# Equivalence and Conditional Independence in Atomic Sheaf Logic

ALEX SIMPSON*, Faculty of Mathematics and Physics, University of Ljubljana
Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

We propose a semantic foundation for logics for reasoning in settings that possess a distinction between equality of variables, a coarser equivalence of variables, and a notion of conditional independence between variables. We show that such relations can be modelled naturally in atomic sheaf toposes. Equivalence of variables is modelled by an intrinsic relation of *atomic equivalence* that is possessed by every atomic sheaf. We identify additional structure on the category generating the atomic topos (primarily, the existence of a system of *independent pullbacks*) that allows the relation of conditional independence to be interpreted in the topos. We then study the logic of equivalence and conditional independence that is induced by the internal logic of the topos. This *atomic sheaf logic* is a classical logic that validates a number of fundamental reasoning principles relating equivalence and conditional independence. As a concrete example of this abstract framework, we use the atomic topos over the category of surjections between finite nonempty sets as our main running example. In this category, the interpretations of equivalence and conditional independence coincide with those given by the multiteam semantics of independence logic, in which the role of equivalence is taken by the relation of mutual inclusion. A major difference from independence logic is that, in atomic sheaf logic, the multiteam semantics of the equivalence and conditional independence relations is embedded within a classical surrounding logic. At the end of the paper, we briefly outline two other instances of our framework, to demonstrate its versatility. The first of these is a category of *probability sheaves*, in which atomic equivalence is equality-in-distribution, and the conditional independence relation is the usual probabilistic one. Our other example is the *Schanuel topos* (equivalent to nominal sets) where equivalence is orbit equality and conditional independence amounts to a relative form of separatedness.

CCS Concepts: • **Theory of computation → Logic**.

Additional Key Words and Phrases: Logics for probability, Categorical probability theory, Conditional independence, Dependence logic, Team semantics, Sheaves, Toposes

**ACM Reference Format:**
Alex Simpson. 2025. Equivalence and Conditional Independence in Atomic Sheaf Logic. *J. ACM* XX, X, Article XXX (X 2025), 53 pages.
https://doi.org/XXXXXXX.XXXXXXX

## 1 Introduction

This paper provides a study of fundamental logical principles for reasoning about relations of *independence* and *conditional independence* together with an associated relation of *equivalence*. The principles, which are obtained via the abstract mathematical framework of *sheaf theory*, are general, in the sense that that they apply uniformly to different

*Research supported by John Templeton Foundation grant number 39465 (2013–14).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 731143.

Author's Contact Information: Alex Simpson, Alex.Simpson@fmf.uni-lj.si, Faculty of Mathematics and Physics, University of Ljubljana
Institute of Mathematics, Physics and Mechanics, Ljubljana , Ljubljana, Slovenia Jadranska 19.

instantiations of the notions of (conditional) independence and equivalence in a number of very different application areas. The paper focuses on the mathematical development of a general theory that is intended to be cross-disciplinary in its applicability, but with computer science as a particularly prominent source of target application areas.

Notions of *independence* and *conditional independence* arise in many scientific areas. One particularly significant area is in probability and statistics, where it has long been recognised that conditional independence relations are subject to subtle rules of inference [7, 40]. Such rules, in a graphical formulation, are crucial in the technology of Bayesian networks [15, 16, 33]. In a more logical form, they have received recent interest in the area of program verification, where, for example, versions of separation logic based on probabilistic independence have been developed [2, 5, 27].

In a different direction, the *dependence* and *independence logics* of Väänänen and Grädel [17, 42] are concerned with purely logical notions of dependence and independence between variables. Such logics are based on *team semantics*, which develops Hodges' compositional approach [20] to the semantics of *independence-friendly logic* [19] into a fully fledged semantic framework. One of the attractions of team semantics is the close relationship it enjoys with database theory and notions of dependence and independence that arise therein [18]. There is also an intriguing aspect to team semantics: it gives rise to logics that are exotic in character. This point is discussed in more detail in Section 10.1.

The starting point of the present paper, in Sections 2 and 3, is the observation that the interpretation given by team semantics, more precisely by its *multiteam* variant [11], to conditional independence statements is equivalent to interpreting these relations in a certain sheaf topos, namely the topos of atomic sheaves on the category $\mathbb{Sur}$ of finite nonempty sets and surjections. This means that the team semantics of conditional independence automatically has a logic canonically associated with it: the internal logic of the topos. Since the topos is atomic, this internal logic is ordinary classical logic, albeit with a nonstandard semantics. We thus obtain a classical logic suitable for reasoning with conditional independence relations endowed with their (multi)team semantics (Section 4).

One advantage of the atomic sheaf perspective on conditional independence is that it is very general. We axiomatise structure, on the generating category of the topos, that gives rise to a canonical interpretation of conditional independence relations. For this, we define, in Section 6, the notion of *independent pullback structure* on a category, closely related to the *conditional independence structure* of [38], but with a much more compact axiomatisation. We also expose a surprisingly rich interplay between independent pullback structure and the induced atomic sheaves. Building on this, in Section 7, we define *atomic conditional independence*, generalising the multiteam conditional-independence relation to any atomic sheaf topos over a generating category with sufficient structure

Along the route to defining conditional independence, we observe, in Section 5, that every object of an atomic topos carries, in addition to the standard equality relation on the object, an additional intrinsic equivalence relation, which we call *atomic equivalence*. Logically, this provides us with a canonical equivalence relation between variables that is, in general, coarser than equality. In the example of the atomic topos on the category $\mathbb{Sur}$, atomic equivalence turns out to coincide with a relation of interest in team semantics, namely the *equiextension* relation.

One important contribution of the paper is the identification of fundamental axioms for relations of equivalence and conditional independence that are validated by the general interpretation of these relations in atomic toposes (over generating categories with enough structure). These axioms include the usual quantifier-free axioms from the literature (for example, axioms formalising the reasoning principles for conditional independence from [7, 40]), but also new first-order axioms that fully exploit the use of atomic sheaf logic. In Sections 5 and 7, we identify five such principles: the *transfer principle*, the *invariance principle*, the *principle of independent equivalence*, the *independent existence principle* and the property of *existence preservation*.

Throughout Sections 3–7, the abstract definitions are illustrated in the case of atomic sheaves over the category $\mathbb{S}\mathrm{ur}$, which is our main running example, chosen because of its connection to (multi)team semantics. In Sections 8 and 9 we present two other examples of our general structure, in order to give some indication of its versatility. Section 8 presents an atomic sheaf topos over a category of probability spaces. The resulting category of *probability sheaves* (first introduced in [37]) includes sheaves of random variables, over which equality coincides with the probabilistic relation of *almost sure equality*, atomic equivalence coincides with the relation of *equality in distribution*, and the atomic conditional independence relation coincides with the usual probabilistic relation. Section 9 very briefly indicates how the Schanuel topos (which is equivalent to the category of nominal sets [14, 35]) fits into our framework. In this case, atomic equivalence is the relation of orbit equality, and conditional independence amounts to a relative form of separatedness.

Finally, in Section 10, we discuss related and potential future work, including a detailed comparison with team and multiteam semantics in Section 10.1, and a discussion of potential computer science applications in Section 10.2.

This paper is an expanded version of a conference paper [39], presented at the thirty-ninth annual ACM/IEEE Symposium on Logic in Computer Science (LICS), held in Tallinn, Estonia in July 2024. In comparison with the conference paper, this journal version includes proofs of all main results, as well as an expanded discussion on sheaves in Section 3 and also a substantially expanded presentation of our second main example, the category of probability sheaves, which occupies Section 8. We further include three new appendices containing lengthy proofs that we prefer not to incorporate into the main body of the paper, where they would interrupt the flow.

## 2 Multiteam semantics

*Dependence logic* [42] and *independence logic* [17] extend first-order logic with new logical primitives expressing notions of dependence and independence between variables. These logics are based on the realisation that such new primitives can be interpreted semantically, by replacing the usual *assignments* used to interpret variables in logical formulas with *teams* (sets of assignments) or with *multiteams* (multisets of assignments). The relevant definitions are as follows, where $A$ is an arbitrary set.

- An $A$-valued *assignment* is a function $\mathcal{V} \to A$ where $\mathcal{V}$ is a (without loss of generality finite) set of variables.
- An $A$-valued *team* [20, 42] is a set of assignments with common variable set $\mathcal{V}$.
- An $A$-valued *multiteam* [11] is a multiset of assignments with common variable set $\mathcal{V}$.

Teams and multiteams give a canonical semantics to a variety of interesting new logical relations between variables, such as those expressing *dependence* $=(\mathsf{x}, \mathsf{y})$, *independence* $\mathsf{x} \perp \mathsf{y}$, *conditional independence* $\mathsf{x} \perp_{\mathsf{z}} \mathsf{y}$, *inclusion* $\mathsf{x} \subseteq \mathsf{y}$, *equiextension* $\mathsf{x} \bowtie \mathsf{y}$ and *exclusion* $\mathsf{x}|\mathsf{y}$, to give a non-exhaustive list. We review this in detail, in the case of multiteams, focusing on two of the above relations: conditional independence and equiextension.

A *multiset* of elements from a set $A$ is a function $m \colon A \to \mathbb{N}$, which assigns to every element $a \in A$ a *multiplicity* $f(a)$. A multiset $m$ is *finite* if its *support* (the set $\mathrm{supp}(m) := \{a \mid m(a) > 0\}$) is finite. A multiset $m \colon A \to \mathbb{N}$ can alternatively be presented by a set $\Omega$ together with a function $M \colon \Omega \to A$ satisfying, for all $a \in A$, the fibre $M^{-1}(a)$ has cardinality $m(a)$. The elements of $\Omega$ can be thought of as names for distinct element occurrences in the multiset (so each element in $A$ has as many names as its multiplicity). Note also that the function $M$ has the support set $\mathrm{supp}(m)$ as its image. Of course a multiset $m \colon A \to \mathbb{N}$ has many different presentations by finite-fibre functions. However, given two such representations $M \colon \Omega \to A$ and $M' \colon \Omega' \to A$, there exists a bijection $i \colon \Omega \to \Omega'$ such that $M = M' \circ i$. (The proof of this statement, although simple, requires the axiom of choice.) So multisets are in one-to-one correspondence with isomorphism classes of presentations.

In the case of a finite multiset $m \colon A \to \mathbb{N}$, the domain set $\Omega$ of a presentation $M \colon \Omega \to A$ is necessarily finite, and all functions with finite domain present finite multisets. Thus there is a one-to-one correspondence between finite multisets and isomorphism classes of finite-domain presentations. (Moreover, because the multisets are now finite, the axiom of choice is no longer needed.)

Since a *multiteam* is a multiset of assignments with a common $\mathcal{V}$. it can be presented by a finite-fibred function of the form

$$M \colon \Omega \to (\mathcal{V} \to A) \ \ .$$

As in [11], we restrict attention to finite multiteams. Henceforth, by *multiteam* we mean a finite multiset of assignments with common $\mathcal{V}$. Such finite multiteams correspond to functions $M$, as above, for which the set $\Omega$ is finite. Equivalently, by transposition, a multiteam can be represented by a function of the form

$$\underline{M} \colon \mathcal{V} \to (\Omega \to A)$$

While this is just a simple set-theoretic reorganisation of the notion of multiteam, it provides an illuminating alternative perspective on multiteam semantics, which we now elaborate.

One can think of a function $X \colon \Omega \to A$ as a *nondeterministic variable* valued in $A$. Here the terminology is motivated by analogy with the notion of *random variable* from probability theory. In our setting, we view the *set* $\Omega$ as a finite *sample set*, a nondeterministic version of a *sample space* in probability theory. The sample set represents a realm of possible nondeterministic choices. With this terminology, a multiteam presented as $\underline{\rho} \colon \mathcal{V} \to (\Omega \to A)$ is simply an assignment of $A$-valued nondeterministic variables (with shared sample set) to logical variables. (In this paper, we restrict to finite sample sets. Nevertheless, the notion of nondeterministic variable obviously generalises to arbitrary sample sets $\Omega$.)

We now use the above formulation of multiteams as assignments of nondeterministic variables to recast definitions from multiteam semantics (as in [11]). Technically, this is simply a straightforward matter of translating the definitions along the equivalence between the two formulations of multiteam. However, even if mathematically equivalent, our formulation of multiteam encourages a different 'local' style of presentation, where the sample sets $\Omega$ play a role similar to that played by *possible worlds* in Kripke semantics and by *forcing conditions* in set theory.

Before addressing semantics, we introduce our syntax. For greater generality, we work with a multi-sorted logic. This also has the advantage that the sorting constraints on logical primitives provide useful information about their generality in scope. Accordingly, we assume a set Sort of basic syntactic *sorts* $\mathsf{A}, \mathsf{B}, \mathsf{C}, \ldots$. Variables $\mathsf{x}^{\mathsf{A}}$ have explicit sorts. We consider three forms of atomic formula.

- If $\mathsf{x}^{\mathsf{A}}, \mathsf{y}^{\mathsf{A}}$ have the same sort, then $\mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}}$ is an atomic formula.
- If $\mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n}$ and $\mathsf{y}_1^{\mathsf{A}_1}, \ldots, \mathsf{y}_n^{\mathsf{A}_n}$ are two lists of variables of the same length $n \geq 0$ with identical sort lists, then
$$\mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n} \sim \mathsf{y}_1^{\mathsf{A}_1}, \ldots, \mathsf{y}_n^{\mathsf{A}_n} \tag{1}$$
is an atomic formula.
- If $\mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_m^{\mathsf{A}_m}$ and $\mathsf{y}_1^{\mathsf{B}_1}, \ldots, \mathsf{y}_n^{\mathsf{B}_n}$ and $\mathsf{z}_1^{\mathsf{C}_1}, \ldots, \mathsf{z}_l^{\mathsf{C}_l}$ are three lists of variables (with $m, n, l \geq 0$) then
$$\mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_m^{\mathsf{A}_m} \perp \mathsf{y}_1^{\mathsf{B}_1}, \ldots, \mathsf{y}_n^{\mathsf{B}_n} \mid \mathsf{z}_1^{\mathsf{C}_1}, \ldots, \mathsf{z}_l^{\mathsf{C}_l} \tag{2}$$

is an atomic formula.

The first formula expresses *equality*, as in ordinary (multi-sorted) first-order logic. The remaining two are atomic constructs borrowed from logics associated with team semantics.

The formula $\vec{x} \sim \vec{y}$ represents what we call *equivalence*, which arises in the team-semantics literature as *equiextension* $\vec{x} \subseteq \vec{y} \wedge \vec{y} \subseteq \vec{x}$, sometimes written with the notation $\vec{x} \bowtie \vec{y}$. Our more neutral notation and terminology reflects the fact that we will later consider other interpretations of the $\sim$ relation. The use of vectors of variables on either side is needed because equivalence is a relation that holds betwen the vectors $\vec{x}$ and $\vec{y}$ *jointly*, and does not reduce to a conjunction of equivalences between components.

The formula $\vec{x} \perp \vec{y} \mid \vec{z}$ represents *conditional independence* from the *independence logic* of [17], where it is written $\vec{x} \perp_{\vec{z}} \vec{y}$. In our syntax, we take the conditioning variables out of the subscript position in order to give them more prominence, adopting a notation that is familiar from probability theory. An important special case of conditional independence is when the sequence $\vec{z}$ is empty. In such cases, we write simply $\vec{x} \perp \vec{y}$ for the resulting relation, which expresses unconditional independence.

It is of course the atomic formulas $\vec{x} \sim \vec{y}$ and $\vec{x} \perp \vec{y} \mid \vec{z}$ that give us *equivalence* and *conditional independence* in the title of this paper.

To define the semantics, we assume we have, for every sort A, an associated set $[\![A]\!]$. In a multi-sorted setting, an assignment for a finite set $\mathcal{V}$ of variables is an element

$$\rho \;\in\; \prod_{x^A \in \mathcal{V}} [\![A]\!] \;\;,$$

and a multiteam is a finite multiset of assignments. In the standard multiteam semantics, a formula $\Phi(x_1^{A_1}, \ldots, x_n^{A_n})$ (i.e., all free variables are in $\{x_1^{A_1}, \ldots, x_n^{A_n}\}$) is given a satisfaction relation

$$\models_m \Phi \;\;, \tag{3}$$

where $m$ is a multiteam of $\{x_1^{A_1}, \ldots, x_n^{A_n}\}$-assignments. If instead we adopt the reformulation of multisets described above, a multiteam is given as a *single* assignment

$$\underline{\rho} \;\in\; \prod_{x^A \in \mathcal{V}} (\Omega \to [\![A]\!]) \tag{4}$$

of nondeterministic variables to logic variables, and the satisfaction relation can then be rewritten as

$$\models_{\underline{\rho}} \Phi \;\;. \tag{5}$$

It turns out to be helpful to make the sample set $\Omega$, that occurs implicitly within $\underline{\rho}$, explicit in the notation, so we write

$$\Omega \Vdash_{\underline{\rho}} \Phi \;\;. \tag{6}$$

We here switch to the 'forcing' notation $\Vdash$, since we shall view $\Omega$ as a 'possible world' or 'condition' (capturing all the nondeterminism that the multiteam uses) that determines the 'local truth' of $\Phi$. We stress that relations (3), (5) and (6) all have exactly the same meaning. The only differences are in the formulation of multiset that is used, and whether or not $\Omega$ is explicit in the notation.

Figure 1 defines the forcing relation $\Omega \Vdash_{\underline{\rho}} \Phi$ directly in terms of our reformulated multiteams, as in (4), for atomic formulas $\Phi$. In the clauses for equivalence and independence we use the notation $\underline{\rho}(\vec{x})$, where $\vec{x}$ is a vector of variables

$$\Omega \Vdash_{\underline{\rho}} \mathsf{x}^{\mathsf{A}} \!=\! \mathsf{y}^{\mathsf{A}} \;\Leftrightarrow\; \underline{\rho}(\mathsf{x}^{\mathsf{A}}) = \underline{\rho}(\mathsf{y}^{\mathsf{A}}) \quad \text{(equal functions } \Omega \to [\![\mathsf{A}]\!])$$
$$\Omega \Vdash_{\underline{\rho}} \vec{\mathsf{x}} \sim \vec{\mathsf{y}} \;\Leftrightarrow\; \underline{\rho}(\vec{\mathsf{x}}) \bowtie \underline{\rho}(\vec{\mathsf{y}})$$
$$\Omega \Vdash_{\underline{\rho}} \vec{\mathsf{x}} \perp \vec{\mathsf{y}} \,|\, \vec{\mathsf{z}} \;\Leftrightarrow\; \underline{\rho}(\vec{\mathsf{x}}) \perp\!\!\!\perp \underline{\rho}(\vec{\mathsf{y}}) \,|\, \underline{\rho}(\vec{\mathsf{z}})$$

Fig. 1. Multiteam semantics of atomic formulas

$\mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n}$, to represent the $([\![\mathsf{A}_1]\!] \times \cdots \times [\![\mathsf{A}_n]\!])$-valued nondeterministic variable

$$\underline{\rho}(\vec{\mathsf{x}}) \;:=\; \omega \mapsto (\underline{\rho}(\mathsf{x}_1^{\mathsf{A}_1}), \ldots, \underline{\rho}(\mathsf{x}_n^{\mathsf{A}_n})) : \Omega \to [\![\mathsf{A}_1]\!] \times \cdots \times [\![\mathsf{A}_n]\!] \;.$$

We also write $X \bowtie Y$ and $X \perp\!\!\!\perp Y \mid Z$ for the semantic relation of *equiextension* and *conditional independence* between nondeterministic variables, as defined below.

*Definition 2.1 (Equiextension for nondeterministic variables).* Two nondeterministic variables $X : \Omega \to A$ and $Y : \Omega \to A$ are *equiextensive* (notation $X \bowtie Y$) if they have equal images, i.e., $X(\Omega) = Y(\Omega)$.

*Definition 2.2 (Conditional independence for nondeterministic variables).* Let $X : \Omega \to A$, $Y : \Omega \to B$ and $Z : \Omega \to C$ be nondeterministic variables. We say that $X$ and $Y$ are *conditionally independent* given $Z$ (notation $X \perp\!\!\!\perp Y \mid Z$) if, for all $a \in A, b \in B, c \in C$,

$$(\exists \omega \in \Omega.\ X(\omega) \!=\! a \text{ and } Z(\omega) \!=\! c) \text{ and } (\exists \omega \in \Omega.\ Y(\omega) \!=\! b \text{ and } Z(\omega) \!=\! c)$$
$$\text{implies } \exists \omega \in \Omega.\ X(\omega) \!=\! a \text{ and } Y(\omega) \!=\! b \text{ and } Z(\omega) \!=\! c \;.$$

In the literature on (in)dependence logics, the semantic clauses for atomic formulas are extended with clauses giving meaning to the logical connectives and quantifiers. A number of inequivalent ways of achieving this appear in the literature [11, 17, 42]. All share the feature that the resulting logics are exotic. We shall discuss this in more detail in Section 10.1.

In this paper, we consider a different approach to embedding the equivalence and conditional independence constructs, with their multiteam semantics, in a full multi-sorted first-order logic. We observe that the multiteam semantics of the atomic constructs lives naturally in a certain *atomic sheaf topos*. and then we make use of the standard *internal logic* of the topos, which in the case of an atomic topos is classical logic.

## 3 Atomic sheaves

In this section, we define the notion of *atomic sheaf topos*, which is a special kind of *Grothendieck topos*. We restrict attention to presenting the definitions and results we shall make use of, attempting to do so in such a way that they can be understood from first principles given knowledge of core category theory. For further contextualisation, [29] is an excellent source.

