$,[15] and each morphism of $\mathscr{T}$-presheaves (in short, $\mathscr{T}$-morphism) $f : M \to N$ to the $\mathscr{R}$-morphism defined by

$$\forall X \in \mathrm{Ob}\,\mathbf{S}, \forall m \in M(X), \quad Ff(X)(1 \otimes m) = 1 \otimes f(m), \quad \text{for } X \in \mathbf{S}. \qquad (86)$$

The natural transformation e mentioned in Proposition A.1.2 corresponds to a collection of $\mathscr{T}$-morphisms $e_{\mathscr{P}} : \mathscr{P} \to \square F(\mathscr{P})$, one for each presheaf $\mathscr{P}$ of $\mathscr{T}$-modules; given X in **S**, we define $e_{\mathscr{P}}(X)(m) = 1 \otimes m$ for each $m \in M(X)$. [16]

Fix now a presheaf $\mathscr{C}$ in $\mathbf{Mod}(\mathscr{R})$. We denote by X a generic element in $\mathrm{Ob}\,\mathbf{S}$. Then, $B_0\mathscr{C}(X) := F\square\mathscr{C}(X) = \mathscr{R}_X \otimes_{\mathscr{T}_X} (\square\mathscr{C}(X))$; this $\mathscr{R}_X$-module is formed by finite $\mathscr{R}_X$-linear combinations of tensors $1 \otimes c$, with $c \in \mathscr{C}$. Generally, an element of $B_n\mathscr{C}(X) = \mathscr{R}_X \otimes \square B_{n-1}\mathscr{C}(X)$, for $n \geq 1$, is a finite $\mathscr{R}_X$-linear combination of tensors $1 \otimes r_1 \otimes r_2 \otimes ... \otimes r_n \otimes c$. The ring $\mathscr{R}_X$ acts on $B_n\mathscr{C}(X)$ by multiplication on the first factor of the tensor product; to highlight this fact, people usually write $r[r_1|r_2|...|r_n|c]$ instead of $r \otimes r_1 \otimes r_2 \otimes ... \otimes r_n \otimes c$. This notation explains the name "bar resolution" adopted above. The definition of e implies that

$$s_{-1}^X : \square\mathscr{C}(X) \to \square B_0\mathscr{C}(X), \quad c \mapsto 1 \otimes c = [c], \qquad (87)$$

and

$$s_n^X : \square B_n\mathscr{C}(X) \to \square B_{n+1}\mathscr{C}(X), \quad r[r_1|r_2|...|r_n|c] \mapsto [r|r_1|r_2|...|r_n|c] \quad \text{for } n \in \mathbb{N}. \qquad (88)$$

---

[15]This is a left $\mathscr{R}$-module with action defined by $r(r' \otimes p) = (rr') \otimes g$. For iterated tensor products, this definition is not canonical; for example, when considering $\mathscr{R}_X \otimes \mathscr{R}_X \otimes \mathscr{P}_X$, the element $(sr) \otimes r' \otimes g$ does not equal $r \otimes (sr') \otimes g$ (for $s \in \mathscr{T}$), unless $\mathscr{T}$ is in the center of $\mathscr{R}$. As in this work we only use commutative rings and algebras, these differences do not pose any problem.

[16]Of course, one has to prove that e is in fact a natural transformation and satisfies the properties required by Proposition A.1.2. This proof is rather trivial but complicated to write, and we omit it.

These equalities determine $s_\bullet$, since these functions are $\mathcal{T}_X$-linear.

Now ϵ is the *unique* $\mathcal{R}$-morphism such that $1_{\Box\mathscr{C}} = \epsilon_\Box e_\mathscr{C}$; this is clearly the case if $\epsilon^X([c]) = c$. Similarly, ∂_1 is the unique $\mathcal{R}$-morphism from $B_1\mathscr{C}$ to $B_0\mathscr{C}$ that satisfies