A *presheaf* on a small category $\mathbb{C}$ is a functor $P : \mathbb{C}^{\mathrm{op}} \to \mathbf{Set}$ (note the contravariance). The *presheaf category* $\mathrm{Psh}(\mathbb{C})$ is the functor category $\mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$. Given a presheaf $P$, object $Y$ of $\mathbb{C}$, element $y \in PY$ and map $f : X \to Y$ in $\mathbb{C}$, we write $y \cdot_P f$ for the element $P(f)(y) \in PX$, or simply $y \cdot f$ when $P$ is clear from the context.

*Example 3.1 (Representable presheaves).* For any object $Z \in \mathbb{C}$, the *representable presheaf* $\mathbf{y}Z := \mathbb{C}(-, Z)$ is defined by

- For any object $X \in \mathbb{C}$, define $\mathbf{y}Z(X) \;:=\; \mathbb{C}(X, Z)$, i.e., the hom set.

- For any map $f\colon Y \to X$ in $\mathbb{C}$ and $g \in (\mathbf{y}Z)(X)$, define $g \cdot f := g \circ f$.

The object mapping $Z \mapsto \mathbf{y}Z$ extends to a full and faithful functor $\mathbf{y} : \mathbb{C} \to \mathrm{Psh}(\mathbb{C})$, the *Yoneda functor* [29].

*Example 3.2 (Product presheaves).* Let $P_1, \dots, P_n$ be presheaves on $\mathbb{C}$. Define the *product presheaf* $P_1 \times \cdots \times P_n$ in $\mathrm{Psh}(\mathbb{C})$ by:

- For any object $X \in \mathbb{C}$, define

$$(P_1 \times \cdots \times P_n)(X) \;:=\; P_1(X) \times \cdots \times P_n(X) \;,$$

  i.e., the product of sets.
- For any map $f\colon Y \to X$ in $\mathbb{C}$ and $(x_1, \dots, x_n) \in (P_1 \times \cdots \times P_n)(X)$, define

$$(x_1, \dots, x_n) \cdot f \;:=\; (x_1 \cdot_{P_1} f, \dots, x_n \cdot_{P_n} f) \;,$$

The above definition generalises to infinite products, and further to arbitrary category-theoretic limits and colimits, all of which are defined on presheaves in a similar (pointwise) way, using the corresponding definitions in the category of sets.

The next example is central to this paper.

*Example 3.3.* Let $\mathbb{Sur}$ be (a small category equivalent to) the category whose objects are non-empty finite sets and whose morphisms are surjective functions. For any set $A$, we have a presheaf $\underline{\mathsf{NV}}(A)$ in $\mathrm{Psh}(\mathbb{Sur})$ of $A$-valued *nondeterministic variables* (in the sense of Section 2), defined as follows.

- For any object $\Omega$ of $\mathbb{Sur}$, define $\underline{\mathsf{NV}}(A)(\Omega)$ to be the set of all functions $\Omega \to A$.
- For any map $p\colon \Omega' \to \Omega$ in $\mathbb{Sur}$, and $X \in \underline{\mathsf{NV}}(A)(\Omega)$ define $X \cdot p$ to be $X \circ p \in \underline{\mathsf{NV}}(A)(\Omega')$.

Grothendieck introduced a very general notion of what it means for a presheaf $P \in \mathrm{Psh}(\mathbb{C})$ to be a *sheaf* relative to a *Grothendieck topology* on $\mathbb{C}$. A Grothendieck topology $\mathcal{J}$ specifies, for every object $X$, a collection $\mathcal{J}_X$ of families of maps with codomain $X$, in which each family of maps $(c_i\colon Y_i \longrightarrow X)_{i\in I} \in \mathcal{J}_X$ is deemed to provide a *covering family* (more briefly *cover*) for $X$. A presheaf $P$ is a $\mathcal{J}$*-sheaf* if, for every such cover, every *matching* family of elements $(y_i \in P(Y_i))_{i\in I}$ has a unique *amalgamation* $x \in P(X)$. The high-level idea is that the *matching* property, which says that the $y_i$ elements agree with each other on overlapping parts of the cover, allows all the $y_i$ to be glued together into a single *amalgamation* $x$, which is an element of $P(X)$. We shall not give the general definitions underlying the emphasised words because, for this paper, it is not necessary to understand the notion of sheaf in its full generality. Nonetheless, there is a point about the general definition worth making. The intuition that is usually presented for the general definition is that the matching condition for the family $(y_i \in P(Y_i))_{i\in I}$ means that the different $y_i$ are compatible with each other, and then the unique amalgamation 'glues' these compatible elements together to form a single element $x \in P(X)$, which is possible because the object $X$ is covered by the family $(Y_i)_{i\in I}$. In this paper, we are going to work only with sheaves for *atomic* Grothendick topologies, for which the usual general intuition for sheaves outlined above is not very helpful. In the case of an atomic topology, covers are single maps $c\colon Y \longrightarrow X$, matching families contain only one element $y \in P(Y)$ (it needs to match only with itself, which turns out to be a nontrivial condition) and the amalgamation $x \in P(X)$ is obtained from $y$ alone, so the usual 'gluing' intuition does not apply. Instead, we shall use the terminology *invariant element* in place of matching family, and *descendent* in place of amalgamation, since this seems more appropriate in the context of an atomic topology.

We first introduce the atomic sheaf concept using the example of $\mathbb{Sur}$, and then follow this with the generalisation to an arbitrary small category $\mathbb{C}$. In the case of $\mathbb{Sur}$, an object $\Omega$ can be thought of as representing a 'world' of currently available nondeterministic choices, and a map $c : \Omega' \to \Omega$ specifies an extension of the existing nondeterministic choices in $\Omega$ to accommodate the additional nondeterminism potentially available in $\Omega'$. Nondeterministic variables form a presheaf $\underline{\mathsf{NV}}(A)$ simply because any nondeterministic variable $X \in \underline{\mathsf{NV}}(A)(\Omega)$ extends via $c$ to a corresponding $\Omega'$-based nondeterministic variable $X \cdot c := X \circ c \in \underline{\mathsf{NV}}(A)(\Omega')$. This latter nondeterministic variable is defined for all nondeterministic choices in $\Omega'$, but only makes use of nondeterminism already available in $\Omega$; that is, $(X \cdot c)(\omega') = (X \cdot c)(\omega'')$ for any $\omega', \omega'' \in \Omega'$ for which $c(\omega') = c(\omega'')$. Furthermore, every element $Y \in \underline{\mathsf{NV}}(A)(\Omega')$, that only makes use of nondeterminism in $\Omega$, arises as $Y = X \cdot c$ for a unique $X \in \underline{\mathsf{NV}}(A)(\Omega)$. In other words, it has a unique representation as a *bona fide* $\Omega$-based nondeterministic variable $X$. In order to formulate this technically, we say that a nondeterministic variable $Y \in \underline{\mathsf{NV}}(A)(\Omega')$ is *c-invariant* if $Y(\omega') = Y(\omega'')$ for any $\omega', \omega'' \in \Omega'$ for which $c(\omega') = c(\omega'')$. The presheaf $\underline{\mathsf{NV}}(A)$ then satisfies: every $c$-invariant $Y \in \underline{\mathsf{NV}}(A)(\Omega')$ arises as $X \cdot c$ for a unique $X \in \underline{\mathsf{NV}}(A)(\Omega)$, which we call the *c-descendent* of $Y$. As we shall see below, the property we have just elucidated asserts that the presheaf $\underline{\mathsf{NV}}(A)$ is a *sheaf* for the *atomic Grothendieck topology* on the category $\mathbb{Sur}$.

A similar story can be told for any small category $\mathbb{C}$ for which an object $X \in \mathbb{C}$ can be thought of as a world of current possibilities, and a map $c \colon Y \to X$ represents a way of extending the current world to another world $Y$ with additional possibilities. Given a presheaf $P$ an element $x \in P(X)$ and map $c \colon Y \to X$, the element $x \cdot c \in P(Y)$ represents the extension of $x$ to incorporate the new possibilities from $Y$. The extended element $x \cdot c$ enjoys the property of *c-invariance* (Definition 3.6 below), which formalises that $x \cdot c$ does not depend on any of the possibilities in $Y$ beyond those already available in $X$. Moreover, for any $y \in P(Y)$ that is $c$-invariant, the definition of *atomic sheaf* (Definition 3.8 below) says that there must exist a unique $x \in P(X)$ that, via the equation $y = x \cdot c$, makes explicit the true dependency of $y$ only on $X$.

The main intuition underpinning the above discussion can be summarised as follows. In the context of a category $\mathbb{C}$, for which we think of maps $c \colon Y \to X$ as extending the possibilities offered by state $X$ to a more refined set of possibilities offered by state $Y$,

- the *presheaf* property of $P$ says that we can *extend* any element $x \in P(X)$, defined using the possibilities at $X$, to a corresponding element $x \cdot c \in P(Y)$ that, although defined at $Y$, does not exploit the potential additional generality of $Y$;
- and the *atomic sheaf* property says that, for any element $y \in P(Y)$, defined using the possibilities at $Y$ in such a way that $y$ does not exploit the potential greater generality afforded by $Y$ over $X$, there exists a unique corresponding element $x \in P(X)$ that makes explicit the dependency of $y$ only on possibilities offered by $X$.

Since we are interested only in atomic topologies, we can define the sheaf property (Definition 3.8 below) directly, without needing to introduce the general notion of Grothendieck topology. However, we do need the atomic Grothendieck topology to exist on the base category $\mathbb{C}$, which happens if and only if the category $\mathbb{C}$ is *coconfluent*.

*Definition 3.4 (Coconfluence).* A category $\mathbb{C}$ is *coconfluent*[1] if for any cospan $X \xrightarrow{f} Z \xleftarrow{g} Y$, there exists a span $X \xleftarrow{u} W \xrightarrow{v} Y$ such that $f \circ u = g \circ v$.

Proposition 3.5. *$\mathbb{Sur}$ is coconfluent.*

[1] In [23, A 2.1.11(h)] $\mathbb{C}$ is said to satisfy the *right Ore condition.*

Proof. Consider any cospan $\Omega_X \xrightarrow{p} \Omega_Z \xleftarrow{q} \Omega_Y$ in $\mathbb{S}\mathrm{ur}$, Define

$$\Omega_W \ := \ \{(x, y) \in \Omega_X \times \Omega_Y \mid p(x) = q(y)\} \ .$$

Then $u := (x, y) \mapsto x$ and $v := (x, y) \mapsto y$ define surjective functions $\Omega_W \twoheadrightarrow \Omega_X$ and $\Omega_W \twoheadrightarrow \Omega_Y$, hence they are maps in $\mathbb{S}\mathrm{ur}$, for which indeed $p \circ u = q \circ v$. (More briefly, the pullback in **Set** is a commuting square in $\mathbb{S}\mathrm{ur}$, though not a pullback in $\mathbb{S}\mathrm{ur}$.) □

Let $P \in \mathrm{Psh}(\mathbb{C})$ be a presheaf.

*Definition 3.6 (Invariant element).* Given $c \colon Y \to X$ and $y \in P(Y)$ we say that $y$ is *$c$-invariant* if, for any parallel pair of maps $d, e \colon Z \to Y$ such that $c \circ d = c \circ e$, it holds that $y \cdot d = y \cdot e$.

*Definition 3.7 (Descendent).* Given $c \colon Y \to X$ and $y \in P(Y)$ we say that $x \in P(X)$ is a *$c$-descendent* of $y$ if $y = x \cdot c$.

It is easily seen that if $x$ is a $c$-descendent of $y$ then $y$ is $c$-invariant. The notion of sheaf imposes a converse.

*Definition 3.8 (Atomic sheaf).* A presheaf $P \in \mathrm{Psh}(\mathbb{C})$ is an *atomic sheaf* if, for every map $c \colon Y \to X$ in $\mathbb{C}$, every $c$-invariant $y \in P(Y)$ has a unique $c$-descendent $x \in P(X)$.

We shall also have use for the following weakening of the notion of sheaf.

*Definition 3.9 (Separated presheaf).* A presheaf $P \in \mathrm{Psh}(\mathbb{C})$ is an *separated* (with respect to the atomic topology) if, for every map $c \colon Y \to X$ in $\mathbb{C}$, every $c$-invariant $y \in P(Y)$ has at most one $c$-descendent $x \in P(X)$

Proposition 3.10. *A presheaf $P \in \mathrm{Psh}(\mathbb{C})$ is separated if and only if, for all $x, y \in P(X)$ and $q \colon Z \to X$, it holds that $x \cdot q = y \cdot q$ implies $x = y$.*

Proof. Suppose $P$ is separated, and $x, y$ and $q$ are such that $x \cdot q = y \cdot q$. It then holds that $x \cdot q$ is $q$-invariant, and $x$ and $y$ are $q$-descendents of $x \cdot q$. So, by separatedness, $x = y$.

The converse implication, showing that separatedness follows from the statement in the proposition, is easy. □

Propositions 3.11 and 3.12 below illustrate the notion of sheaf in the case of $\mathbb{C} = \mathbb{S}\mathrm{ur}$.

Proposition 3.11. *For any set $A$ the presheaf $\underline{\mathsf{NV}}(A)$ in $\mathrm{Psh}(\mathbb{S}\mathrm{ur})$ is an atomic sheaf.*

Proof. Consider any map $c \colon \Omega' \to \Omega$ in $\mathbb{S}\mathrm{ur}$ and $c$-invariant $Y \in \underline{\mathsf{NV}}(A)(\Omega')$, i.e., function $Y \colon \Omega' \to A$. Define

$$\Omega'' := \{(\omega', \omega'') \in \Omega' \times \Omega' \mid c(\omega') = c(\omega'')\}\,,$$

and $u := (\omega', \omega'') \mapsto \omega' \colon \Omega'' \to \Omega'$ and $v := (\omega', \omega'') \mapsto \omega'' \colon \Omega'' \to \Omega'$. Clearly $c \circ u = c \circ v$. So, since $Y$ is $c$-invariant, $Y \circ u = Y \cdot u = Y \cdot v = Y \circ v$. That is, for any $(\omega', \omega'') \in \Omega''$, we have $Y(\omega') = Y(\omega'')$; i.e., for any $\omega \in \Omega$, the function $Y$ is constant on $c^{-1}(\omega)$. Define $X \in \underline{\mathsf{NV}}(A)(\Omega)$, i.e., $X \colon \Omega(A)$ by:

$$X(\omega) \ := \ Y(\omega') \text{ where } \omega' \in c^{-1}(\omega) \ . \tag{7}$$

Since $c$ is surjective, this is a good definition by the constancy property remarked above. By definition, $Y = X \circ c = X \cdot c$, so $X$ is a $c$-descendent of $Y$. It is the unique such, because, for any $c$-descendent $X$, the surjectivity of $c$ forces (7). □

Proposition 3.12. *For any finite set $\Omega$ the representable presheaf $\mathsf{y}(\Omega)$ in $\mathrm{Psh}(\mathbb{S}\mathrm{ur})$ is an atomic sheaf.*

We omit the proof, which is very similar to the previous. This last proposition asserts that the atomic topology on $\mathbb{S}\mathbb{u}\mathbb{r}$ is *subcanonical*.

As a final set of examples, it is standard (and also easily verified) that if $\underline{P_1}, \dots, \underline{P_n}$ are sheaves then the product presheaf $\underline{P_1} \times \cdots \times \underline{P_n}$ is also a sheaf, the *product sheaf*. (A similar fact applies more generally to arbitrary category-theoretic limits of sheaves.) In this statement, we introduce a notational convention we shall often adopt. We shall typically use underlined names for sheaves (as with $\underline{\mathsf{NV}}(A)$) in order to emphasise that they are sheaves not just presheaves.

Assuming the small category $\mathbb{C}$ is coconfluent, we write $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ for the full subcategory of atomic sheaves in $\mathrm{Psh}(\mathbb{C})$. While the coconfluence condition was not actually used in the definition of atomic sheaf above, it nonetheless plays a critical role. For the benefit of readers who know the relevant category theory, we reiterate that the coconfluence condtion is equivalent to the collection of atomic covers in $\mathbb{C}$ forming a Grothendieck topology, which in turn means that $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ is a *Grothendieck topos*, and the inclusion functor $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C}) \to \mathrm{Psh}(\mathbb{C})$ has a left adjoint $\mathbf{a} : \mathrm{Psh}(\mathbb{C}) \to \mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$, the *associated sheaf* functor [29]. Composing with the Yoneda functor, we obtain a functor $\mathbf{ay} : \mathbb{C} \to \mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$. Because we are working with atomic topologies, every map in $\mathbb{C}$ is a *cover*, i.e., it is mapped by $\mathbf{ay}$ to an epimorphism in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$. It thus follows from the Yoneda lemma that a necessary condition for every representable presheaf to be a sheaf (i.e., for the atomic topology to be subcanonical) is that all maps in $\mathbb{C}$ are epimorphic.

## 4 Atomic sheaf logic

For the next two sections, let $\mathbb{C}$ be an arbitrary coconfluent small category. We present a fragment of the internal logic of the topos $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ of atomic sheaves, which we will extend later with equivalence and conditional independence formulas. The fragment we consider is simply multi-sorted first-order logic. Let Sort be a collection of sorts. We assume a collection of primitive relation symbols, where each relation symbol $\mathsf{R}$ has an *arity* given as a finite sequence of sorts $\mathrm{arity}(\mathsf{R}) \in \mathrm{Sort}^*$. As in Section 2, variables $\mathsf{x}^{\mathsf{A}}$ have explicit sorts. The rules for forming atomic formulas are:

- if $\mathrm{arity}(\mathsf{R}) = \mathsf{A}_1 \dots \mathsf{A}_n$ and $\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}$ is a list of variables of the corresponding sorts, then $\mathsf{R}(\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n})$ is a formula;
- if $\mathsf{x}^{\mathsf{A}}, \mathsf{y}^{\mathsf{A}}$ have the same sort then $\mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}}$ is a formula.

The grammar for formulas extends atomic formulas with the usual constructs of first-order logic.

$$\Phi \ ::= \mathsf{R}(\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}) \mid \mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}} \mid \neg\Phi \mid \Phi \wedge \Phi \mid \Phi \vee \Phi \mid \Phi \to \Phi \mid \exists \mathsf{x}^{\mathsf{A}}.\, \Phi \mid \forall \mathsf{x}^{\mathsf{A}}.\, \Phi \ .$$

We write $\mathrm{FV}(\Phi)$ for the set of free variables of a formula $\Phi$.

*Definition 4.1 (Semantic interpretation).* A *semantic interpretation* in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ is given by a function mapping every sort $\mathsf{A}$ to an atomic sheaf $\underline{\mathsf{A}}$ (i.e., to an object of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$), and a function mapping every relation symbol $\mathsf{R}$ of arity $\mathsf{A}_1 \dots \mathsf{A}_n$ to a *subsheaf* $\underline{\mathsf{R}} \subseteq \underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_1}$.

*Definition 4.2 (Subpresheaf/subsheaf).* For $P, Q \in \mathrm{Psh}(\mathbb{C})$, we say that $Q$ is a *subpresheaf* of $P$ (notation $Q \subseteq P$) if:

- for every object $X \in \mathbb{C}$, we have $Q(X) \subseteq P(X)$, and
- for every map $f : Y \to X$ in $\mathbb{C}$ and element $x \in Q(X)$, it holds that $x \cdot_Q f = x \cdot_P f$.

For sheaves $\underline{P}, \underline{Q}$ with $\underline{Q} \subseteq \underline{P}$, we say $\underline{Q}$ is a *subsheaf* of $\underline{P}$.

The following is standard, and also easily verified.

Proposition 4.3. *Given a presheaf $P \in \mathrm{Psh}(\mathbb{C})$ and a function $Q$ mapping every object $X \in \mathbb{C}$ to a subset of $P(X)$, the function $Q$ determines a (necessarily unique) subpresheaf of $P$ if and only if:*

- *for every $f\colon Y \to X$ in $\mathbb{C}$ and $x \in Q(X)$, it holds that $x \cdot_P f \in Q(Y)$.*

*If the above holds and $P$ is also a sheaf, then the uniquely determined subpresheaf $Q$ is itself a sheaf if and only if*

- *for every $f\colon Y \to X$ in $\mathbb{C}$ and $x \in P(X)$, if $x \cdot_P f \in Q(Y)$ then $x \in Q(X)$.*

*(This characterisation is valid in the form above because we are considering only sheaves for the atomic topology.)*

The three propositions below illustrate the notion of subsheaf. The first two observe that the relations of equiextension and conditional independence of nondeterministic variables (Definitions 2.1 and 2.2) form subsheaves, a fact which will enable us to extend atomic sheaf logic with equivalence and conditional-independence relations at the end of the present section. Although the proofs are straightforward, we include them to help give readers who are not familiar with sheaves some feeling for the subsheaf property.

Proposition 4.4. *The subsets*

$$\{(X, Y) \mid X \bowtie Y\} \subseteq (\underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A))(\Omega)$$

*define a subsheaf $\bowtie_A \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A)$ via Proposition 4.3.*

Proof. For the subpresheaf property, suppose $(X, Y) \in (\underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A))(\Omega)$ are such that $(X, Y) \in \bowtie_A(\Omega)$; i.e., we have equality of images $X(\Omega) = Y(\Omega)$. Let $q\colon \Omega' \to \Omega$ be a map in $\mathbb{S}\mathrm{ur}$. We need to show that $(X \cdot q, Y \cdot q) \in \bowtie_A(\Omega')$. But indeed

$$(X \cdot q)(\Omega') = X(q(\Omega')) = X(\Omega) = Y(\Omega) = Y(q(\Omega')) = (Y \cdot q)(\Omega') \ ,$$

where the second and fourth equalities hold because $q$ is surjective.

For the subsheaf property, suppose we have $(X, Y) \in (\underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A))(\Omega)$ and map $q\colon \Omega' \to \Omega$ in $\mathbb{S}\mathrm{ur}$ such that $(X \cdot q, Y \cdot q) \in \bowtie_A(\Omega')$. By the definition of equiextension, $X(q(\Omega')) = Y(q(\Omega'))$. Because $q$ is surjective, $X(\Omega) = Y(\Omega)$. That is, $(X, Y) \in \bowtie_A(\Omega)$, as required by Proposition 4.3 to show the subsheaf property. □

Proposition 4.5. *The subsets*

$$\{(X, Y, Z) \mid X \perp\!\!\!\perp Y | Z\} \subseteq (\underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(B) \times \underline{\mathsf{NV}}(C))(\Omega)$$

*define a subsheaf $\perp\!\!\!\perp_{A,B|C} \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(B) \times \underline{\mathsf{NV}}(C)$ via Prop. 4.3.*

Proof. We leave the subpresheaf property to the reader and verify just the subsheaf property. Suppose we have $(X, Y, Z) \in (\underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(B) \times \underline{\mathsf{NV}}(C))(\Omega)$ and map $q\colon \Omega' \to \Omega$ in $\mathbb{S}\mathrm{ur}$ such that $(X \cdot q,\ Y \cdot q,\ Z \cdot q) \in \perp\!\!\!\perp_{A,B|C}(\Omega')$; i.e., $X \cdot q \perp\!\!\!\perp Y \cdot q \,|\, Z \cdot q$. We need to show that $(X, Y, Z) \in \perp\!\!\!\perp_{A,B|C}(\Omega)$; i.e., $X \perp\!\!\!\perp Y \,|\, Z$.

Suppose that there exists $\omega_1 \in \Omega$ such that $X(\omega_1) = a$ and $Z(\omega_1) = c$, and there exists $\omega_2 \in \Omega$ such that $Y(\omega_2) = b$ and $Z(\omega_2) = c$. Using the surjectivity of $q$, let $\omega_1', \omega_2' \in \Omega'$ be such that $q(\omega_1') = \omega_1$ and $q(\omega_2') = \omega_2$. Then $(X \cdot q)(\omega_1') = a$ and $(Z \cdot q)(\omega_1') = c$. Similarly $(Y \cdot q)(\omega_2') = b$ and $(Z \cdot q)(\omega_2') = c$. Because $X \cdot q \perp\!\!\!\perp Y \cdot q \,|\, Z \cdot q$, there exists $\omega' \in \Omega'$ such that $(X \cdot q)(\omega') = a$ and $(Y \cdot q)(\omega') = b$ and $(Z \cdot q)(\omega') = c$. So $\omega := q(\omega')$ satisfies $X(\omega) = a$ and $Y(\omega) = b$ and $Z(\omega) = c$, showing that indeed $X \perp\!\!\!\perp Y \,|\, Z$. □

As further interesting examples of subsheaves, we show how subsheaves of the sheaf $\underline{\mathsf{NV}}(A)$ can be defined by using modalities to lift properties $P \subseteq A$ to properties of $A$-valued nondeterministic variables.

$$
\begin{aligned}
X \Vdash_{\underline{\rho}} \mathsf{R}(\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}) &\iff (\underline{\rho}(\mathsf{x}_1^{\mathsf{A}_1}), \dots, \underline{\rho}(\mathsf{x}_n^{\mathsf{A}_n})) \in \underline{\mathsf{R}}(X) \\
X \Vdash_{\underline{\rho}} \mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}} &\iff \underline{\rho}(\mathsf{x}^{\mathsf{A}}) = \underline{\rho}(\mathsf{y}^{\mathsf{A}}) \\
X \Vdash_{\underline{\rho}} \neg\Phi &\iff X \nVdash_{\underline{\rho}} \Phi \\
X \Vdash_{\underline{\rho}} \Phi \wedge \Psi &\iff X \Vdash_{\underline{\rho}} \Phi \text{ and } X \Vdash_{\underline{\rho}} \Psi \\
X \Vdash_{\underline{\rho}} \Phi \vee \Psi &\iff X \Vdash_{\underline{\rho}} \Phi \text{ or } X \Vdash_{\underline{\rho}} \Psi \\
X \Vdash_{\underline{\rho}} \Phi \rightarrow \Psi &\iff X \nVdash_{\underline{\rho}} \Phi \text{ or } X \Vdash_{\underline{\rho}} \Psi \\
X \Vdash_{\underline{\rho}} \exists \mathsf{x}^{\mathsf{A}}.\, \Phi &\iff \exists Y.\, \exists f \colon Y \to X.\, \exists x \in \underline{\mathsf{A}}(Y).\; Y \Vdash_{(\underline{\rho}\cdot f)[\mathsf{x}^{\mathsf{A}}:=x]} \Phi \\
X \Vdash_{\underline{\rho}} \forall \mathsf{x}^{\mathsf{A}}.\, \Phi &\iff \forall Y.\, \forall f \colon Y \to X.\, \forall x \in \underline{\mathsf{A}}(Y).\; Y \Vdash_{(\underline{\rho}\cdot f)[\mathsf{x}^{\mathsf{A}}:=x]} \Phi
\end{aligned}
$$

Fig. 2. Semantics of atomic sheaf logic

Proposition 4.6. *For any set $A$ and subset $P \subseteq A$, the definitions*

$$
\begin{aligned}
\underline{\Box P}(\Omega) &:= \{X : \Omega \to A \mid \forall \omega \in \Omega.\, X(\omega) \in P\} \\
\underline{\Diamond P}(\Omega) &:= \{X : \Omega \to A \mid \exists \omega \in \Omega.\, X(\omega) \in P\}
\end{aligned}
$$

*define subsheaves $\underline{\Box P}$ and $\underline{\Diamond P}$ of $\underline{\mathrm{NV}}(A)$ in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$, by Proposition 4.3.*

This time we omit the proof, since the modality subsheaves will not play any further role in the paper. We mention, however, that the constructions in Proposition 4.6 can be used as the basis for an interesting modal extension of the first-order atomic sheaf logic of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$, in which the modalities mediate between the ordinary first-order logic of variables valued in $A$ and the sheaf logic of nondeterministic variables valued in $\underline{\mathrm{NV}}(A)$.

Returning to the general semantic interpretation of atomic sheaf logic in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$, the semantics of formulas is given by a forcing relation

$$X \Vdash_{\underline{\rho}} \Phi \ ,$$

where $\Phi$ is a formula, $X$ is an object of $\mathbb{C}$ and

$$\underline{\rho} \in \prod_{\mathsf{x}^{\mathsf{A}} \in \{\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}\}} \underline{A}(X)$$

is what we call an *$X$-assignment*: it maps every variable $\mathsf{x}^{\mathsf{A}}$ in a set $\{\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}\} \supseteq \mathrm{FV}(\Phi)$ to an element $\underline{\rho}(\mathsf{x}^{\mathsf{A}}) \in \underline{\mathsf{A}}(X)$, where $\underline{\mathsf{A}}$ is the sheaf interpreting the sort $\mathsf{A}$ of the variable.

The definition of the forcing relation is presented in Fig. 2. In the quantifier clauses, we write write $\underline{\rho} \cdot f$ for the $Y$-assignment $\mathsf{z}^{\mathsf{B}} \mapsto \underline{\rho}(\mathsf{z}^{\mathsf{B}}) \cdot f$, where $\underline{\rho}$ is an $X$-assignment and $f \colon Y \to X$ is a map in $\mathbb{C}$,

The clauses for the propositional connectives in Fig. 2 look remarkably simple-minded. They are, nonetheless, equivalent to the more involved clauses that appear in the *sheaf semantics* for logic in a sheaf topos [29]. The simplification in formulation is possible because we are working in the special case of atomic sheaves. The clauses for the existential and universal quantifier are also taken from sheaf semantics, and do not admit further simplification. Their non-local nature (they involve a change of world along $f \colon Y \to X$) is the key feature that will give atomic sheaf logic its character, when we later include equivalence and conditional independence formulas.

The next results summarise fundamental properties of the forcing relation and the logic it induces. The first is very basic, but we include it explicitly because the notion of *locality* it addresses, namely the dependency of semantics only on assignments to the free variables appearing in a formula, has been a delicate issue in the context of (in)dependence logics.

Proposition 4.7 (Locality). *For any formula* $\Phi$*, object* $X$ *of* $\mathbb{C}$ *and* $X$*-assignments* $\underline{\rho}, \underline{\rho}'$ *that are defined and coincide on* $\mathsf{FV}(\Phi)$*.*

$$X \Vdash_{\underline{\rho}} \Phi \text{ if and only if } X \Vdash_{\underline{\rho}'} \Phi.$$

Proposition 4.8 (Sheaf property). *For any formula* $\Phi$*, map* $f\colon Y \to X$ *in* $\mathbb{C}$*, and* $X$*-assignment* $\underline{\rho}$ *defined on* $\mathsf{FV}(\Phi)$*.*

$$X \Vdash_{\underline{\rho}} \Phi \text{ if and only if } Y \Vdash_{\underline{\rho}\cdot f} \Phi. \tag{8}$$

Proposition 4.8 is called the sheaf property because it is equivalent to the statement that, for every formula $\Phi$ with $\mathsf{FV}(\Phi) \subseteq \{\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n}\}$, it holds that

$$\{(x_1, \dots, x_n) \mid X \Vdash_{\mathsf{x}_i^{\mathsf{A}_i} \mapsto x_i} \Phi\} \subseteq (\underline{A_1} \times \cdots \times \underline{A_n})(X) \tag{9}$$

defines a subsheaf of $\underline{A_1} \times \ldots \times \underline{A_n}$ via Proposition 4.3.

Propositions 4.7 and 4.8 are both proved by induction on the structure of the formula. We omit the proof of Proposition 4.7, which is straightforward. Proposition 4.8 asserts that the *monotonicity* and *local character* properties from [29, §VI.7] hold. In *loc. cit.*, these properties are shown to hold for arbitrary Grothendiek topologies, whereas Proposition 4.8 concerns just the special case of atomic topologies. Nevertheless, we give a direct proof of Proposition 4.8, both for the benefit of readers who do not know general sheaf theory, and also to demonstrate the crucial role played by the coconfluence property of $\mathbb{C}$.

Proof of Proposition 4.8. By induction on the structure of $\Phi$.

In the case that $\Phi$ is an atomic formula of the form $\mathsf{R}(\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n})$, property (8) holds because $\underline{\mathsf{R}}$ is a subsheaf of $\underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_1}$.

If $\Phi$ is an equality $\mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}}$, then the left-to-right implication of (8) is immediate. For the right-to-left implication, suppose $Y \Vdash_{\underline{\rho}\cdot f} \mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}}$; that is, $\underline{\rho}(\mathsf{x}^{\mathsf{A}}) \cdot f = \underline{\rho}(\mathsf{y}^{\mathsf{A}}) \cdot f$. Since $\underline{\mathsf{A}}$ is a sheaf, hence separated, we have $\underline{\rho}(\mathsf{x}^{\mathsf{A}}) = \underline{\rho}(\mathsf{y}^{\mathsf{A}})$, by Proposition 3.10. That is, $X \Vdash_{\underline{\rho}} \mathsf{x}^{\mathsf{A}} = \mathsf{y}^{\mathsf{A}}$, as required.

The cases for the propositional connectives are all easy. We note only that, for the cases of negation and implication, in which there are negated clauses in the definition of the forcing relation (Fig. 2), the induction hypothesis is used in the opposite direction of (8) to the implication being proved.

In the case that $\Phi$ is an existentially quantified formula $\exists \mathsf{x}^{\mathsf{A}}.\, \Phi'$, we prove the left-to-right implication of (8). Accordingly, suppose that $X \Vdash_{\underline{\rho}} \exists \mathsf{x}^{\mathsf{A}}.\, \Phi'$. By the forcing clause for the existential quantifier, there exist $g : Z \to X$ and $x \in \underline{A}(Z)$ such that $Z \Vdash_{(\underline{\rho}\cdot g)[\mathsf{x}^{\mathsf{A}}:=x]} \Phi'$. By coconfluence, there exists a span $Y \xleftarrow{f'} W \xrightarrow{g'} Z$ such that $g \circ g' = f \circ f'$. By the induction hypothesis, $W \Vdash_{(\underline{\rho}\cdot g\cdot g')[\mathsf{x}^{\mathsf{A}}:=x\cdot g']} \Phi'$; i.e., $W \Vdash_{(\underline{\rho}\cdot f\cdot f')[\mathsf{x}^{\mathsf{A}}:=x\cdot g']} \Phi'$. Whence, by the forcing clause for the existential quantifier, $Y \Vdash_{\underline{\rho}\cdot f} \exists \mathsf{x}^{\mathsf{A}}.\, \Phi'$, as required. We leave the easier right-to-left implication of (8), which does not involve coconfluence, to the reader.

The proof for the universal quantifier is similar. (It can also be bypassed, by noting that the forcing interpretation of $\exists \mathsf{x}^{\mathsf{A}}.\, \Phi'$ is equivalent to that for $\neg\exists \mathsf{x}^{\mathsf{A}}.\, \neg\Phi'$.) □

It is standard that sheaf semantics, for an arbitrary Grothendieck topology, always validates intuitionistic logic. In the special case of an atomic topology, the *law of excluded middle* $\Phi \vee \neg\Phi$ is also validated, hence atomic sheaf logic is classical. In more detail, atomic topologies are special cases of dense Grothendieck topologies, and categories of sheaves for the latter are always boolean, hence classical logic is validated. This whole picture is explained in [29]. We shall not, however, assume familiarity with this abstract picture. Accordingly, we give a brief, direct explanation of how atomic sheaf logic validates classical logic.

A formula $\Phi$ is said to be *true (in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$*) under all assignments* if, for every object $X$ of $\mathbb{C}$ and $X$-assignments $\underline{\rho}$ defined on $\mathrm{FV}(\Phi)$,it holds that $X \Vdash_{\underline{\rho}} \Phi$.

Theorem 4.9 (Classical logic). *If $\Phi$ is a theorem of (multisorted) classical logic then it is true in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ *under all assignments.*

Proof (outline). It follows trivially from the definition of the forcing relation Fig. 2 that every classical propositional tautology (including every instance of the law of excluded middle $\Phi \vee \neg\Phi$) is true under all assignments (assuming, as we do, that we are working in a classical meta-theory).

The verification of the validity of the axioms and inference rules pertaining to quantifiers takes a little more work, but is not difficult. Since we are working in a special case of sheaf semantics, where such facts are anyway well established in far greater generality, we do not go into details. A sceptical reader may enjoy verifying this for themselves, using their preferred formulation of the axioms and rules of multi-sorted first-order logic. □

By Theorem 4.9, atomic sheaf logic is just multisorted first-order classical logic with a nonstandard semantics. The logic includes the equality relation, which is given a canonical interpretation. The nonstandard semantics allows relation symbols to be interpreted as arbitrary subsheaves of product sheaves. Atomic sheaf categories possess interesting such subsheaves that have no analogue in the standard semantics of first-order logic. Our main examples of this phenomenon are the two relations from the title: equivalence and conditional independence.

To end this section, we observe that, in the case of our running example $\mathrm{Sh}_{\mathrm{at}}(\mathbb{S}\mathrm{ur})$, atomic sheaf logic can incorporate the relations of equivalence and conditional independence from multiteam semantics, as in Section 2. Syntactically, we simply extend the logic with equivalence and conditional independence formulas (1) and (2), as in Section 2. Actually, we can do this simply by including equivalence and conditional independence as particular relation symbols, so the equivalence and conditional independence formulas are then instances of atomic formulas of the form $\mathsf{R}(\mathsf{x}_1^{\mathsf{A}_1}, \dots, \mathsf{x}_n^{\mathsf{A}_n})$. Specifically, for equivalence, we include relation symbols $\sim_{\mathsf{A}_1 \dots \mathsf{A}_n}$ with $\mathrm{arity}(\sim_{\mathsf{A}_1 \dots \mathsf{A}_n}) = \mathsf{A}_1 \dots \mathsf{A}_n \mathsf{A}_1 \dots \mathsf{A}_n$. Similarly, for conditional independence, we include relation symbols $\perp_{\mathsf{A}_1 \dots \mathsf{A}_l, \mathsf{B}_1 \dots \mathsf{B}_m | \mathsf{C}_1 \dots \mathsf{C}_n}$ with $\mathrm{arity}(\perp_{\mathsf{A}_1 \dots \mathsf{A}_l, \mathsf{B}_1 \dots \mathsf{B}_m | \mathsf{C}_1 \dots \mathsf{C}_n}) = \mathsf{A}_1 \dots \mathsf{A}_l \mathsf{B}_1 \dots \mathsf{B}_m \mathsf{C}_1 \dots \mathsf{C}_n$.

To interpret the extended logical language, we instantiate the semantic interpretation of Definition 4.1, in the special case of the category $\mathrm{Sh}_{\mathrm{at}}(\mathbb{S}\mathrm{ur})$, by requiring that every sort $\mathsf{A}$ is interpreted by a sheaf of nondeterministic variables $\underline{\mathsf{A}} := \underline{\mathsf{NV}}(\llbracket \mathsf{A} \rrbracket)$ for some set $\llbracket \mathsf{A} \rrbracket$. We then interpret each relation $\sim_{\mathsf{A}_1 \dots \mathsf{A}_n}$ as the subsheaf of $\left(\prod_{i=1}^n \underline{\mathsf{A}_i}\right) \times \left(\prod_{i=1}^n \underline{\mathsf{A}_i}\right)$ that is isomorphic to the subsheaf $\bowtie_{(\prod_{i=1}^n \llbracket \mathsf{A}_i \rrbracket)}$ of $\underline{\mathsf{NV}}(\prod_{i=1}^n \llbracket A_i \rrbracket) \times \underline{\mathsf{NV}}(\prod_{i=1}^n \llbracket \mathsf{A}_i \rrbracket))$ from Proposition 4.4 along the canonical isomorphism between the two product sheaves. A similar procedure, using $\perp\!\!\!\perp_{A,B|C}$ from Proposition 4.5, defines the semantics of conditional independence formulas as subsheaves. These rather convoluted definitions are equivalent to simply interpreting equivalence and conditional independence formulas directly using the conditions given in Figure 1. The benefit of the convoluted explanation in terms of subsheaves is that it presents the extended logic as a special case of general atomic sheaf logic, and in doing so explains why the meta-logical properties (locality, sheaf property, classical logic) hold for the extended logic.

## 5 Atomic equivalence

The interpretation of equivalence formulas at the end of Section 4 was given only for the relation of equiextension of nondeterministic variables, interpreted over the sheaves of nondeterministic variables in $\mathrm{Sh_{at}}(\mathbb{Sur})$ using Proposition 4.4.

Atomic sheaves offer, however, a much more general perspective on the notion of equivalence. Every category $\mathrm{Sh_{at}}(\mathbb{C})$ of atomic sheaves possesses a canonical notion of equivalence, which we call *atomic equivalence*. Specifically, for every sheaf $\underline{P}$, there is an associated subsheaf $\sim_{\underline{P}} \subseteq \underline{P} \times \underline{P}$ that is an equivalence relation in $\mathrm{Sh_{at}}(\mathbb{C})$. (A subsheaf $\underline{E} \subseteq \underline{P} \times \underline{P}$ is an *equivalence relation in* $\mathrm{Sh_{at}}(\mathbb{C})$ if $\underline{E}(X) \subseteq \underline{P}(X) \times \underline{P}(X)$ is an equivalence relation, for every $X \in \mathbb{C}$.)

Theorem 5.1 (Atomic equivalence). *Let $\underline{P}$ be any sheaf in $\mathrm{Sh_{at}}(\mathbb{C})$.*

$$\sim_{\underline{P}} (X) \; := \{(x, x') \in \underline{P}(X) \times \underline{P}(X) \mid \exists Z, \exists u, u' : Z \to X.\; x \cdot u = x' \cdot u'\}$$

*defines a subsheaf $\sim_{\underline{P}} \subseteq \underline{P} \times \underline{P}$ via Proposition 4.3. Moreover, this is an equivalence relation in $\mathrm{Sh_{at}}(\mathbb{C})$.*

Proof. For the subpresheaf property, suppose $(x, x') \in \sim_{\underline{P}} (X)$. Thus, for some $u, u' : Z \to X$, we have $x \cdot u = x' \cdot u'$. Consider any $f : Y \to X$. By coconfluence, there exist $g : W \to Z$ and $v : W \to Y$ such that $f \circ v = u \circ g$. Similarly, there exist $g' : W' \to Z$ and $v' : W' \to Y$ such that $f \circ v' = u' \circ g'$. Again by coconfluence, there exist $h : V \to W$ and $h' : V \to W'$ such that $g \circ h = g' \circ h'$. Then:

$$x \cdot f \cdot v \cdot h = x \cdot u \cdot g \cdot h = x' \cdot u' \cdot g' \cdot h' = x' \cdot f \cdot v' \cdot h' \quad .$$

So $v \circ h$ and $v' \circ h' : V \to Y$ show that $(x \cdot f, x' \cdot f) \in \sim_{\underline{P}}(Y)$.

For the subsheaf property, consider any $(x, x') \in \underline{P}(X) \times \underline{P}(X)$ and $f : Y \to X$ such that $(x \cdot f, x' \cdot f) \in \sim_{\underline{P}} (Y)$; i.e., there exist $u, u' : Z \to Y$ such that $x \cdot f \cdot u = x' \cdot f \cdot u'$. Thus $f \circ u$ and $f \circ u' : Z \to X$ show that indeed $(x, x') \in \sim_{\underline{P}}(X)$.