$$\partial_{1\Box}s_0 = 1 - s^X_{-1}\epsilon_\Box \tag{89}$$

Since $B_1\mathscr{C}(X)$ is generated as a $\mathcal{R}_X$-module by the elements $[r|c]$, and $s^X_0(r[c]) = [r|c]$, the equation (89) defines ∂_1 completely. Just remark that $\epsilon(r[c]) = r\epsilon([c]) = r\epsilon(1\otimes c) = rc$ and $s_1(rc) = [rc]$. We conclude that

$$\partial_1([r|c]) = r[c] - [rc]. \tag{90}$$

It can be proved by recursion that (cf. [Mac Lane, 1994, p. 281])

$$\partial[r_1|...|r_n|c] = r_1[r_2|...|r_n|c] + \sum_{k=1}^{n-1}(-1)^k[r_1|...|r_kr_{k+1}|...|r_n|c] + (-1)^n[r_1|...|r_{n-1}|r_nc]. \tag{91}$$

In virtue of Proposition A.1.4, we obtain in this way a free allowable resolution of $\mathscr{C}$.

References

S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, 2011.

S. Abramsky and L. Hardy. Logical Bell inequalities. *Phys. Rev. A*, 85:062114, Jun 2012.

S. Abramsky, R. S. Barbosa, K. Kishida, R. Lal, and S. Mansfield. Contextuality, Cohomology and Paradox. In S. Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 211–228, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

J. Aczél and Z. Daróczy. *On Measures of Information and Their Characterizations*. Mathematics in Science and Engineering. Academic Press, 1975. ISBN 9780120437603.

M. Artin, A. Grothendieck, and J.-L. Verdier. *Théorie des Topos et Cohomologie Étale des Schémas: Séminaire de Géométrie Algébrique du Bois-Marie 1963/64 - SGA 4. Tome 1*. Lecture notes in mathematics. Springer-Verlag, 1972a.

M. Artin, A. Grothendieck, and J.-L. Verdier. *Théorie des Topos et Cohomologie Étale des Schémas: Séminaire de Géométrie Algébrique du Bois-Marie 1963/64 - SGA 4. Tome 2*. Lecture notes in mathematics. Springer-Verlag, 1972b.

J. C. Baez, T. Fritz, and T. Leinster. A characterization of entropy in terms of information loss. *Entropy*, 13(11):1945–1957, 2011.

P. Baudot and D. Bennequin. The homological nature of entropy. *Entropy*, 17(5):3253–3318, 2015.

D. Bennequin and J. P. Vigneaux. A functional equation related to generalized entropies and the modular group. *Aequationes Mathematicae*, 2020.

S. Bloch and H. Esnault. The additive dilogarithm. *Kazuya Kato*, page 131, 2003.

J.-L. Cathelineau. Sur l'homologie de SL_2 à coefficients dans l'action adjointe. *Mathematica Scandinavica*, pages 51–86, 1988.

J.-L. Cathelineau. Remarques sur les différentielles des polylogarithmes uniformes. *Annales de l'Institut Fourier*, 46(5):1327–1347, 1996.

D. Cohn. *Measure Theory: Second Edition*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer New York, 2013.

A. Connes and C. Consani. Homological algebra in characteristic one. *Higher Structures*, 3(1), 2019.

C. M. Constantin and A. Döring. Contextual entropy and reconstruction of quantum states. *arXiv preprint arXiv:1208.2046*, 2012.

T. Cover and J. Thomas. *Elements of Information Theory*. A Wiley-Interscience publication. Wiley, 2006. ISBN 9780471748816.

I. Csiszár. Axiomatic characterizations of information measures. *Entropy*, 10(3):261–273, 2008.

J. Curry. *Sheaves, cosheaves and applications*. PhD thesis, The University of Pennsylvania, 2013. arXiv:1303.3255.

Z. Daróczy. Generalized information functions. *Information and control*, 16(1):36–51, 1970.

N. De Silva and R. S. Barbosa. Contextuality and noncommutative geometry in quantum mechanics. *Communications in Mathematical Physics*, 365(2):375–429, 2019.

P. Elbaz-Vincent and H. Gangl. On poly(ana)logs I. *Compositio Mathematica*, 130(2):161–214, 2002.

P. Elbaz-Vincent and H. Gangl. Finite polylogarithms, their multiple analogues and the Shannon entropy. In *International Conference on Geometric Science of Information*, pages 277–285. Springer, 2015.