For the equivalence relation property, reflexivity and symmetry are trivial. For transitivity, suppose $(x, x') \in \sim_{\underline{P}} (X)$ and $(x', x'') \in \sim_{\underline{P}} (X)$; i.e., there exist $u, u' : Z \to X$ such that $x \cdot u = x' \cdot u'$ and $v, v' : Z' \to X$ such that $x' \cdot v = x'' \cdot v'$. By coconfluence, there exist $w : W \to Z$ and $w' : W \to Z'$ such that $u' \circ w = v \circ w'$. Then

$$x \cdot u \cdot w = x' \cdot u' \cdot w = x' \cdot v \cdot w' = x'' \cdot v' \cdot w' \quad .$$

So $u \circ w$ and $v' \circ w'$ show that $(x, x'') \in \sim_{\underline{P}} (X)$. ◻

In the special case of sheaves $\underline{\mathsf{NV}}(A)$ of nondeterministic variables in $\mathrm{Sh_{at}}(\mathbb{Sur})$, the canonical equivalence $\sim_{\underline{\mathsf{NV}}(A)}$ coincides with the equiextension subsheaf $\bowtie_A$ defined in Proposition 4.4.

Proposition 5.2. *The subsheaf $\sim_{\underline{\mathsf{NV}}(A)} \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A)$ in $\mathrm{Sh_{at}}(\mathbb{Sur})$ coincides with $\bowtie_A \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(A)$.*

Proof. Consider any $X, X' : \Omega \to A$.

Suppose there exist $u, u' : \Omega' \to \Omega$ such that $X \cdot u = X' \cdot u'$; i.e., $X \circ u = X' \circ u'$. Then $X \bowtie X'$ because

$$X(\Omega) \; = \; X(u(\Omega')) \; = \; X'(u'(\Omega')) \; = \; X'(\Omega) \;\; ,$$

using the surjectivity of $u$ and $u'$ for the first and last equalities.

Conversely, suppose $X \bowtie X'$, i.e., $X(\Omega) = X'(\Omega)$. Define $\Omega_A := X(\Omega)$, which is a finite nonempty set hence (up to isomorphism) an object of $\mathbb{Sur}$. The functions $X$ and $X'$ are surjective from $\Omega$ to $\Omega_A$, hence give morphisms

$$\vec{\mathsf{x}} \sim \vec{\mathsf{x}} \tag{10}$$

$$\vec{\mathsf{x}} \sim \vec{\mathsf{y}} \ \rightarrow\ \vec{\mathsf{y}} \sim \vec{\mathsf{x}} \tag{11}$$

$$\vec{\mathsf{x}} \sim \vec{\mathsf{y}} \ \wedge\ \vec{\mathsf{y}} \sim \vec{\mathsf{z}} \ \rightarrow\ \vec{\mathsf{x}} \sim \vec{\mathsf{z}} \tag{12}$$

$$\vec{\mathsf{x}} \sim \vec{\mathsf{y}} \ \rightarrow\ \pi(\vec{\mathsf{x}}) \sim \pi(\vec{\mathsf{y}}) \tag{13}$$

$$\vec{\mathsf{x}}, \mathsf{x} \sim \vec{\mathsf{y}}, \mathsf{y} \ \rightarrow\ \vec{\mathsf{x}} \sim \vec{\mathsf{y}} \tag{14}$$

$$\vec{\mathsf{x}} \sim \vec{\mathsf{y}} \ \wedge\ \Phi(\vec{\mathsf{x}}) \ \rightarrow\ \Phi(\vec{\mathsf{y}}) \tag{15}$$

$$\vec{\mathsf{x}} \sim \vec{\mathsf{x}'} \ \rightarrow\ \exists \mathsf{y}'.\, (\vec{\mathsf{x}}, \mathsf{y} \sim \vec{\mathsf{x}'}, \mathsf{y}') \tag{16}$$

Fig. 3. Axioms for equivalence

$X, X' \colon \Omega \to \Omega_A$ in $\mathbb{S}\mathrm{ur}$. By coconfluence, there exist maps $p, q \colon \Omega' \to \Omega$ such that $X \circ p = X' \circ q$. But this means that $X \cdot p = X' \cdot q$, hence $(X, X') \in \sim_{\underline{\mathsf{NV}}(A)} (\Omega)$. □

Using the notion of atomic equivalence, we give a canonical semantics to equivalence formulas (1) in any atomic sheaf topos. As at the end of Section 4, we include such formulas by considering them as given by relation symbols $\sim_{\mathsf{A}_1 \ldots \mathsf{A}_n}$ with $\mathrm{arity}(\sim_{\mathsf{A}_1 \ldots \mathsf{A}_n}) = \mathsf{A}_1 \ldots \mathsf{A}_n \mathsf{A}_1 \ldots \mathsf{A}_n$. The general semantic interpretation of sorts and relations (Definition 4.1) is then extended to require that each relation symbol $\sim_{\mathsf{A}_1 \ldots \mathsf{A}_n}$ is interpreted as the subsheaf

$$\underline{\sim_{\mathsf{A}_1 \ldots \mathsf{A}_n}} := \sim_{\underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_n}} \subseteq (\underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_n}) \times (\underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_n}) \ .$$

The forcing relation $X \Vdash_{\underline{\rho}} \mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n} \sim \mathsf{y}_1^{\mathsf{A}_1}, \ldots, \mathsf{y}_n^{\mathsf{A}_n}$ is then covered by the general clause for relation symbols R in Figure 2. This is equivalent to defining:

$$X \Vdash_{\underline{\rho}} \mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n} \sim \mathsf{y}_1^{\mathsf{A}_1}, \ldots, \mathsf{y}_n^{\mathsf{A}_n} \ \Leftrightarrow\ (\, (\underline{\rho}(\mathsf{x}_1^{\mathsf{A}_1}), \ldots, \underline{\rho}(\mathsf{x}_n^{\mathsf{A}_n}))\, ,\, (\underline{\rho}(\mathsf{y}_1^{\mathsf{A}_1}), \ldots, \underline{\rho}(\mathsf{y}_n^{\mathsf{A}_n}))\, ) \ \in\ \sim_{\underline{A_1} \times \cdots \times \underline{A_n}}(X).$$

By Proposition 5.2, the above definition generalises the multiteam interpretation of independence as the equiextension relation, in the case $\mathbb{C} = \mathbb{S}\mathrm{ur}$ and $\underline{\mathsf{A}} = \underline{\mathsf{NV}}([\![\mathsf{A}]\!])$, that was given in Section 4.

We now explore the logic of atomic equivalence, valid in any category of atomic sheaves. Fig. 3 lists formulas that are valid in our semantics, which we identify as axioms for equivalence. In them, we have abbreviated variable sequences by vectors. It is implicitly assumed that the lengths and sorts of the variable sequences match so that the equivalence formulas are legitimate. Axioms (10)–(12) simply state that $\sim$ is an equivalence relation. The next two assert structural properties. In (13), $\pi$ is any permutation of the variable sequence, and the axiom asserts that equivalence is preserved if one permutes variables in the same way on both sides. By axiom (14), equivalence is also preserved if one drops identically positioned variables from both sides. Axiom (15) is more interesting: equivalence enjoys a substitutivity property, similar to the substitutivity property of equality. However, an important restriction is hidden in the notation. It is assumed that all free variables in $\Phi$ are contained in a sequence $\vec{\mathsf{z}}$ of *distinct* variables matching in length and sorting with $\vec{\mathsf{x}}$, and hence also with $\vec{\mathsf{y}}$. We then write $\Phi(\vec{\mathsf{x}})$ for the substitution $\Phi(\vec{\mathsf{x}}/\vec{\mathsf{z}})$, and similarly for $\Phi(\vec{\mathsf{y}})$. We call (15) the *invariance principle*, as it states that properties not involving extraneous variables are invariant under equivalence. Axiom (16) is called the *transfer principle*. If $\vec{\mathsf{x}}$ and $\vec{\mathsf{x}'}$ are jointly equivalent, then for any variable $\mathsf{y}$ there exists a (necessarily equivalent) variable $\mathsf{y}'$ such that $\vec{\mathsf{x}}, \mathsf{y}$ and $\vec{\mathsf{x}'}, \mathsf{y}'$ are jointly equivalent.

This soundness of axioms (10) to (14) is straightforward. The soundness of the invariance principle (15) is a consequence of the following simple lemma.

Lemma 5.3. *For any* $\underline{P} \in \mathrm{Sh_{at}}(\mathbb{C})$ *with subsheaf* $\underline{Q} \subseteq \underline{P}$. *If* $x, x' \in \underline{P}(X)$ *are such that* $(x, x') \in \sim_{\underline{P}}(X)$ *and* $x \in \underline{Q}(X)$ *then* $x' \in \underline{Q}(X)$.

Proof. Because $(x, x') \in \sim_{\underline{P}}(X)$, we have that there exist $u, u' \colon Y \to X$ such that $x \cdot u = x' \cdot u'$. As $x \in \underline{Q}(X)$ and $\underline{Q}$ is a subpresheaf, we have $x \cdot u \in \underline{Q}(Y)$, that is $x' \cdot u' \in \underline{Q}(Y)$. Hence, since $\underline{Q}$ is a subsheaf, $x' \in \underline{Q}(X)$. □

The invariance principle follows from the lemma, because $\Phi$ defines a subsheaf of $\underline{\mathsf{A}_1} \times \cdots \times \underline{\mathsf{A}_n}$ via (9), where $\mathsf{A}_1, \ldots, \mathsf{A}_n$ are the sorts of the vector $\vec{\mathsf{x}} = \mathsf{x}_1^{\mathsf{A}_1}, \ldots, \mathsf{x}_n^{\mathsf{A}_n}$ (and hence also of $\vec{\mathsf{y}}$) in (15).

The soundness of the transfer principle (16) is a consequence of the lemma below.

Lemma 5.4. *Let* $\underline{P}, \underline{Q}$ *be sheaves and let* $x, x' \in P(X)$ *such that* $(x, x') \in \sim_{\underline{P}}(X)$. *For any* $y \in P(X)$, *there exists* $p \colon Z \to X$ *and* $y' \in \underline{Q}(Z)$ *such that* $((x \cdot p, y \cdot p), (x' \cdot p, y')) \in \sim_{\underline{P} \times \underline{Q}}(Z)$.

Proof. Since $(x, x') \in \sim_{\underline{P}}(X)$, there exist maps $u, u' \colon Y \to X$ such that $x \cdot u = x' \cdot u'$. By coconfluence, let $v, v' \colon Z \to Y$ be such that $u \circ v = u' \circ v'$. Define $p := u \circ v$ and $y' := y \cdot u \cdot v'$. By coconfluence again, let $w, w' \colon W \to Z$ be such that $w \circ v = w' \circ v'$. Then $w, w'$ show that $((x \cdot p,\ y \cdot p),\ (x' \cdot p,\ y')) \in \sim_{\underline{P} \times \underline{Q}}(Z)$, because:

$$x \cdot p \cdot w = x \cdot u \cdot v \cdot w = x' \cdot u' \cdot v' \cdot w' = x' \cdot u \cdot v \cdot w' \ = \ x' \cdot p \cdot w'$$

and

$$y \cdot p \cdot w \ = \ y \cdot u \cdot v \cdot w \ = \ y \cdot u \cdot v' \cdot w' \ = \ y' \cdot w' \ .$$

□

## 6 Independent pullbacks

Whereas Section 5 has given equivalence formulas a canonical interpretation in an arbitrary atomic sheaf topos $\mathrm{Sh_{at}}(\mathbb{C})$, the interpretation of conditional independence formulas (seemingly) requires additional structure on the generating category $\mathbb{C}$. Primary amongst this is that $\mathbb{C}$ possess *independent pullbacks*, as defined below.

*Definition 6.1 (Independent pullbacks).* A system of *Independent pullbacks* on a category $\mathbb{C}$ is given by a collection of commuting squares in $\mathbb{C}$, called *independent squares*. A commuting square

$$\begin{array}{ccc} X & \xrightarrow{\ f\ } & Y \\ {\scriptstyle g}\big\downarrow & & \big\downarrow{\scriptstyle u} \\ Z & \xrightarrow{\ v\ } & W \end{array} \tag{17}$$

is then defined to be an *independent pullback* if it is independent and it satisfies the usual pullback property restricted to independent squares; i.e., for every independent square

$$\begin{array}{ccc} X' & \xrightarrow{\ f'\ } & Y \\ {\scriptstyle g'}\big\downarrow & & \big\downarrow{\scriptstyle u} \\ Z & \xrightarrow{\ v\ } & W \end{array}$$

there exists a unique $q : X' \to X$ such that $f \circ p = f'$ and $g' \circ p = g$. The assumed collection of independent squares and derived collection of independent pullbacks are together required to satisfy the five conditions below.

**(IP1)** Every commuting square of the form below is independent.

$$\begin{array}{ccc} X & \longrightarrow & Y \\ \downarrow & & \downarrow \\ Z & \xrightarrow{\mathrm{id}_Z} & Z \end{array}$$

**(IP2)** If the left square below is independent then so is the right.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle u} \\ Z & \xrightarrow{v} & W \end{array} \qquad \begin{array}{ccc} X & \xrightarrow{g} & Z \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle v} \\ Y & \xrightarrow{u} & W \end{array}$$

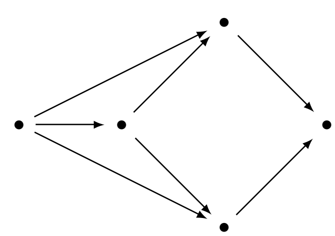
**(IP3)** If $(A)$ and $(B)$ below are independent, then so is the composite rectangle $(AB)$.

$$\begin{array}{ccccc} X & \xrightarrow{s} & Y & \xrightarrow{t} & Z \\ {\scriptstyle p}\downarrow & (A) & {\scriptstyle q}\downarrow & (B) & \downarrow{\scriptstyle r} \\ U & \xrightarrow[u]{} & V & \xrightarrow[v]{} & W \end{array} \tag{18}$$

**(IP4)** If the composite rectangle $(AB)$ above is independent and $(B)$ is an independent pullback then $(A)$ is independent.

**(IP5)** Every cospan $Y \xrightarrow{u} W \xleftarrow{v} Z$ has a completion to a commuting square (17) that is an independent pullback.

It is an easy consequence of axioms (IP1) and (IP3) that, in any commuting diagram as below, if the right square is independent then so is the outer kite.



(19)

A straightforward consequence of this property in turn is that, in any independent pullback square (17), the span $f, g$ is *jointly monic*, i.e., for all parallel pairs $s, t : V \to X$, if both $f \circ s = f \circ t$ and $g \circ s = g \circ t$ then $s = t$.

*Definition 6.2 (Descent property).* We say that independent-pullback structure has the *descent property* if, in any commuting diagram of the form (19) above, if the outer kite is independent then so is the right-hand square.

As a first (trivial) example of independent pullbacks, in any category $\mathbb{C}$ with pullbacks the collection of all commuting squares defines an independent pullback structure on $\mathbb{C}$ satisfying the descent property, for which the independent pullbacks are exactly the pullbacks. The category $\mathbb{Sur}$ (which does not have pullbacks) provides a nontrivial example.

*Definition 6.3 (Independent square in $\mathbb{Sur}$).* Define a commuting square in $\mathbb{Sur}$

$$\begin{array}{ccc} \Omega_X & \xrightarrow{p} & \Omega_Y \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle r} \\ \Omega_Z & \xrightarrow{s} & \Omega_W \end{array}$$

to be *independent* if $p \perp\!\!\!\perp q \mid r \circ p$, using conditional independence of nondeterministic variables (Definition 2.2).

Proposition 6.4. *Definition 6.3 endows* $\mathbb{Sur}$ *with independent pullback structure satisfying the descent property.*

Proof. Because the square is commuting and the maps are surjective, the condition of Definition 2.2 simplifies to: for all $\omega_Y \in \Omega_Y$ and $\omega_Z \in \Omega_Z$, we have $r(\omega_Y) = s(\omega_Z)$ implies there exists $\omega_X \in \Omega_X$ such that $p(\omega_X) = \omega_Y$ and $q(\omega_X) = \omega_Z$.

The easy verification of properties (IP1) and (IP2) is left to the reader.

For (IP3), suppose (A) and (B) in diagram (18) are independent. We show that $t \circ s \perp\!\!\!\perp p \mid r \circ t \circ s$, using the characterisation above. Accordingly, suppose $\omega_Z \in \Omega_Z$ and $\omega_U \in \Omega_U$ are such that $r(\omega_Z) = v(u(\omega_U))$. We need to find $\omega_X \in \Omega_X$ such that $t(s(\omega_X)) = \omega_Z$ and $p(\omega_X) = \omega_U$. Because $r(\omega_Z) = v(u(\omega_U))$, the independence of (B) gives us $\omega_Y \in \Omega_Y$ such that $t(\omega_Y) = \omega_Z$ and $q(\omega_Y) = u(\omega_U)$. By the latter equation and the independence of (A), there exists $\omega_X \in \Omega_X$ such that $s(\omega_X) = \omega_Y$ and $p(\omega_Z) = \omega_U$. We then have $t(s(\omega_X)) = t(\omega_Y) = \omega_Z$ as required.

For (IP4), we verify the stronger property that if the composite rectangle (AB) in diagram (18) is independent and if $t, q$ are jointly monic then (A) is independent. In the category $\mathbb{Sur}$ the joint monicity of $t, q$ means that, for all $\omega_Y, \omega'_Y \in \Omega_Y$, if both $t(\omega_Y) = t(\omega'_Y)$ and $q(\omega_Y) = q(\omega'_Y)$ then $\omega_Y = \omega'_Y$. To prove that (A) is independent, suppose $\omega_Y \in \Omega_Y$ and $\omega_U \in \Omega_U$ are such that $q(\omega_Y) = u(\omega_U)$. Then $r(t(\omega_Y)) = v(q(\omega_Y)) = v(u(\omega_U))$. So, by the independence of (AB), there exists $\omega_X \in \Omega_X$ such that $t(s(\omega_X)) = t(\omega_Y)$ and $p(\omega_X) = \omega_U$. We then have $q(s(\omega_X)) = u(p(\omega_X)) = u(\omega_U) = q(\omega_Y)$. It follows, by the joint monicity of $t, q$, that $s(\omega_X) = \omega_Y$. Together with the equation $p(\omega_X) = \omega_U$, this verifies the independence of (A).

For (IP5), the construction in the proof of Proposition 3.5 completes any cospan to an independent pullback square, as is easily verified.

We leave the straightforward verification of the descent property to the reader. □

A more abstract way of describing the independent pullback structure on $\mathbb{Sur}$ is that a commuting square in $\mathbb{Sur}$ is independent if and only if it is a weak[2] pullback in **Set**, and it is an independent pullback if and only if it is a pullback in **Set**. One can use this to give a more abstract verification that (IP1)–(IP5) and descent hold.

We end this section with some general consequences of the definition of independent pullback structure. The first such consequence is that an analogue of the pullback lemma holds for independent pullbacks.

Lemma 6.5 (Independent-pullback lemma). *Suppose* $\mathbb{C}$ *has independent pullback structure.*

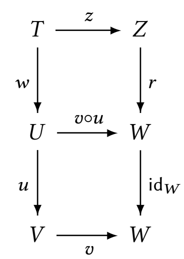
(1) *If* $(A)$ *and* $(B)$ *in* (18) *are both independent pullbacks then so is the composite rectangle* $(AB)$.
(2) *If* $(B)$ *and the composite rectangle* $(AB)$ *in* (18) *are both independent pullbacks then so is* $(A)$.

Proof. The proof has the same structure as that of the standard pullback lemma, with the additional burden of having to verify that various commuting squares are independent. We give the proof of statement 1 insofar as it involves independence properties, leaving the standard uniqueness argument and the proof of statement 2 to the reader.

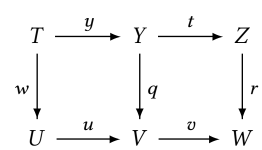
Suppose (A) and (B) are independent pullbacks. We need to verify that (AB) is an independent pullback. Accordingly, suppose that $z : T \to Z$ and $w : T \to U$ are such that the top square i the diagram below is independent. We need to

[2]A *weak limit* is a cone that enjoys the existence property but not necessarily the uniqueness property of a limit.

show that there exists a unique map $x : T \to X$ such that $p \circ x = w$ and $t \circ s \circ x = z$.




By axioms (IP1) and (IP2), the bottom square above is independent, hence, by (IP2) and (IP3), so is the composite rectange. Since (B) is an independent pullback, there exists a unique $y : T \to Y$ such that $t \circ y = z$ and $q \circ y = u \circ w$. This means that the top square in the diagram above factorises as




Since the composite rectangle is independent and the right-hand square is (B), which is an independent pullback, the left-hand square is independent by (IP4). Since (A) is an independent pullback, there exists a unique $x : T \to X$ such that $p \circ x = w$ and $s \circ x = y$, whence $t \circ s \circ x = t \circ y = z$. The proof that $x$ is the unique map satisfying $p \circ x = w$ and $t \circ s \circ x = z$ then proceeds as usual. □

By axiom (IP5), any category with independent pullbacks is *a fortiori* coconfluent, hence we can consider the category $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ of atomic sheaves, for small such $\mathbb{C}$. The remaining results in this section demonstrate a pleasing interplay between atomic sheaves and independent pullback structure. They are aimed at readers who are interested in the general category-theoretic framework. Readers keen to arrive at the atomic sheaf logic of conditional independence may prefer to skip to the next section.

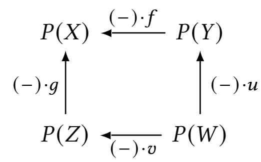
Theorem 6.6. *Suppose $\mathbb{C}$ is a small category with independent pullback structure. The following are equivalent, for every $P \in \mathrm{Psh}(\mathbb{C})$.*

(1) *$P$ is an atomic sheaf.*
(2) *$P$ maps independent squares in $\mathbb{C}$ to pullbacks in **Set**.*

Note that, by contravariance, $P$ maps an independent square of the form (17) to a pullback square in **Set** with apex $PW$.

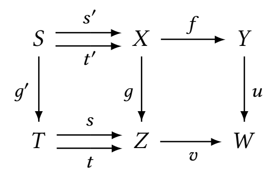
The proof of Theorem 6.6 is an adaptation to the axiomatic structure of independent pullbacks of a standard argument (see, e.g., [23, A 2.1.11(h)]) that sheaves in the Schanuel topos can be characterised as pullback preserving functors from the category $\mathbb{I}$ of finite sets and injective functions to **Set**.

Proof. For the (1) ⇒ (2) implication, suppose that $P$ is an atomic sheaf. We first show that $P$ maps independent pullbacks in $\mathbb{C}$ to pullbacks in **Set**. Consider any independent pullback of the form (17). We need to show that the square below is a pullback in **Set**.




Accordingly, let $y \in P(Y)$ and $z \in P(Z)$ be such that $y \cdot f = z \cdot g$. We need to show that there exists a unique $w \in P(W)$ such that $w \cdot u = y$ and $w \cdot v = z$.

We show that $z$ is $v$-invariant. Let $s, t : T \longrightarrow Z$ be such that $v \circ s = v \circ t$. By the independent-pullback lemma, we can construct an independent pullback of $u$ along $v \circ s = v \circ t$, either by composing the independent pullback (17) with the independent pullback of $g$ along $s$, or by composing (17) with the independent pullback of $g$ along $t$. By a straightforward argument, this means the independent pullbacks of $g$ along $s$ and $t$ can be given the same left edge $g'$ as in the diagram below, which comprises three independent pullback squares (one with $f$ and $v$, one with $s'$ and $s$ and one with $t'$ and $t$).




We have:

$$z \cdot s \cdot g' = z \cdot g \cdot s' = y \cdot f \cdot s' = y \cdot f \cdot t' = z \cdot g \cdot t' = z \cdot t \cdot g' \quad .$$

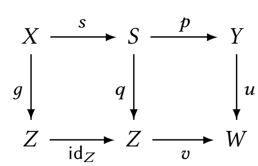
Since $P$ is separated (Definition 3.9) it follows that $z \cdot s = z \cdot t$. Thus $z$ is indeed $v$-invariant.

By the sheaf property there exists $w \in P(W)$ such that $z = w \cdot v$. Then:

$$w \cdot u \cdot f = w \cdot v \cdot g = z \cdot g = y \cdot f \quad .$$

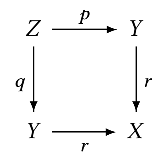
So, by separatedness, we have found $w$ such that $w \cdot u = y$ and $w \cdot v = z$. Such a $w$ is unique by separatedness.

Having established that $P$ maps independent pullbacks in $\mathbb{C}$ to pullbacks in **Set**, we show that it more generally maps all independent squares to pullbacks. Accordingly, suppose (17) is an independent square. By taking the independent pullback of $u$ along $v$, we can obtain (17) as a composite:




Since the right-hand square is an independent pullback, it is mapped by $P$ to a pullback in **Set**. The left-hand square is mapped by $P$ to a commuting square in **Set** with an identity in a position that makes it a trivial pullback. Thus $P$ maps the composite square (17) to a composition of pullbacks, hence to a pullback.

For the (2) ⇒ (1) implication, let $y \in P(Y)$ and $r\colon Y \longrightarrow X$ in $\mathbb{C}$ be such that $y$ is $r$-invariant. Consider an independent pullback of $r$ along itself




Because $y$ is $r$-invariant, $y \cdot p = y \cdot q$. By assumption, $P$ maps the above square to a pullback in **Set**. Hence, there exists a unique $x \in P(X)$ such that $y = x \cdot r$, as required by the sheaf property. □

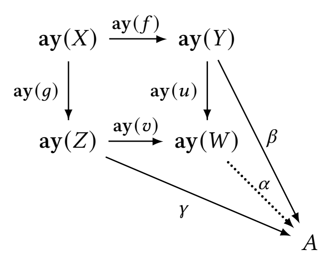
Corollary 6.7. *The functor* $\mathbf{ay} : \mathbb{C} \to \mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ *maps independent squares in* $\mathbb{C}$ *to pushouts in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$.

Proof. This is a straightforward consequence of Theorem 6.6 on account of the bijections

$$\mathrm{Sh}_{\mathrm{at}}(\mathbf{ay}(X), \underline{A}) \cong \mathrm{Psh}(\mathbf{y}(X), \underline{A}) \cong \underline{A}(X) \ ,$$

natural in $X$ and $\underline{A}$, given by the left-adjoint property of the associated sheaf functor and by the Yoneda lemma.

In more detail, consider any independent square in $\mathbb{C}$ of the form (17). Suppose we have maps $\beta$ and $\gamma$ in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ making the outside kite below commute.




The natural bijections above mean that $\beta$ and $\gamma$ correspond respectively to $y \in \underline{A}(Y)$ and $z \in \underline{A}(Z)$ satisfying $y \cdot f = g \cdot v$. Since, by Theorem 6.6, $\underline{A}$ maps the square (17) to a pullback in **Set**, there exists a unique $w \in \underline{A}(w)$ such that $w \cdot u = y$ and $w \cdot v = z$. Translating back along the natural bijections, there exists a unique map $\alpha\colon \mathbf{ay}(W) \longrightarrow \underline{A}$ such that $\alpha \cdot \mathbf{ay}(u) = \beta$ and $\alpha \cdot \mathbf{ay}(v) = \gamma$, as required. □

Corollary 6.8. *The following are equivalent for a small category* $\mathbb{C}$ *with independent pullbacks.*

(1) *Every representable presheaf is an atomic sheaf.*
(2) *Every independent square in* $\mathbb{C}$ *is a pushout.*

Proof. For the (1) $\Longrightarrow$ (2) direction, suppose every representable is an atomic sheaf. Then **ay** and **y** are naturally isomorphic, hence $\mathbf{ay} : \mathbb{C} \to \mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ is full and faithful. As a fully faithful functor, **ay** reflects (co)limits in general, and so pushouts in particular. Thus independent squares are pushouts in $\mathbb{C}$ by Corollary 6.7.

For the (2) $\Longrightarrow$ (1) direction, it holds from the definition of $\mathbf{y}(X)$ as $\mathbb{C}(-, X)$ that every representable presheaf $\mathbf{y}(X) : \mathbb{C}^{\mathrm{op}} \to \mathbf{Set}$ maps any colimit of a $D$-shaped diagram in $\mathbb{C}$ to a limit of the induced $D^{\mathrm{op}}$-shaped diagram in **Set**. In particular, $\mathbf{y}(X)$ maps pushouts in $\mathbb{C}$ to pullbacks in **Set**. So, if every independent square is a pushout in $\mathbb{C}$, then representables map independent squares to pullbacks in **Set**, and it follows from Theorem 6.6 that representables are atomic sheaves. □

## 7 Atomic conditional independence

A main goal of this section is to define a canonical subsheaf $\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}} \subseteq \underline{A} \times \underline{B} \times \underline{C}$ representing a conditional independence relation between sheaves $\underline{A}, \underline{B}, \underline{C}$ in atomic sheaf toposes $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$. To achieve this, we shall require that $\mathbb{C}$ have independent pullbacks. We shall also need to assume that the sheaves $\underline{A}, \underline{B}, \underline{C}$ enjoy the special property of having *supports*, a notion that we now define.

*Definition 7.1 (Supports).* A *representable factorisation* of an element $x \in P(X)$, where $P \in \mathrm{Psh}(\mathbb{C})$, is given by a triple $(Y, q, y)$ such that: $q : X \to Y$ is a map in $\mathbb{C}$, we have $y \in P(Y)$ and $x = y \cdot q$. A *morphism* from one representable factorisation $(Y, q, y)$ of $x$ to another $(Y', q', y')$ is given by a map $r\colon Y \to Y'$ in $\mathbb{C}$ such that $r \circ q = q'$ and $y' \cdot r = y$. A representable factorisation $(Y, q, y)$ is called a *support* for $x$ when it is a terminal object in the category of representable factorisations of $x$. A presheaf $P \in \mathrm{Psh}(\mathbb{C})$ is said to have *supports* if, for every $X \in \mathbb{C}$, it holds that every $x \in P(X)$ has a support.

For readers familiar with the *category of elements* $\int P$ of a presheaf $P$, we remark that a support for $x \in P(X)$ is the same thing as a terminal object in the co-slice category $(X, x)/\int P$. This elegant formulation is used as the definition of support in [28] (there called *minimal support*).

Lemma 7.2. *Suppose all maps in $\mathbb{C}$ are epimorphic and that $P \in \mathrm{Psh}(\mathbb{C})$ has supports. Then, for any $x \in P(X)$ and map $Y \xrightarrow{q} X$ in $\mathbb{C}$, a representable factorisation $(Z, t, z)$ of $x$ is a support for $x$ if and only if $(Z, t \circ q, z)$ is a support for $x \cdot q$.*

Proof. Suppose $(Z, t, z)$ is a support for $x$. Let $(W, u, w)$ be a support for $x \cdot q$. Because $(Z, t \circ q, z)$ is a representable factorisation of $x \cdot q$, there exists a unique map $r : Z \to W$ that is a morphism from $(Z, t \circ q, z)$ to $(W, u, w)$. Then $(W, r \circ t, w)$ is a representable factorisation of $x$. So there exists a unique map $s : W \to Z$ that is a morphism from $(W, r \circ t, w)$ to $(Z, t, z)$. That is, $r$ is the unique map such that $r \circ t \circ q = u$ and $w \cdot r = z$, and $s$ is the unique map such that $s \circ r \circ t = t$ and $z \cdot s = w$. Since $t$ is an epi, the equation $s \circ r \circ t = t$ implies $s \circ r = \mathrm{id}_Z$. Then we have $t \circ q = s \circ r \circ t \circ q = s \circ u$, which means that $s$ is a morphism of $x \cdot q$ factorisations from $(W, u, w)$ to $(Z, t \circ q, z)$. So $r \circ s$ is a morphism from $(W, u, w)$ to itself. Since $(W, u, w)$ is the terminal $x \cdot q$ factorisation, $r \circ s = \mathrm{id}_W$. Thus $r$ and $s$ are mutual inverses, and $t$ is an isomorphism of $x \cdot q$ factorisations from $(W, u, w)$ to $(Z, t \circ q, z)$. Hence $(Z, t \circ q, z)$ is also a support for $x \cdot q$.

Conversely, suppose $(Z, t \circ q, z)$ is a support for $x \cdot q$. Let $(V, v, w)$ be a representable factorisation of $x$. Then $(V, v \circ q, w)$ is a representable factorisation of $x \cdot q$. So there exists a unique map $s : W \to Z$ that is a morphism from $(V, v \circ q, w)$ to $(Z, t \circ q, z)$, that is, $s \circ v \circ p = t \circ p$ and $z \cdot s = w$. Since $p$ is an epi, $s \circ v = t$, and so $s$ is a (clearly unique) morphism from $(V, v, w)$ to $(Z, t, z)$. This shows that $(Z, t, z)$ is a support for $x$. ◻

We shall also require presheaves with supports to be closed under finite products. This follows from a further property of the category $\mathbb{C}$ (dual to the existence of $\mathcal{M}$-*images* as defined in [41, §5.1]).

*Definition 7.3 (Pairings).* A *pair factorisation* of a span $Y \xleftarrow{f} X \xrightarrow{g} Z$ in a category $\mathbb{C}$ is given by $(X', q', f', g')$ where $q' : X \to X'$ and $Y \xleftarrow{f'} X' \xrightarrow{g'} Z$ are maps in $\mathbb{C}$ that satisfy $f' \circ q' = f$ and $g' \circ q' = g$. A *morphism* from a pair factorisation $(X', q', f', g')$ of $f, g$ to another $(X'', q'', f'', g'')$ is a map $r : X' \to X''$ in $\mathbb{C}$ such that $r \circ q' = q''$, $f'' \circ r = f'$ and $g'' \circ r = g'$. A pair factorisation $(X', q', f', g')$ is said to be a *pairing* for $f, g$ if it is a terminal object in the category of pair factorisations of $f, g$. We say that the category $\mathbb{C}$ has *pairings* if every span $f, g$ has a pairing.

Proposition 7.4. *Suppose all maps in $\mathbb{C}$ are epimorphic and that $\mathbb{C}$ has pairings. If $P, Q \in \mathrm{Psh}(\mathbb{C})$ both have supports, then so does the product $P \times Q$.*

Proof. Consider any element $(x, y) \in (P \times Q)(X)$. Let $(U, u, x')$ be support for $x$ and $(V, v, y')$ support for $y$. Let $(W, w, u', v')$ be a pairing for $u, v$. We show that $(W, w, (x' \cdot u', y' \cdot v'))$ is support for $(x, y)$.

Let $(Z, t, (x'', y''))$ be any representable factorisation of $(x, y)$. Then $(Z, t, x'')$ is a representable factorisation of $x$, so there exists a unique map $r : Z \longrightarrow U$ that is a morphism from $(Z, t, x'')$ to $(U, u, x')$, i.e., such that $r \circ t = u$ and $x' \cdot r = x''$. Similarly, there exists a unique map $s : Z \longrightarrow V$ such that $s \circ t = v$ and $y' \cdot s = y''$. Since $(X, t, r, s)$ is a pair factorisation of $u, v$, there exists a unique $w' : Z \to W$ such that $w' \circ t = w$ and $u' \circ w' = r$ and $v' \circ w' = s$. We claim that $w' : Z \to W$ is the unique morphism from $(Z, t, (x'', y''))$ to $(W, w, (x' \cdot u', y' \cdot v'))$. We have seen that $w' \circ t = w$. Since $t$ is epimorphic, this determines $w'$ uniquely. It also holds that $x' \cdot u' \cdot w' = x' \cdot r = x''$ and $y' \cdot v' \cdot w' = y' \cdot s = y''$. So $w'$ is indeed a morphism of representable factorisations. □

We explore the above properties in the case of our running example $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$.

Proposition 7.5. *In $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$ every sheaf of the form $\underline{\mathsf{NV}}(A)$ has supports.*

Proof. Consider any $X \in \underline{\mathsf{NV}}(A)(\Omega)$, i.e., $X \colon \Omega \to A$. Factorise $X$ as a composite $\Omega \xrightarrow{p} \Omega' \xrightarrow{X'} A$ where $p$ is surjective and $X'$ injective. It is easy to verify that $(\Omega', p, X')$ is a support for $X$. □

Proposition 7.6. *The category $\mathbb{Sur}$ has pairings.*

Proof. Consider a span $\Omega_Y \xleftarrow{p} \Omega \xrightarrow{q} \Omega_Z$ in $\mathbb{Sur}$. Factorise the function $(p, q) : \Omega \to \Omega_Y \times \Omega_Z$ as $\Omega \xrightarrow{r} \Omega' \xrightarrow{(p', q')} \Omega_Y \times \Omega_Z$ where $r$ is surjective and $(p', q')$ injective. Then $(\Omega', r, p', q')$ is a pairing of $p, q$. □

Proposition 7.5 is in fact subsumed by a much more general result, which however has a far more involved proof..

Theorem 7.7. *In $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$ every sheaf has supports.*

Because this theorem is not central to the development, we relegate its proof to Appendix A

We henceforth impose global assumptions on our category $\mathbb{C}$.

*Definition 7.8.* We say that a small category $\mathbb{C}$ has the *requisite structure* if: every map in $\mathbb{C}$ is an epimorphism, it has pairings, and it has independent-pullback structure satisfying the descent property.

The reason for imposing the assumption that every map is an epimorphism is that it allows us to apply Lemma 7.2 and Proposition 7.4. Because the role of $\mathbb{C}$ is to serve as the gateway to the category $\mathrm{Sh}_{\mathrm{at}}\mathbb{C}$ of atomic sheaves, this assumption is very mild. As discussed at the end of Section 3, it is weaker than assuming that all representable presheaves are atomic sheaves. Moreover, every atomic sheaf topos is equivalent to $\mathrm{Sh}_{\mathrm{at}}\mathbb{C}$ for some coconfluent small category $\mathbb{C}$ in which every map is an epimorphism.

Since it is obvious that every map in $\mathbb{Sur}$ is epimorphic, Propositions 7.6 and 6.4 show that the category $\mathbb{Sur}$ has the requisite structure.

**For the remainder of the present section, let $\mathbb{C}$ be a small category with the requisite structure.**

We define a general *atomic conditional independence* relation for atomic sheaves $\underline{A}, \underline{B}, \underline{C}$ on $\mathbb{C}$ with supports. For any $X \in \mathbb{C}$, define

$$\perp\!\!\!\perp_{A,B|C}(X) \subseteq (\underline{A} \times \underline{B} \times \underline{C})(X) \tag{20}$$

to consist of those triples $(x, y, z) \in (\underline{A} \times \underline{B} \times \underline{C})(X)$ that satisfy the condition: there exists an independent square $r \circ p = s \circ q$ in $\mathbb{C}$ (as in the diagram below), and there exist elements $(x', u') \in (\underline{A} \times \underline{C})(X_x)$, and $(y', v') \in (\underline{A} \times \underline{C})(X_y)$ and $z' \in \underline{C}(X_z)$ such that $x' \cdot p = x$ and $y' \cdot q = y$ and $z' \cdot r = u'$ and $z' \cdot s = v'$ and $(X_z,\ r \circ p,\ z')$ is support for $z$. The data in the condition above is illustrated by the hybrid diagram below, where the symbol $\perp\!\!\!\perp$ indicates that the square is independent.

$$\begin{array}{ccccc} X & \xrightarrow{p} & X_x & \xrightarrow{(x',u')} & \underline{A}\times\underline{C} \\ {\scriptstyle q}\downarrow & \perp\!\!\!\perp & \downarrow{\scriptstyle r} & & \\ X_y & \xrightarrow[s]{} & X_z & & \downarrow{\scriptstyle \pi_2} \\ {\scriptstyle (y',v')}\downarrow & & & \searrow{\scriptstyle z'} & \\ \underline{B}\times\underline{C} & & \xrightarrow[\pi_2]{} & & \underline{C} \end{array} \tag{21}$$

The above diagram is *hybrid* in the sense that the arrows in it represent three distinct kinds of entity. Arrows of the form $X \to Y$ between objects of $\mathbb{C}$ represent maps in $\mathbb{C}$. Arrows of the form $X \to \underline{A}$, from an object $X$ of $\mathbb{C}$ to a sheaf $\underline{A}$, represent elements of the set $\underline{A}(X)$. Arrows of the form $\underline{A} \to \underline{B}$ between sheaves represent maps in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$. By the Yoneda lemma, such hybrid diagrams can equivalently be interpreted as ordinary diagrams in the presheaf category $\mathrm{Psh}(\mathbb{C})$, with objects $X$ of $\mathbb{C}$ being interpreted as representable presheaves $\mathbf{y}X$.

Lemma 7.9. *In the definition of* $\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(X)$*, we can, without loss of generality, choose the data so that* $(X_x, p, (x',u'))$ *is support for* $(x,z)$ *and* $(X_y, q, (y',v'))$ *is support for* $(y,z)$*.*

Proof. Suppose we have:

$$\begin{array}{ccccc} X & \xrightarrow{p} & X_x & \xrightarrow{(x',u')} & \underline{A}\times\underline{C} \\ {\scriptstyle q}\downarrow & \perp\!\!\!\perp & {\scriptstyle r}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ X_y & \xrightarrow[s]{} & X_z & \xrightarrow[z']{} & \underline{C} \end{array} \tag{22}$$

where $(X_z,\ r \circ p,\ z')$ is a support for $z$. Let $(X'_x, t, (x'',u''))$ be support for $(x',u') \in (\underline{A}\times\underline{C})(X_x)$. Then $(X'_x, t, u'')$ is a representable factorisation of $z' \cdot r$. By Lemma 7.2, $(X_z, r, z')$ is a support for $z' \cdot r$. So there exists $r' : X'_x \to X_z$ such that $r' \circ t = r$ and $z' \cdot r' = u''$. We have thus obtained the data in the hybrid diagram below.

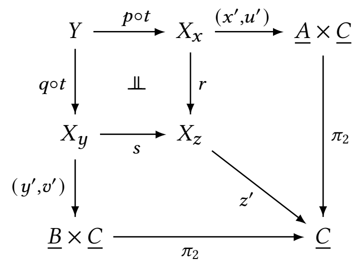
$$\begin{array}{ccccccc} X & \xrightarrow{p} & X_x & \xrightarrow{t} & X'_x & \xrightarrow{(x'',u'')} & \underline{A}\times\underline{C} \\ {\scriptstyle q}\downarrow & \perp\!\!\!\perp & {\scriptstyle r}\downarrow & \perp\!\!\!\perp & {\scriptstyle r'}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ X_y & \xrightarrow[s]{} & X_z & \xrightarrow[\mathrm{id}_{X_z}]{} & X_z & \xrightarrow[z']{} & \underline{C} \end{array} \tag{23}$$

Morover, by Lemma 7.2, it holds that $(X'_x,\ t \circ p, (x'',u''))$ is support for $(x',u') \cdot p = (x,z)$. We have thus shown how diagram (22), gives rise to diagram (23), in which the composite independent square satisfies the desired support property for $(x,z)$.

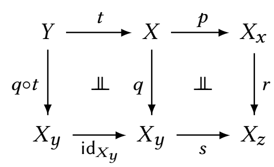
By starting with the new diagram and repeating the same argument in a vertical rather than horizontal direction, one similarly satisfies the required support property for $(y,z)$. ☐

Theorem 7.10. *Suppose* $\underline{A}, \underline{B}, \underline{C}$ *are atomic sheaves with supports. Then* $\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(X) \subseteq (\underline{A}\times\underline{B}\times\underline{C})(X)$ *defines a subsheaf via Prop. 4.3.*

Proof of Theorem 7.10. We first show that $\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}$ is a subpresheaf. Suppose $(x,y,z) \in \perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(X)$ and $t : Y \longrightarrow X$ is a map in $\mathbb{S}$; that is, we have the data in diagram (21) and $(X_z,\ r \circ p,\ z')$ is a support for $z$. We need to show that $(x \cdot r,\ y \cdot r,\ z \cdot r) \in \perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(Y)$. This holds on account of the data illustrated below.