T. Fritz and R. Chaves. Entropic inequalities and marginal problems. *IEEE transactions on Information Theory*, 59(2):803–817, 2013.

M. Gromov. In a search for a structure, part 1: On entropy. *Preprint available at http://www. ihes. fr/gromov*, 2012.

A. Grothendieck. Sur quelques points d'algèbre homologique, I. *Tohoku Mathematical Journal*, 9(2):119–221, 1957.

P. R. Halmos. *Finite-Dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, USA, reprint of 2nd edition, 1958.

H. Hatori. A note on the entropy of a continuous distribution. *Kodai Mathematical Seminar Reports*, 10(4):172–176, 1958.

J. Havrda and F. Charvát. Quantification method of classification processes. Concept of structural *a*-entropy. *Kybernetika*, 3(1):30–35, 1967.

P. Kannappan and C. T. Ng. Measurable solutions of functional equations related to information theory. *Proceedings of the American Mathematical Society*, 38(2):pp. 303–310, 1973.

A. Khinchin. *Mathematical Foundations of Information Theory*. Dover Books on Mathematics Series. Dover Publications, 1957.

P. M. Lee. On the axioms of information theory. *The Annals of Mathematical Statistics*, 35(1):415–418, 03 1964.

G. L. Litvinov. Maslov dequantization, idempotent and tropical mathematics: A brief introduction. *Journal of Mathematical Sciences*, 140(3):426–444, 2007.

S. Mac Lane. *Homology*. Springer-Verlag, Germany, reprint of the 1975 edition, 1994.

S. Mac Lane. *Categories for the Working Mathematician*. Graduate Texts in Mathematics. Springer-Verlag, New York, USA, 1998.

H. Matsuda. Information theoretic characterization of frustrated systems. *Physica A: Statistical Mechanics and its Applications*, 294(1-2):180–190, 2001.

R. Matveev and J. W. Portegies. Asymptotic dependency structure of multiple signals. *Information Geometry*, 1(2):237–285, 2018.

M. Mézard and A. Montanari. *Information, Physics, and Computation*. Oxford Graduate Texts. OUP Oxford, 2009. ISBN 9780198570837.

A. Otáhal. Finiteness and continuity of differential entropy. In *Asymptotic Statistics*, pages 415–419. Springer, 1994.

A. Pelizzola. Cluster variation method in statistical physics and probabilistic graphical models. *Journal of Physics A: Mathematical and General*, 38(33):R309, 2005.

I. Z. Ruzsa. Sumsets and entropy. *Random Structures & Algorithms*, 34(1):1–10, 2009.

P. Schapira. Algebra and topology. *Lecture Notes from Paris VI*, 2008.

C. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.

T. Stacks Project Authors. *Stacks Project.* https://stacks.math.columbia.edu, 2018.

T. Tao. *Topics in Random Matrix Theory.* American Mathematical Society, 2012.

C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52(1-2):479–487, 1988.

H. Tverberg. A new derivation of the information function. *Mathematica Scandinavica*, 6:297–298, 1958.

S. Varadhan. Large deviations and entropy. In A. Greven, G. Keller, and G. Warnecke, editors, *Entropy*, chapter 9, pages 199–214. Princeton University Press, Princeton and Oxford, 2003.

J. P. Vigneaux. *Topology of Statistical Systems: A Cohomological Approach to Information Theory.* PhD thesis, Université de Paris, 2019.

P. O. Vontobel. Counting in graph covers: A combinatorial characterization of the Bethe entropy function. *IEEE Transactions on Information Theory*, 59(9):6018–6048, 2013.

C. A. Weibel. *An introduction to homological algebra.* Cambridge University Press, Printed in USA, 1994.

Université de Paris, Sorbonne Université, CNRS, Institut de Mathématiques de Jussieu-Paris Rive Gauche (IMJ-PRG), 75013, Paris, France.
Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstraße 22, 04013, Leipzig
Email: vigneaux@mis.mpg.de

This article may be accessed at http://www.tac.mta.ca/tac/