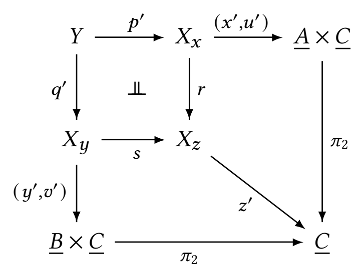


Indeed, $(X_z,\ r \circ p \circ t,\ z')$ is support for $z \circ r$ on account of Lemma 7.2, and the marked square is independent since it is a composition of two independent squares:



For the subsheaf property, suppose $(x, y, z) \in (\underline{A} \times \underline{B} \times \underline{C})(X)$ and $(x \cdot t,\ y \cdot t,\ z \cdot t) \in \perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(Y)$ where $t : Y \longrightarrow X$ is a map in $\mathbb{S}$. We need to show that $(x, y, z) \in \perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(X)$.

The assumption gives us the data below



where, $x' \cdot p' = x \cdot t$ and $y' \cdot q' = y \cdot t$ and $(X_z,\ r \circ p',\ z')$ is support for $z \cdot t$. By Lemma 7.9, we can assume that $(X_x, p', (x', u'))$ is support for $(x \cdot t, z \cdot t)$ and $(X_y, q', (y', v'))$ is support for $(y \cdot t, z \cdot t)$. Since $(X, t, (x, z))$ is a representable factorisation of $(x \cdot t, z \cdot t)$, we have $p' = p \circ t$ and $(x, z) = (x' \cdot p,\ u' \cdot p)$, for some $p \colon X \longrightarrow X_x$. Similarly, $q' = q \circ t$ and $(y, z) = (y' \cdot q,\ v' \cdot q)$, for some $q \colon X \longrightarrow X_y$. Then

$$r \circ p \circ t = t \circ p' = s \circ q' = s \circ q \circ t \ .$$

Since $t$ is epimorphic, $r \circ p = s \circ q$ is a commuting square, which is independent by the descent property. Accordingly, we have precisely the data in diagram (21). Moreover, since $(X_z,\ r \circ p \circ t,\ z') = (X_z,\ r \circ p',\ z')$ is support for $z \cdot t$, it follows from Lemma 7.2 that $(X_z,\ r \circ p,\ z')$ is support for $z$, as required. □

In the special case of sheaves $\underline{\mathsf{NV}}(A)$ of nondeterministic variables in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$, the general atomic conditional independence defined above coincides with the multiteam conditional independence from Proposition 4.5.

Proposition 7.11. *The subsheaf*

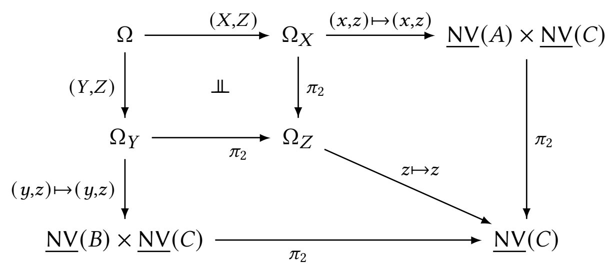
$$\perp\!\!\!\perp_{\underline{\mathsf{NV}}(A),\underline{\mathsf{NV}}(B)|\underline{\mathsf{NV}}(C)} \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(B) \times \underline{\mathsf{NV}}(C)$$

*in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$ *coincides with* $\perp\!\!\!\perp_{A,B|C} \subseteq \underline{\mathsf{NV}}(A) \times \underline{\mathsf{NV}}(B) \times \underline{\mathsf{NV}}(C)$ *from Proposition 4.5.*

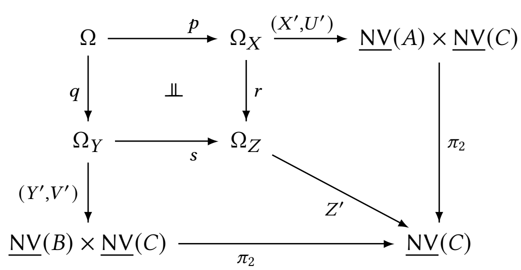
Proof. Suppose $X : \Omega \to A,\ Y : \Omega \to B,\ Z : \Omega \to C$ are nondeterministic variables such that $X \perp\!\!\!\perp Y \mid Z$, according to Definition 2.2. Define

$$\begin{aligned} \Omega_X &:= \{(x, z) \in A \times C \mid \exists \omega \in \Omega.\ x = X(\omega) \text{ and } z = Z(\omega)\} \\ \Omega_Y &:= \{(y, z) \in B \times C \mid \exists \omega \in \Omega.\ y = Y(\omega) \text{ and } z = Z(\omega)\} \\ \Omega_Z &:= Z(\Omega) \end{aligned}$$

Then the hybrid diagram below, shows that $(X, Y, Z)$ belongs to the atomic conditional independence $\perp\!\!\!\perp_{\underline{\mathsf{NV}}(A),\underline{\mathsf{NV}}(B)|\underline{\mathsf{NV}}(C)}(\Omega)$, since $(\Omega_Z,\ Z,\ z \mapsto z)$ is support for $Z$, by the definition of $\Omega_Z$.




Conversely, suppose $(X, Y, Z) \in \perp\!\!\!\perp_{\underline{\mathsf{NV}}(A),\underline{\mathsf{NV}}(B)|\underline{\mathsf{NV}}(C)}(\Omega)$. That is, we have the data in the hybrid diagram below, where $X' \cdot p = X$ and $Y' \cdot q = Y$ and $(\Omega_Z,\ r \circ p,\ Z')$ is support for $Z$.




We show that $X \perp\!\!\!\perp Y \mid Z$, according to Definition 2.2. Suppose we have $\omega', \omega'' \in \Omega$ such that $X(\omega') = a$ and $Z(\omega') = c$ and $Y(\omega'') = b$ and $Z(\omega'') = c$. Then

$$Z'(r(p(\omega'))) = Z(\omega') = Z(\omega'') = Z'(s(q(\omega''))) \ .$$

Since $(\Omega_Z,\ r \circ p,\ Z')$ is support for $Z$, the function $Z' : \Omega_Z \to C$ is injective, hence $r(p(\omega')) = s(q(\omega''))$. Since the top-left square is independent, there exists $\omega \in \Omega$ such that $p(\omega) = p(\omega')$ and $q(\omega) = q(\omega'')$. Then $X(\omega) = X'(p(\omega)) = X'(p(\omega')) = X(\omega') = a$. Similarly, $Y(\omega) = b$, and $Z(\omega) = c$. □

We now turn to the extension of the atomic sheaf logic of Sections 4 and 5 with conditional independence formulas (2). Once again, we view this extension as being obtained by including a family of relation symbols. In this case we add relations $\perp_{\vec{\mathsf{A}},\vec{\mathsf{B}}|\vec{\mathsf{C}}}$, and require that each such relation is interpreted as the subsheaf

$$\perp\!\!\!\perp_{\underline{\vec{\mathsf{A}}},\underline{\vec{\mathsf{B}}}|\underline{\vec{\mathsf{C}}}} \subseteq \underline{\vec{\mathsf{A}}} \times \underline{\vec{\mathsf{B}}} \times \underline{\vec{\mathsf{C}}} \ ,$$

$$\vec{x} \perp \vec{y} \mid \vec{z} \;\rightarrow\; \pi(\vec{x}) \perp \pi'(\vec{y}) \mid \pi''(\vec{z}) \tag{24}$$

$$\vec{x} \perp \vec{y} \mid \vec{y} \tag{25}$$

$$\vec{x} \perp \vec{y} \mid \vec{z} \;\rightarrow\; \vec{y} \perp \vec{x} \mid \vec{z} \tag{26}$$

$$\vec{x} \perp \vec{y}, \vec{z} \mid \vec{w} \;\rightarrow\; \vec{x} \perp \vec{y} \mid \vec{w} \tag{27}$$

$$\vec{x} \perp \vec{y}, \vec{z} \mid \vec{w} \;\rightarrow\; \vec{x} \perp \vec{y} \mid \vec{z}, \vec{w} \tag{28}$$

$$\vec{x} \perp \vec{y} \mid \vec{z}, \vec{w} \;\wedge\; \vec{x} \perp \vec{z} \mid \vec{w} \;\rightarrow\; \vec{x} \perp \vec{y}, \vec{z} \mid \vec{w} \tag{29}$$

$$\exists \vec{y}.\; (\vec{y}, \vec{w} \sim \vec{x}, \vec{w} \;\wedge\; \vec{y} \perp \vec{z} \mid \vec{w}) \tag{30}$$

Fig. 4. Axioms for conditional independence

where we write, e.g., $\underline{\vec{A}}$ for the product $\prod_{i=1}^{n} \underline{A_n}$, where $\vec{A}$ is the vector of sorts $A_1, \dots, A_n$. To ensure that $\perp\!\!\!\perp_{\underline{\vec{A}}, \underline{\vec{B}} \mid \underline{\vec{C}}}$ is well defined, we require that every sort A is interpreted as a sheaf $\underline{A}$ with supports.

Figure 4 lists formulas valid in this semantics that we single out as a suitable list of axioms for reasoning about conditional independence. Axiom (24) asserts that conditional independence is preserved under permutations within each of the three lists of variables involved. This axiom, together with axioms (25)–(29) are all standard axioms for conditional independence, appearing in closely related forms in [7, Theorem 3.1 and Lemmas 4.1–4.3], in [40, Theorem 1] and in the work of Pearl, Paz and Geiger [15, 16, 33] (in which only conditional independence statements of the restricted form $\vec{x} \perp \vec{y} \mid \vec{z}$ for three disjoint sets of variables $\vec{x}$, $\vec{y}$ and $\vec{z}$ are considered). The axioms appear more explicitly in their present form in Dawid's axioms for the notion of *separoid* [8]. We leave the straightforward verification of the soundness of axioms (24)–(27) to the reader. The soundness of axioms (28) and (29) is more technical. To avoid encumbering the main development with these technical proofs, they are given in Appendix B.

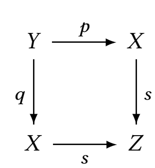
Whereas axioms (24)–(29) concern conditional independence in isolation, axiom (30) captures the interaction between conditional independence and atomic equivalence. Axiom (30) makes essential use of the existential quantifier of atomic sheaf logic to capture a key first-order property: given variables $\vec{x}, \vec{z}, \vec{w}$ one can always find variables $\vec{y}$ that are conditionally independent from $\vec{z}$ given $\vec{w}$, but such that $\vec{y}, \vec{w}$ is jointly equivalent to $\vec{x}, \vec{w}$. We call this property the *independent existence principle*: independent variables with any desired distribution always exist. The validity of the principle of independent existence (30) is established by Lemma 7.12 below.

Lemma 7.12. *Given $x \in \underline{A}(X)$, $z \in \underline{B}(X)$ and $w \in \underline{C}(X)$, there exist $p \colon Y \to X$ and $y \in \underline{A}(Y)$ such that*

$$((y,\ w \cdot p),\ (x \cdot p,\ w \cdot p)) \in\ \sim_{\underline{A} \times \underline{C}}\ (Y) \tag{31}$$

$$(y,\ z \cdot p,\ w \cdot p) \in\ \perp\!\!\!\perp_{\underline{A}, \underline{B} \mid \underline{C}}(Y)\ \ . \tag{32}$$

Proof. Let $(Z, s, w')$ be support for $w$, and consider the independent pullback of $s : X \to Z$ along itself:




We have:

$$w \cdot p = w' \cdot s \cdot p = w' \cdot s \cdot q = w \cdot q\ \ .$$

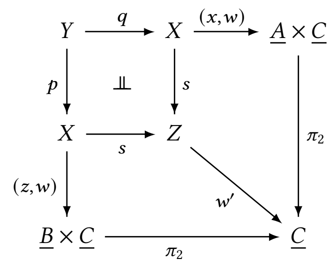
Define $y := x \cdot q$.

By the independent pullback above, there is a unique map $t : Y \longrightarrow Y$ such that $p \circ t = q$ and $q \circ t = p$. So:

$$(y,\ w \cdot p) = (x \cdot q,\ w \cdot q) \ = \ (x \cdot p \circ t,\ w \circ p \circ t) \ .$$

Thus the pair $\mathrm{id}_Y, t \colon Y \longrightarrow Y$ shows that (31) holds.

For the independence statement, we have:




The first component of the top side is $x \cdot q = y \in \underline{A}(Y)$. The first component of the left side is $z \cdot p \in \underline{B}(Y)$. Moreover, by Lemma 7.2, $(Z,\ s \circ p,\ w')$ is support for $w \cdot p$. Thus we indeed have (32). □

As an interesting consequence of the axioms, we prove that existence properties are preserved under conditional independence, in the sense of the result below. This provides a first-order reasoning principle for conditional independence, whose scope potentially extends beyond atomic sheaf logic to more general contexts in which there is a conditional independence relation but no analogue of the relation $\sim$ of atomic equivalence.

Theorem 7.13 (Existence preservation). *The schema below follows from the axioms in Figs. 3 and 4.*

$$(\exists \vec{\mathsf{y}}.\ \Phi(\vec{\mathsf{x}}, \vec{\mathsf{y}}, \vec{\mathsf{w}})) \ \longrightarrow \ \forall \vec{\mathsf{z}}.\ (\, \vec{\mathsf{x}} \perp \vec{\mathsf{z}} \,|\, \vec{\mathsf{w}} \ \longrightarrow \ \exists \vec{\mathsf{y}}.\ (\, \vec{\mathsf{x}}, \vec{\mathsf{y}} \perp \vec{\mathsf{z}} \,|\, \vec{\mathsf{w}} \ \wedge \ \Phi(\vec{\mathsf{x}}, \vec{\mathsf{y}}, \vec{\mathsf{w}})))$$

*Here we adopt the same convention as in the invariance principle. In $\Phi(\vec{\mathsf{x}}, \vec{\mathsf{y}}, \vec{\mathsf{w}})$ every free variable in $\Phi$ has been substituted by one of the variables in $\vec{\mathsf{x}}, \vec{\mathsf{y}}, \vec{\mathsf{w}}$.*

Proof. Let $\vec{\mathsf{y}}$ be such that

$$\Phi(\vec{\mathsf{x}}, \vec{\mathsf{y}}, \vec{\mathsf{w}})\ . \tag{33}$$

Consider any $\vec{\mathsf{z}}$. By the independent existence principle (30), there exists $\vec{\mathsf{y}'}$ such that

$$\vec{\mathsf{y}'}, \vec{\mathsf{x}}, \vec{\mathsf{w}} \sim \vec{\mathsf{y}}, \vec{\mathsf{x}}, \vec{\mathsf{w}} \tag{34}$$

and

$$\vec{\mathsf{y}'} \perp \vec{\mathsf{z}} \,|\, \vec{\mathsf{x}}, \vec{\mathsf{w}}\ . \tag{35}$$

Suppose

$$\vec{\mathsf{x}} \perp \vec{\mathsf{z}} \,|\, \vec{\mathsf{w}}\ . \tag{36}$$

Then (35) and (36) combine to give $\vec{\mathsf{x}}, \vec{\mathsf{y}'} \perp \vec{\mathsf{z}} \,|\, \vec{\mathsf{w}}$, by the axioms for conditional independence.

Further, (33) and (34) combine to give $\Phi(\vec{\mathsf{x}}, \vec{\mathsf{y}'}, \vec{\mathsf{w}})$, by the invariance principle (15). □

## 8 Probability sheaves

In this long section, we present another instance of our axiomatic structure: atomic sheaves over *standard Borel probability spaces*. The idea is that such spaces take the role of sample spaces, and random variables over such sample spaces collectively form an atomic sheaf. More precisely, for any standard Borel space $A$, we shall obtain a sheaf $\underline{\mathrm{RV}}(A)$ of all $A$-valued random variables. For this aim, the standard-Borel assumption serves three purposes. Firstly, it is sufficiently general that it encompass both discrete and continuous probability. Secondly, it provides a *small* category of sample spaces to build atomic sheaves over. Finally, it also provides useful technical machinery (such as *disintegrations* of random variables), which would be unavailable in general if arbitrary probability and measurable spaces were used. This machinery is essential in showing that the category of sample spaces has independent pullback structure. When interpreted over the sheaves of random variables $\underline{\mathrm{RV}}(A)$, atomic sheaf logic provides logical principles governing the relations of almost sure equality, of equality in distribution and of conditional independence with its standard probabilistic meaning, since these three relations are respectively encapsulated as equality, atomic equivalence and atomic conditional independence in the logic.

In order to fully understand the technical development in the present section, it is necessary to have some background in probability and measure theory. Nevertheless, we try to also explain the main ideas informally, so help readers without the necessary background to follow the line of development at a high level.

Standard Borel spaces will be the value spaces of random variables, and they will also be the structures over which we build sample spaces.

*Definition 8.1 (Standard Borel space).* A *standard Borel space* (SBS) is a measurable space $(A, \mathcal{B}_A)$ where $A$ is a Borel subset of a Polish space $T$ (i.e., a complete separable metric space) and $\mathcal{B}_A$ is the $\sigma$-algebra $\{S \cap A \mid S \subseteq T \text{ is Borel}\}$. A *morphism* of standard Borel spaces from $(A, \mathcal{B}_A)$ to $(B, \mathcal{B}_B)$ is a function $f : A \to B$ that is *measurable*, i.e., $f^{-1}(S) \in \mathcal{B}_A$ for all $S \in \mathcal{B}_B$.

When $(A, \mathcal{B}_A)$ is a standard Borel space, we shall refer to the sets in $\mathcal{B}_A$ as the *Borel subsets* of $A$, which is justified because $A$ can always itself be given a Polish topology in which $\mathcal{B}_A$ is the Borel $\sigma$-algebra. As is well known, the image $f(C)$ of a Borel subset $C \subseteq A$ under a measurable function $f : A \to B$, where $(B, \mathcal{B}_B)$ is also standard Borel, need not itself be a Borel subset of $B$, but $f(C)$ is always an *analytic* subset of $B$.

On the one hand, the collection of standard Borel spaces is very rich, as it incorporates most measurable spaces that arise naturally in mathematics. On the other, it is also limited, since there are only two types of standard Borel spaces: (i) spaces $(A, \mathcal{P}(A))$, where $A$ is a *countable* (possibly finite) set with its full powerset $\mathcal{P}(A)$ as the $\sigma$-algebra; and (ii) spaces $(A, \mathcal{B}_A)$ that are isomorphic to the real numbers with the Borel $\sigma$-algebra $(\mathbb{R}, \mathcal{B})$. As a consequence of this classification, every standard Borel space has a measurable embedding into the interval $[0, 1]$ with the Borel $\sigma$-algebra $\mathcal{B}_{[0,1]}$.

Standard Borel probability spaces will act as our sample spaces. As such, they will provide the objects of the category of sample spaces over which we shall consider atomic sheaves.

*Definition 8.2 (Standard Borel probability space).* A *standard Borel probability space* (SBPS) is a triple $(\Omega, \mathcal{B}_\Omega, P_\Omega)$ where $(\Omega, \mathcal{B}_\Omega)$ is an SBS and $P_\Omega \colon \mathcal{B}_\Omega \to [0, 1]$ is a probability measure. A *morphism* of standard Borel probability spaces from $(\Omega, \mathcal{B}_\Omega, P_\Omega)$ to $(\Omega', \mathcal{B}_{\Omega'}, P_{\Omega'})$ is an SBS morphism $q$ from $(\Omega, \mathcal{B}_\Omega)$ to $(\Omega', \mathcal{B}_{\Omega'})$ that *preserves measure*; i.e., $q_*(P_\Omega) = P_{\Omega'}$, where $q_*(P)$ is the *pushforward measure* $S \mapsto P_\Omega(q^{-1}(S)) : \mathcal{B}_{\Omega'} \to [0, 1]$ .

As with standard Borel spaces, standard Borel probability spaces include the most common probability spaces that one naturally encounters in mathematics. Any standard Borel probability space $(\Omega, \mathcal{B}_\Omega, P_\Omega)$ can be decomposed uniquely into its *discrete* and *continuous* parts, moreover the continuous part has a very constrained form. In detail, there exist unique Borel measures $\delta, \mu : \mathcal{B}_\Omega \to [0, 1]$ such that $P_\Omega = \delta + \mu$, the measure $\delta$ is *discrete* (i.e., $\delta(B) = \sum_{x \in B} \delta(\{x\})$

for every $B \in \mathcal{B}_\Omega$), and either $\mu = 0$ or $(\Omega, \mathcal{B}_\Omega, \mu)$ is isomorphic, via measure-preserving functions, to the interval $([0,c], \mathcal{B}_{[0,c]}, \lambda)$, where $c := P_\Omega(\Omega)$, with the Borel $\sigma$-algebra $\mathcal{B}_{[0,c]}$ and the (Borel restriction of) Lebesgue measure $\lambda : \mathcal{B}_{[0,c]} \to [0,c]$.

In probability theory, a random variable is a measurable function from a probability space, called the sample space, to a measurable space, the value space. In this paper, we restrict ourselves to the case in which these spaces are both standard Borel. This is broad enough to incorporate both the discrete and continuous random variables arising most commonly in mathematics.

*Definition 8.3 (Random variable).* If $\Omega$ is an SBPS and $A$ is an SBS (for notational convenience we here and henceforth abbreviate $(A, \mathcal{B}_A)$ as $A$ and $(\Omega, \mathcal{B}_\Omega, P_\Omega)$ as $\Omega$), a *random variable* $X : \Omega \to A$ is a measurable function from $(\Omega, \mathcal{B}_\Omega)$ to $(A, \mathcal{B}_A)$. The SBPS $\Omega$ is called the *sample space* of $X$, and the SBS $A$ is called the *value space*.

We next define the three main relations between random variables we shall be interested in: *almost-sure equality*, *equidistribution* and *conditional independence*.

In general, we say that a property of elements $\omega \in \Omega$ holds *for $P_\Omega$-almost-all $\omega$* if there exists $S \in \mathcal{B}_\Omega$ with $P_\Omega(S) = 1$ such that the property holds for every $\omega \in S$.

*Definition 8.4 (Almost-sure equality).* Two random variables $X, Y : \Omega \to A$ are *almost surely equal* (notation $X =_{\text{a.s.}} Y$) if $X(\omega) = Y(\omega)$ holds for $P_\Omega$-almost-all $\omega$. (Since $A$ is a standard Borel space, the set $\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}$ is measurable, and the above condition is equivalent to asking that $P_\Omega(\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}) = 1$.)

The *distribution* (or *law*) of a random variable $X : \Omega \to A$ is the probability measure $P_X : \mathcal{B}_A \to [0,1]$ defined as the pushforward $P_X := X_*(P_\Omega)$.

*Definition 8.5 (Equidistribution).* Two random variables $X, Y : \Omega \to A$ are *equidistributed* (notation $X \overset{d}{=} Y$) if $P_X = P_Y$.

An important consequence of only considering random variables between standard Borel spaces is that random variables have *disintegrations*. We state this property as Fact 8.6 below. A proof of can be found in [10]. We mention also that an equivalent statement to Fact 8.6 appears as Theorem 6 of [6].

*Fact 8.6.* Every random variable $X \colon \Omega \to A$ has a *disintegration*; that is, a Markov kernel $D_X : A \times \mathcal{B}_\Omega \to [0,1]$

$$(x, S) \;\mapsto\; P_{X^{-1}(x)}(S) \;:\; A \times \mathcal{B}_\Omega \to [0,1]$$

satisfying the two properties below.

**(D1)** $P_{X^{-1}(x)}(X^{-1}(x)) = 1$ for $P_X$-almost all $x \in A$, and

**(D2)** for every $S \in \mathcal{B}_\Omega$,

$$P_\Omega(S) \;=\; \int P_{X^{-1}(x)}(S)\; \mathrm{d}P_X(x) \;\; .$$

By the Markov kernel property, the function $S \mapsto P_{X^{-1}(x)}(S)$ is a probability measure $P_{X^{-1}(x)} : \mathcal{B}_\Omega \to [0,1]$, for every $x \in A$. By (D1), $P_{X^{-1}(x)}$ can be thought of as a probability measure on the fibre set $X^{-1}(x) \in \mathcal{B}_\Omega$, which, by (D2), represents the conditional probability distribution on $\omega \in \Omega$ under the condition $X(\omega) = x$. Properties (D1) and (D2) together characterise the mapping $x \mapsto P_{X^{-1}(x)}$ up to $P_\Omega$-almost-sure equality.

Exploiting disintegrations, we give a definition of conditional independence that is a transparent generalisation of the elementary probabilistic definition of unconditional independence.

*Definition 8.7 (Conditional independence).* For random variables $X : \Omega \to A,\ Y : \Omega \to B$ and $Z : \Omega \to C$, we say that $X$ and $Y$ are *conditionally independent given* $Z$ (notation $X \perp\!\!\!\perp Y \mid Z$) if, for every $S \in \mathcal{B}_A$ and $T \in \mathcal{B}_B$, and for $P_Z$-almost all $z \in C$,

$$P_{Z^{-1}(z)}(X^{-1}(S) \cap Y^{-1}(T)) \ = \ P_{Z^{-1}(z)}(X^{-1}(S)) \cdot P_{Z^{-1}(z)}(Y^{-1}(T)) \ .$$

Our goal in this section is to recover the three principal relations between random variables (almost-sure equality, equidistribution and conditional independence) as the relations of equality, atomic equivalence and atomic conditional independence in a suitable atomic sheaf topos. In order to be able to construct sheaves of random variables, the category over which sheaves will be taken is a category of sample spaces. In fact we consider two such categories.

*Definition 8.8 (The categories $\mathbb{SBP}$ and $\mathbb{SBP}_0$).* We write $\mathbb{SBP}$ for a small category of standard Borel probability spaces, that contains every such space up to isomorphism. We write $\mathbb{SBP}_0$ for the quotient category, with the same objects, in which morphisms are equivalence classes $[p]$ of maps modulo almost-sure equality $=_{\text{a.s.}}$.

It is an interesting fact that one can take the category of atomic sheaves over either category, $\mathbb{SBP}$ or $\mathbb{SBP}_0$, and in doing so one obtains equivalent categories of sheaves. Sheaves for the atomic topology on $\mathbb{SBP}$ were introduced in [37] as *probability sheaves*. In the present paper, it will be convenient to instead take atomic sheaves over $\mathbb{SBP}_0$. Since the two categories of sheaves are equivalent, we shall continue to use the name *probability sheaves*. The equivalence of the two categories will be shown in a separate paper.

An important advantage of working with $\mathbb{SBP}_0$ is the property below, which fails for $\mathbb{SBP}$.

Proposition 8.9. *Every morphism in $\mathbb{SBP}_0$ is an epimorphism.*

Proof. We first observe that every map $q : \Omega \to \Omega'$ in $\mathbb{SBP}$ is *almost surjective* in the sense that, for any $S \in \mathcal{B}_\Omega$ with $P_\Omega(S) = 1$, there exists $T \subseteq q(S)$ such that $T \in \mathcal{B}_{\Omega'}$ and $P_{\Omega'}(T) = 1$. This holds because the image $q(S)$ is an analytic subset of $\Omega'$ with outer measure 1. Since all analytic sets are measurable with respect to the completion of the Borel measure $P_{\Omega'}$, the image $q(S)$ also has inner measure 1, meaning that there exists $T \subseteq q(\Omega)$ with the required properties.

To prove that every morphism in $\mathbb{SBP}_0$ is epimorphic, suppose we have $[q] : \Omega \to \Omega'$ and $[r], [r'] : \Omega' \to \Omega''$ such that $[r] \circ [q] = [r'] \circ [q]$; i.e., $r \circ q =_{\text{a.s.}} r' \circ q$. Let $S \subseteq \Omega$ be Borel such that $P_\Omega(S) = 1$ and $(r \circ q) \restriction_S = (r' \circ q) \restriction_S$. By the almost surjectivity of $q$, let $T \subseteq q(S)$ be such that $T \in \mathcal{B}_{\Omega'}$ and $P_{\Omega'}(T) = 1$. Then $r \restriction_T = r' \restriction_T$; i.e., $r =_{\text{a.s.}} r'$. Equivalently $[r] = [r']$ as required. □

Proposition 8.10. *The category $\mathbb{SBP}_0$ has pairings.*

Proof. Given any span $\Omega_Y \xleftarrow{[p]} \Omega_X \xrightarrow{[q]} \Omega_Z$ in $\mathbb{SBP}_0$, its pairing is given by $(\Omega,\ [(p,q)],\ \pi_1, \pi_2)$, where $\Omega := (\Omega_Y \times \Omega_Z, \mathcal{B}_{\Omega_Y \times \Omega_Z}, P_{(p,q)})$, using the product standard Borel space and the probability distribution of the paired random variables $p$ and $q$. The properties of a pairing are easily verified, using Proposition 8.9 for uniqueness. □

*Definition 8.11 (Independent square in $\mathbb{SBP}_0$).* Define a commuting square in $\mathbb{SBP}_0$

$$\begin{array}{ccc} \Omega_X & \xrightarrow{[p]} & \Omega_Y \\ {\scriptstyle [q]}\downarrow & & \downarrow{\scriptstyle [r]} \\ \Omega_Z & \xrightarrow{[s]} & \Omega_W \end{array} \tag{37}$$

to be *independent* if $p \perp\!\!\!\perp q \mid r \circ p$, using conditional independence of random variables (Definition 8.7).

Proposition 8.12. *Definition 8.11 endows $\mathbb{SBP}_0$ with independent pullback structure satisfying the descent property.*

The proof of Proposition 8.12, which is intricate, can be found in Appendix C. In the present section, we content ourselves with exhibiting the construction needed to complete any cospan $\Omega_Y \xrightarrow{[r]} \Omega_W \xleftarrow{[s]} \Omega_Z$ to an independent pullback. Using the disintegrations for $r$ and $s$, we endow the standard Borel product $(\Omega_Y \times \Omega_Z, \mathcal{B}_{\Omega_Y \times \Omega_Z})$ with the probability measure $P$ defined as:

$$U \;\mapsto\; \int (P_{r^{-1}(\omega)} \otimes P_{s^{-1}(\omega)})(U)\; \mathrm{d}P_{\Omega_W}(\omega)\,, \tag{38}$$

where $P_{r^{-1}(\omega)} \otimes P_{s^{-1}(\omega)}$ is the product probability measure. Then

$$(\Omega_Y \times \Omega_Z, \mathcal{B}_{\Omega_Y \times \Omega_Z}, P)$$

together with the two projections, which are measure preserving, gives the required independent pullback. We write the resulting independent pullback square as

$$\begin{array}{ccc} \Omega_Y \otimes_{\Omega_W} \Omega_Z & \xrightarrow{[p_1]} & \Omega_Y \\ {\scriptstyle [p_2]}\downarrow & & \downarrow{\scriptstyle [r]} \\ \Omega_Z & \xrightarrow{[s]} & \Omega_W \end{array}$$

In combination, Propositions 8.9, 8.10 and 8.12 show that the category $\mathbb{SBP}_0$ has the requisite structure (Definition 7.8).

We next define the anticipated sheaves of random variables, first by defining them as presheaves, and then subsequently verifying the atomic sheaf property.

*Definition 8.13 (Presheaf of random variables $\underline{\mathrm{RV}}(A)$).* Let $A$ be a standard Borel space. Define a presheaf $\underline{\mathrm{RV}}(A) \in \mathrm{Psh}(\mathbb{SBP}_0)$ of $A$=valued random variables (modulo $=_{\text{a.s.}}$) by:

- $\underline{\mathrm{RV}}(A)(\Omega) :=$ equivalence classes of random variables $X : \Omega \longrightarrow A$ modulo $=_{\text{a.s.}}$.
- For $[X] \in \underline{\mathrm{RV}}(A)(\Omega)$ and $[q] \colon \Omega' \longrightarrow \Omega$, define $[X] \cdot [q] := [X \circ q]$.

We remark that a similar definition can be used to define a presheaf of $A$-valued random variables modulo $=_{\text{a.s.}}$ over the base category $\mathbb{SBP}$. In the case that $\mathbb{SBP}$ is used as the base category, one can also define an alternative presheaf of random variables, in which random variables are not quotiented modulo $=_{\text{a.s.}}$, an option which is not available when $\mathbb{SBP}_0$ is used as the base category. The $\mathbb{SBP}$ -presheaf of unquotiented $A$-valued random variables is not, however, an atomic sheaf. In contrast, irrespective of the choice of base category, $\mathbb{SBP}$ or $\mathbb{SBP}_0$, the presheaf of random variables modulo $=_{\text{a.s.}}$ does form a sheaf. We prove this in the case of our chosen base category, $\mathbb{SBP}_0$.

Proposition 8.14. *For any standard Borel space $A$, it holds that $\underline{\mathrm{RV}}(A)$ is an atomic sheaf.*

Proof. Suppose $[Y] \in \underline{\mathrm{RV}}(A)(\Omega')$ is $[q]$-invariant where $\Omega' \xrightarrow{q} \Omega$. is a map in $\mathbb{SBP}_0$. Consider the independent pullback square

$$\begin{array}{ccc} \Omega' \otimes_{\Omega} \Omega' & \xrightarrow{[p_1]} & \Omega' \\ {\scriptstyle [p_2]}\downarrow & & \downarrow{\scriptstyle [q]} \\ \Omega' & \xrightarrow{[q]} & \Omega \end{array}$$

By $[q]$-invariance, $[Y] \cdot [p_1] = [Y] \cdot [p_2]$, i.e., $Y \circ p_1 =_{\text{a.s.}} Y \circ p_2$. That is, the measure of

$$U \;:=\; \{(\omega_1', \omega_2') \in \Omega' \times \Omega' \mid Y(\omega_1) = Y(\omega_2)\}$$

in $\Omega' \otimes_\Omega \Omega'$ is 1. Equivalently, using (38),

$$\int (P_{q^{-1}(\omega)} \otimes P_{q^{-1}(\omega)})(U) \, \mathrm{d}P_\Omega(\omega) \;=\; 1.$$

So, for $P_\Omega$-almost all $\omega \in \Omega$, we have

$$(P_{q^{-1}(\omega)} \otimes P_{q^{-1}(\omega)})(U) \;=\; 1 \;.$$

For any such $\omega$, by the definition of product measure,

$$\int \int \mathbb{1}_U(\omega'_1, \omega'_2) \, \mathrm{d}P_{q^{-1}(\omega)}(\omega'_1) \, \mathrm{d}P_{q^{-1}(\omega)}(\omega'_2) \;=\; 1 \quad ,$$

where $\mathbb{1}_U$ is the indicator function for the set $U$. So for $P_{q^{-1}(\omega)}$-almost all $\omega'_1$ and $P_{q^{-1}(\omega)}$-almost all $\omega'_2$, we have $(\omega'_1, \omega'_2) \in U$, i.e., $Y(\omega'_1) = Y(\omega'_2)$. By arguing using the decomposability property of $P_{q^{-1}(\omega)}$ discussed beneath Definition 8.2, it follows there exists a Borel subset $C_\omega \subseteq \Omega'$ with $P_{q^{-1}(\omega)}(C_\omega) = 1$ such that $Y$ is constant on $C_\omega$. By the first property of disintegrations, $P_{q^{-1}(\omega)}(q^{-1}(\omega)) = 1$. Defining $D_\omega := C_\omega \cap q^{-1}(\omega)$, it holds that $P_{q^{-1}(\omega)}(D_\omega) = 1$, the function $q$ has constant value $\omega$ on $D_\omega$, and $Y$ is also constant on $D_\omega$. Let $d_\omega$ be the constant value of $Y$ on $D_\omega$. Note that we have obtained such $d_\omega$ and $D_\omega$, for $P_\Omega$-almost-all $\omega$.

Next we show that there exists a measurable function $X : \Omega \to A$ such that $X(\omega) = d_\omega$, for $P_\Omega$-almost all $\omega$. We first show this in the special case that $A \subseteq \mathbb{R}$ is a closed bounded interval, so all $A$-valued random variables are integrable with their integrals taking values in $A$. Using integrability, we define

$$X(\omega) \;:=\; \int Y(\omega') \, \mathrm{d}P_{q^{-1}(\omega)}(\omega') \;. \tag{39}$$

For $P_\Omega$-almost all $\omega$, we have

$$\int Y(\omega') \, \mathrm{d}P_{q^{-1}(\omega)}(\omega') \;=\; d_\omega \;, \tag{40}$$

because $Y(\omega') = d_\omega$, for $P_{q^{-1}(\omega)}$-almost all $\omega' \in D_\omega$. So we indeed have the required measurable function $X$ in the case of a closed bounded interval $A$. In the case of an arbitrary standard Borel space $A$, one takes some measurable embedding of $A$ into $[0, 1]$ (see the discussion after Definition 8.1), and then the definition of $X$ given above can be used to obtain a measurable function $\Omega \to [0, 1]$ that lands with probability 1 in the image of the embedding of $A$ in $[0, 1]$, meaning that it restricts (modulo redefining it on a null set) to the required map $X : \Omega \to A$.

We next verify that $X \circ q =_{\text{a.s.}} Y : \Omega' \to A$. Consider the Borel set $E := \{\omega' \in \Omega' \mid X(q(\omega')) = Y(\omega')\}$. We claim that, for $P_\Omega$-almost-every $\omega$, it holds that $D_\omega \subseteq E$. Indeed, for $P_\Omega$-almost-all $\omega$, we have that $\omega' \in D_\omega$ implies both $q(\omega') = \omega$ and $Y(\omega') = d_\omega$, hence $X(q(\omega')) = Y(\omega')$ follows, i.e., $\omega' \in E$. Because $D_\omega \subseteq E$, we have

$$P_{q^{-1}(\omega)}(E) \;=\; P_{q^{-1}(\omega)}(D_\omega) \;=1 \;.$$

By the definition of disintegrations,

$$P_{\Omega'}(E) \;=\; \int P_{q^{-1}(\omega)}(E) \, \mathrm{d}P_\Omega(\omega) \;=\; \int 1 \, \mathrm{d}P_\Omega(\omega) \;=\; 1 \;.$$

So indeed $X \circ q =_{\text{a.s.}} Y : \Omega' \to A$. That is, $[X] \cdot [q] = [Y]$. So $[X]$ is a $[q]$-descendent of $[Y]$.

That $[X]$ is the unique $[q]$-descendent of $[Y]$ holds because $q$ is almost surjective, as in the proof of Proposition 8.9. ◻

Corollary 8.15. *For any SBPS $\Omega$ the representable presheaf $\mathbf{y}\Omega$ is an atomic sheaf,.*

Proof. For any SBPS $\Omega'$, we have that $(\mathbf{y}\Omega)(\Omega') \subseteq \underline{\mathrm{RV}}(\Omega)(\Omega')$; indeed it is the subset of measure-preserving functions. It is then easily verified using Proposition 4.3 that $\mathbf{y}\Omega$ is a subsheaf of $\underline{\mathrm{RV}}(\Omega)$. In particular, $\mathbf{y}\Omega$ is a sheaf. ◻

We end this section by showing as promised that the three atomic forms of atomic formula of our general atomic sheaf logic are, in the case that sorts are interpreted as sheaves of random variables, correctly interpreted as the expected probabilistic relations between random variables. Firstly, that equality in the logic corresponds to almost sure equality of random variables is immediate from the definition of the sheaf $\underline{\mathsf{RV}}(A)$, in which random variables are explicitly identified modulo $=_{\text{a.s.}}$. Secondly, Proposition 8.16 below shows that atomic equivalence is interpreded as the equidistribution relation $\stackrel{d}{=}$.

Proposition 8.16. *For any SBS $A$, the atomic equivalence subsheaf $\sim_{\underline{\mathsf{RV}}(A)} \subseteq \underline{\mathsf{RV}}(A) \times \underline{\mathsf{RV}}(A)$ from Theorem 5.1 satisfies:*

$$\sim_{\underline{\mathsf{RV}}(A)} (\Omega) = \{([X],[X']) \in (\underline{\mathsf{RV}}(A) \times \underline{\mathsf{RV}}(A))(\Omega) \mid X \stackrel{d}{=} X'\} \ .$$

Proof. Consider any $[X],[X'] \in \underline{\mathsf{RV}}(A)(\Omega)$.

Suppose we have $[u],[u'] \colon \Omega' \to \Omega$ with $[X] \cdot [u] = [X'] \cdot [u']$, i.e., $X \circ u =_{\text{a.s.}} X' \circ u'$. Then $(X \circ u)_*(P_{\Omega'}) = (X' \circ u')_*(P_{\Omega'})$. Whence

$$X_*(P_\Omega) \ = \ X_*(u_*(P_{\Omega'})) \ = \ X'_*(u'_*(P_{\Omega'})) \ = \ X'_*(P_\Omega) \ ,$$

which shows $X \stackrel{d}{=} X'$.

Conversely, suppose $X \stackrel{d}{=} X'$; i.e., $X_*(P_\Omega) = X'_*(P_\Omega)$. We write $\Omega_A$ for the SBP space given by $A$ together with the probability measure $P_S := X_*(P_\Omega)$. With this probability measure, the functions $X \colon \Omega \to \Omega_A$ and $X' \colon \Omega \to \Omega_A$ are morphisms in $\mathbb{SBP}$. By coconfluence, there exist $p,q \colon \Omega' \to \Omega$ such that $X \circ p =_{\text{a.s.}} X' \circ q$, which implies $[X] \cdot [p] = [X'] \cdot [q]$. So indeed $([X],[X']) \in \sim_{\underline{\mathsf{RV}}(A)}(\Omega)$. □

The remaining form of atomic formula in our logic is conditional independence. Proposition 8.18 below shows that atomic conditional independence is indeed interpreted as the probabilistic relation of conditional independence (Definition 8.7). Before this, in order to be able to make sense of the relation of atomic conditional independence, we need to verify that the sheaves $\underline{\mathsf{RV}}(A)$ have supports (Definition 7.1).

Proposition 8.17. *For any standard Borel space $A$, it holds that $\underline{\mathsf{RV}}(A)$ has supports.*

Proof. Consider any $[X] \in \underline{\mathsf{RV}}(A)(\Omega)$. Define a standard Borel probability space by

$$\Omega_X \ := \ A \text{ with probability measure } P_{\Omega_X} := X_*(P_\Omega)\,.$$

It is easily checked that $(\Omega_X,\ [X],\ [x \mapsto x])$ is a support for $[X]$, using the almost surjectivity of $[X] : \Omega \to \Omega_X$, as in the proof of Proposition 8.9, for uniqueness. □

Proposition 8.18. *For any SBSs $A, B, C$, the atomic conditional independence subsheaf*

$$\perp\!\!\!\perp_{\underline{\mathsf{RV}}(A),\underline{\mathsf{RV}}(B)\mid\underline{\mathsf{RV}}(C)} \ \subseteq \ \underline{\mathsf{RV}}(A) \times \underline{\mathsf{RV}}(B) \times \underline{\mathsf{RV}}(C)$$

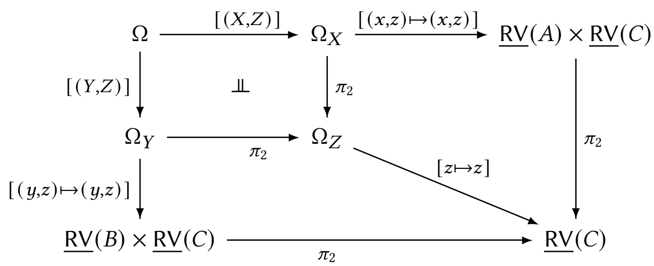
*from Theorem 7.10 satisfies:*

$$\perp\!\!\!\perp_{\underline{\mathsf{RV}}(A),\underline{\mathsf{RV}}(B)\mid\underline{\mathsf{RV}}(C)}(\Omega) \ = \ \{([X],[Y],[Z]) \in (\underline{\mathsf{RV}}(A) \times \underline{\mathsf{RV}}(B) \times \underline{\mathsf{RV}}(C))(\Omega) \mid X \perp\!\!\!\perp Y \mid Z\} \ .$$

Proof. Suppose $[X] \in \underline{\mathrm{RV}}(A)(\Omega)$, $[Y] \in \underline{\mathrm{RV}}(B)(\Omega)$ and $[Z] \in \underline{\mathrm{RV}}(C)(\Omega)$ are such that $X \perp\!\!\!\perp Y \mid Z$ according to Definition 8.7. Define

$$\begin{aligned} \Omega_X &:= A \times C \text{ with probability measure } P_{\Omega_X} := (X, Z)_*(P_\Omega) \\ \Omega_Y &:= B \times C \text{ with probability measure } P_{\Omega_Y} := (Y, Z)_*(P_\Omega) \\ \Omega_Z &:= C \text{ with probability measure } P_{\Omega_Z} := Z_*(P_\Omega)\,. \end{aligned}$$

Then the hybrid diagram below, shows that the triple $([X], [Y], [Z])$ belongs to the atomic conditional independence relation $\perp\!\!\!\perp_{\underline{\mathrm{RV}}(A), \underline{\mathrm{RV}}(B) \mid \underline{\mathrm{RV}}(C)}(\Omega)$.




In this diagram, $(\Omega_Z, [Z], [z \mapsto z])$ is support for $[Z]$, by the definition of $\Omega_Z$, and the top-left square is independent, because $(X, Z) \perp\!\!\!\perp (Y, Z) \mid Z$ holds, which follows from $X \perp\!\!\!\perp Y \mid Z$.

Conversely, suppose $([X], [Y], [Z]) \in \perp\!\!\!\perp_{\underline{\mathrm{RV}}(A), \underline{\mathrm{RV}}(B) \mid \underline{\mathrm{RV}}(C)}(\Omega)$. Defining $\Omega_X, \Omega_Y$ and $\Omega_Z$ as above, we have that $(\Omega_X, [(X, Z)], [(x, z) \mapsto (x, z)])$ is support for $[(X, Z)]$ and $(\Omega_Y, [(Y, Z)], [(y, z) \mapsto (y, z)])$ is support for $[(Y, Z)]$ $(\Omega_Z, [Z], [z \mapsto z])$ is support for $[Z]$. So, by Lemma 7.9, these supports fit into the hybrid diagram above. Since the top-left square is independent, we have $(X, Z) \perp\!\!\!\perp (Y, Z) \mid Z$. From this, $X \perp\!\!\!\perp Y \mid Z$ follows, as required. □

## 9 The Schanuel topos

We give a very condensed outline, without proofs, of one more example in which we have an atomic sheaf logic of equivalence and conditional independence: the Schanuel topos, which is equivalent to the category of nominal sets of Gabbay and Pitts [14, 35].

Let $\mathbb{I}$ be (a small version of) the category whose objects are finite sets and whose morphisms are injective functions. We consider the topos of atomic sheaves over the category $\mathbb{I}^{\mathrm{op}}$. Since all maps in $\mathbb{I}$ are obviously monomorphic, all maps in $\mathbb{I}^{\mathrm{op}}$ are epimorphic.

Proposition 9.1. *The category $\mathbb{I}^{\mathrm{op}}$ carries independent pullback structure satisfying the descent property. and it has pairings.*

Description of structure. Define a commuting square in $\mathbb{I}^{\mathrm{op}}$ to be *independent* if the associated square (with opposite orientation) of functions in $\mathbb{I}$ is a pullback in $\mathbb{I}$ (or equivalently in **Set**). A commuting square in $\mathbb{I}^{\mathrm{op}}$ is then an independent pullback if and only if the associated square of functions in $\mathbb{I}$ is a pushout in **Set** (but not necessarily in $\mathbb{I}$). Every cospan in $\mathbb{I}^{\mathrm{op}}$ completes to an independent pullback by taking the pushout in **Set** of the associated span of functions in $\mathbb{I}$. □

Proposition 9.2. *The category $\mathbb{I}^{\mathrm{op}}$ has pairings.*

Description of structure. A span in $\mathbb{I}^{op}$ gives rise to a cospan of functions in $\mathbb{I}$. The pairing in $\mathbb{I}^{op}$ is given by the pushout in **Set** of the pullback in $\mathbb{I}$ (or **Set**) of this cospan of functions. □

A presheaf $P \in \mathrm{Psh}(\mathbb{I}^{op})$ is just a covariant functor $P\colon \mathbb{I} \to \mathbf{Set}$. The description of independent squares above, means that Theorem 6.6, in the case of $\mathbb{C} = \mathbb{I}^{op}$, specialises to the well-known characterisation that a presheaf $P \in \mathrm{Psh}(\mathbb{I}^{op})$ is an atomic sheaf if and only if the covariant functor $P\colon \mathbb{I} \to \mathbf{Set}$ preserves pullbacks (see, e.g., [23, A 2.1.11(h)]). This property enables the result below to be established by constructing supports in $\mathbb{I}^{op}$ as a multiple pullbacks in $\mathbb{I}$ over all representable factorisations, of which there are only finitely many.

Proposition 9.3. *Every atomic sheaf in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{op})$ *has supports.*

For a sheaf $\underline{A}$ in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{op})$, the support of an element $x \in \underline{A}(X)$ corresponds to a smallest subset $\mathrm{supp}(x) \subseteq X$ for which there exists $y \in \underline{A}(\mathrm{supp}(x))$ such that $x = y \cdot i$, where $i : X \to \mathrm{supp}(x)$ in $\mathbb{I}^{op}$ is given by the inclusion function $\mathrm{supp}(x) \to X$. Proposition 9.3 is well known. For example, it plays a key role in Fiore's presentation of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{op})$ as a Kleisli category [13, 30]. An analogous property is also prominent in presentations of the equivalent category of nominal sets [14, 35].

Proposition 9.4. *For any* $\underline{A}$ *in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{op})$*, the atomic equivalence subsheaf* $\sim_{\underline{A}} \subseteq \underline{A}\times\underline{A}$ *from Theorem 5.1 satisfies:*

$$\sim_{\underline{A}} (X) = \{(x, y) \in (\underline{A} \times \underline{A})(X) \mid \exists X \xrightarrow{i} X.\quad y = x \cdot i\} \ .$$

Proposition 9.5. *For any* $\underline{A}, \underline{B}, \underline{C}$ *in* $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{op})$*, the atomic conditional independence subsheaf*

$$\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}} \subseteq \underline{A}\times\underline{B}\times\underline{C}$$

*from Theorem 7.10 satisfies:*

$$\perp\!\!\!\perp_{\underline{A},\underline{B}|\underline{C}}(X) \;=\; \{(x, y, z) \in (\underline{A} \times \underline{B} \times \underline{C})(X) \mid \mathrm{supp}(x) \cap \mathrm{supp}(y) \subseteq \mathrm{supp}(z)\ \} \ .$$

## 10 Discussion and related work

### 10.1 Relationship with (multi)team semantics

Our main running example throughout the paper has been the category of atomic sheaves over the category $\mathbb{Sur}$, in which the interpretations of atomic equivalence and conditional independence, when applied to the sheaves $\underline{\mathrm{NV}}(A)$ of nondeterministic variables, coincide with the multiteam interpretations of those relations from the *(in)dependence logics* of [11, 17, 42]. For our logic, we use the canonical internal logic of an atomic sheaf topos, whose semantics is provided by the forcing relation of Figure 2, and whose underlying logic is ordinary classical logic.

In our route to atomic sheaf logic in Sections 2–4, the use of multiteams seems essential. Indeed, it is the presentation of multiteams as finite-fibred functions in Section 2 that forms the basis for the connection with the category $\mathbb{Sur}$, whence with atomic sheaves. This contrasts with the majority of work on (in)dependence logic, from [17, 42] onwards, which is largely based on teams rather than on multiteams. It is accordingly worth observing, that it is possible to reformulate the atomic sheaf logic of Figure 2 directly in terms of teams. To see this, note that any finite team trivially gives rise to a canonical finite multiteam, in which every assignment has multiplicity 1. Conversely, the support of any finite multiteam is a team. Under the correspondence between $\mathcal{V}$-multiteams, and $\mathcal{V}$-assignments of nondeterministic variables, discussed in Section 2, we can reformulate these two statements in the following way. Every finite $\mathcal{V}$-team gives rise to $\underline{\rho}\colon \mathcal{V} \to (\Omega \to A)$ enjoying the *team property*: for all $\omega, \omega' \in \Omega$, if $\underline{\rho}(\mathsf{x})(\omega) = \underline{\rho}_S(\mathsf{x})(\omega')$, for all $\mathsf{x} \in \mathcal{V}$, then

$\omega = \omega'$. Moreover, for every $\mathcal{V}$-multiteam $\underline{\rho'} : \mathcal{V} \to (\Omega' \to A)$ there exists a unique up to isomorphism $q : \Omega' \to \Omega$ and $\underline{\rho} : \mathcal{V} \to (\Omega \to A)$ such that $\underline{\rho}$ satisfies the team property and $\underline{\rho'} = \underline{\rho} \cdot q$. It thus follows from the sheaf property of forcing (Proposition 4.8) that the behaviour of the relation $\Omega \Vdash_{\underline{\rho}} \Phi$ in $\mathrm{Sh_{at}}(\mathbb{Sur})$, for any formula $\Phi$, is determined entirely by its behaviour on teams $\underline{\rho}$. Moreover, it is easy to unwind the clauses in Figure 2 and to reformulate them directly in terms of ordinary teams qua sets of assignments. Thus atomic sheaf logic over $\mathbb{Sur}$ could equivalently be presented in terms of teams rather than multiteams.

If one carries out such a reformulation in the case of conjunction and of the existential quantifier, one obtains the standard team interpretation of the former [42], and the *lax* interpretation of the latter, which is often the preferred team interpretation [17, 18]. The clauses for the other connectives and for the universal quantifier are different however. Whereas the clauses in Figure 2 validate the laws of classical logic, it is well known that the standard team semantics of the other connectives and the universal quantifier leads to some logically exotic behaviour. For example, disjunction is not an idempotent operation. Abramsky and Väänänen [1] provide an illuminating explanation for such behaviour, by showing that the dependence logic connectives and quantifiers can be naturally understood as fitting into the framework of Pym and O'Hearn's *logic of bunched implications (BI)* [32, 36]. We now review this perspective and then discuss how it might be adapted to atomic sheaf logic.

The approach of [1] is based on Lawvere's notion of *hyperdoctrine* [26, 34]. Recall that the contravariant poswerset functor $\mathsf{P}$ on sets, can be viewed as a functor $\mathsf{P} \colon \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$, where $\mathbf{Pos}$ is the category of partially ordered sets and monotone functions. Specifically, $\mathsf{P}$ maps any set $X$ to its set of subsets partially ordered set ordered by subset inclusion. The functor $\mathsf{P} \colon \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$ is then a hyperdoctrine. Propositional logic for propositions over a set $X$ is modelled by the boolean algebra structure on $\mathsf{P}(X)$. For any function $f \colon X \to Y$, the *reindexing function* $\mathsf{P}(f) := f^{-1} \colon \mathsf{P}(Y) \to \mathsf{P}(X)$ preserves the boolean algebra structure. The quantifiers $\exists : \mathsf{P}(X \times Y) \to \mathsf{P}(X)$ and $\forall : \mathsf{P}(X \times Y) \to \mathsf{P}(X)$, quantifying over a set $Y$, are modelled as left and right adjoints respectively to the monotone function (considered qua functor) $\pi_1^{-1} : \mathsf{P}(X) \to \mathsf{P}(X \times Y)$, where $\pi_1 \colon X \times Y \to X$ is the projection map.

The main construction in [1], adapts the above hyperdoctrine for classical logic to team semantics, by composing $\mathsf{P} \colon \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$ with the functor $\mathcal{L} \colon \mathbf{Pos} \to \mathbf{Pos}$ given by the operation $\mathcal{L}$ that maps any partial order $B$ to its lattice $\mathcal{L}(B)$ of down-closed sets. The composite functor $\mathcal{L}\mathsf{P} \colon \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$ then has the following properties. For every set $X$, the fibre poset $\mathcal{L}\mathsf{P}(X)$ is, in a canonical way, a *BI algebra*, that is an algebraic model of the *logic of bunched implications* BI [32, 36]. In the case $X = A^{\mathcal{V}}$, the elements of $\mathcal{L}\mathsf{P}(A^{\mathcal{V}})$ are precisely down-closed (in the subset ordering) sets of $A$-valued teams with variable set $\mathcal{V}$. Each connective of BI is modelled algebraically as a function of appropriate arity on $\mathcal{L}\mathsf{P}(A^{\mathcal{V}})$. For example, the *multiplicative conjunction* $\otimes$, is modelled as a certain canonically generated function $\otimes : \mathcal{L}\mathsf{P}(A^{\mathcal{V}}) \times \mathcal{L}\mathsf{P}(A^{\mathcal{V}}) \to \mathcal{L}\mathsf{P}(A^{\mathcal{V}})$. Writing $\Phi$ and $\Psi$ for elements of $\mathcal{L}\mathsf{P}(A^{\mathcal{V}})$ (which can be thought of as an abstract set of propositions), and writing $S \Vdash \Phi$ to mean $S \in \Phi$, the function $\otimes$ can be characterised by

$$S \Vdash \Phi \otimes \Psi \iff \exists T, U,\ S = T \cup U \text{ and } T \Vdash \Phi \text{ and } U \Vdash \Psi\,.$$

This is exactly the semantic clause for the *disjunction* connective of team semantics. The exotic behaviour of the disjunction of dependence logic is thus nicely explained as a manifestation of the expected behaviour of the multiplicative conjunction of BI, whose multiplicative connectives have a natural resource-sensitive interpretation. A further consequence of the hyperdoctrine construction in [1] is that the embedding of dependence logic in BI enriches the former with additional logical connectives, such as both additive (intuitionistic) and multiplicative implications. Lastly, the hyperdoctrine formulation of dependence logic provides an elegant explanation for the team semantics interpretation of the quantifiers $\exists$ and $\forall$: $\mathcal{L}\mathsf{P}(A^{\mathcal{V}\uplus\{x\}}) \to \mathcal{L}\mathsf{P}(A^{\mathcal{V}})$, which are characterised in the desired way [26, 34] as respectively left and right adjoints to $\mathcal{L}\mathsf{P}(\rho \mapsto \rho|_{\mathcal{V}}) : \mathcal{L}\mathsf{P}(A^{\mathcal{V}}) \to \mathcal{L}\mathsf{P}(A^{\mathcal{V}\uplus\{x\}})$.

The above hyperdoctrine construction from [1] works for the original dependence logic [42], but not for independence logic [17], because teams satisfying independence atoms are not down-closed in the subset order. This means that the $\mathcal{L}$ functor cannot be used to interpret formulas involving independence. An alternative is to combine the contravariant powerset functor $\mathsf{P}$ with the covariant powerset functor $\mathsf{P}_!$ (with direct image as its functorial action). It turns out that if one considers the composition in the order $\mathsf{P}\mathsf{P}_! : \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$, then the left and right adjoints to the monotone function $\mathsf{P}\mathsf{P}_!(\rho \mapsto \rho|_{\mathcal{V}}) : \mathsf{P}\mathsf{P}_!(A^{\mathcal{V}}) \to \mathsf{P}\mathsf{P}_!(A^{\mathcal{V}\uplus\{x\}})$ correspond respectively to the existential and universal quantifier with (the team version of) the forcing clauses from Figure 2. Further, the boolean algebra structure on $\mathsf{P}\mathsf{P}_!(A^{\mathcal{V}})$ corresponds to (the team version of) the forcing clauses for the propositional connectives in Figure 2, and this structure is preserved by all *reindexing maps* $\mathsf{P}\mathsf{P}_!(f)$. The hyperdoctrine $\mathsf{P}\mathsf{P}_! : \mathbf{Set}^{\mathrm{op}} \to \mathbf{Pos}$ thus recovers the team version of atomic sheaf logic as in Figure 2. It would be interesting to investigate this construction in more detail, for example to explore how independence and equivalence formulas interact with the hyperdoctrine formulation, and also the extent to which the logic BI logic is relevant in this picture. Both points are potentially subtle. The standard hyperdoctrine desideratum that logical structure should be preserved by reindexing maps provides a constraint on which atomic primitives are admissible. Moreover, the relevance of BI logic is less *a priori* apparent than in [1], because the switch in the order of composition ($\mathsf{P}\mathsf{P}_!$ has the covariant functor as the inner functor, whereas $\mathcal{L}\mathsf{P}$ has its covariant functor as the outer functor) means that the outermost functor is no longer given by a canonical BI-algebra construction.

A different source of exotic behaviour in (in)dependence logics concerns interaction between the universal quantifier and (in)dependence atoms. One particularly striking example is provided by the sentence below.

$$\forall \mathsf{x}^{\mathsf{A}}, \forall \mathsf{y}^{\mathsf{B}}.\ \ (\mathsf{x}^{\mathsf{A}} \perp \mathsf{y}^{\mathsf{B}}) \tag{41}$$

According to the usual team semantics of the universal quantifier, the above sentence is valid. Nevertheless, one can easily exhibit example teams $S$ for which it is not the case that $S \Vdash \mathsf{x}^{\mathsf{A}} \perp \mathsf{y}^{\mathsf{B}}$, and rightly so, because there would be little point in independence logic if independence were a universally valid relation. We view the validity of (41) (and other examples like it) as showing that if one is to use (in)dependence logic as a basis for reasoning about (in)dependence properties then the associated rules of inference will have to be unusual.

Nevertheless, independence logics and their team semantics have been successfully applied in the direction of reasoning about dependence and conditional independence. For example, Hannula and Kontinen axiomatise the valid implications involving *inclusion* and *embedded multivalued dependencies* in database theory in terms of inclusion and conditional independence formulas with their team semantics [18]. An interesting observation about this work is that it takes place in the fragment of independence logic comprising conjunction and (lax) existential quantification as the only logical operators. Since these are exactly the logical operators for which the semantic interpretations in independence logic and atomic sheaf logic coincide, the same development can be imported verbatim into atomic sheaf logic in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{S}\mathrm{ur})$ extended with the inclusion relation (which indeed defines a subsheaf of $\underline{\mathrm{NV}}(A) \times \underline{\mathrm{NV}}(A)$). One advantage of such a reformulation is that the axiomatised rules of inference in [18] can be expressed as individual formulas, using the general implication connective of atomic sheaf logic, rather than left as entailments. For example, the rule of *inclusion introduction*, which concerns the inclusion relation, has an obvious (derivable) analogue for the equivalence (equiextension) relation, namely: if one has already derived an equivalence formula $\vec{\mathsf{x}} \sim \vec{\mathsf{x}}'$ then one can infer the formula $\exists \mathsf{y}'^{\mathsf{A}}.\ (\vec{\mathsf{x}}, \mathsf{y}^{\mathsf{A}} \sim \vec{\mathsf{x}}', \mathsf{y}'^{\mathsf{A}})$. In atomic sheaf semantics, this rule can be formulated as an implication. Indeed, it is none other than the *transfer principle* (16) from Figure 3, valid in any atomic sheaf topos. The same transfer principle can also be found in mainstream probability theory. The interpretation of (16) in the category $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$ of probability sheaves is very close to the *transfer theorem* of [24, Theorem 5.10], and arguably captures the essence of that theorem in logical form.

The interpretation of atomic sheaf logic in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$ also connects with a body of work on adapting team semantics to probability-based scenarios. For example, an $A$-valued *measure team* in [21] is a measurable map $\Omega \to (\mathcal{V} \to A)$, for some probability space $\Omega$ and set of variables $\mathcal{V}$. This can equivalently be presented as a map $\mathcal{V} \to (\Omega \to A)$, which is almost the same thing as a variable assignment in atomic sheaf logic over $\mathbb{SBP}_0$, i.e., a mapping from variables to elements of $\underline{\mathrm{RV}}(A)(\Omega)$. There are however two key differences: random variables in $\underline{\mathrm{RV}}(A)$ are identified up to almost sure equality, and objects in $\mathbb{SBP}_0$ are restricted to probability spaces $\Omega$ that are standard Borel. Although these differences may seem minor, they are crucial to the interpretation of atomic sheaf logic in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$. For example, it is because of the restriction to standard Borel spaces that the category $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$ is coconfluent. The failure of coconfluence for general probability spaces makes it difficult to extend the measure-team semantics of atomic formulas in [21] to include the logical connectives and atoms of independence logic. In the literature, such extensions have been given only for probabilistic teams based on discrete probability [12]. It is worth remarking that discrete probability fits in equally well with the approach of the present paper. One can consider atomic sheaves over the category of finite probability spaces, or alternatively over the category of countable probability spaces, both of which are full subcategories of $\mathbb{SBP}_0$. Such examples further substantiate our thesis that atomic sheaf categories provide a unifying framework configurable to diverse settings for conditional independence. It would be interesting to compare our approach with the semiring-based framework of [4], which provides a different unifying approach to varieties of team semantics, which encompasses both ordinary teams and discrete probabilistic teams.

### 10.2 Computer science applications

In this section we outline possible computer science applications for atomic sheaf logic. Rather than trying to be comprehensive, we instead focus on a few illustrative examples, beginning with reasoning about probabilistic programs.

An almost surely terminating imperative probabilistic program $C$ can be modelled as a probabilistic map between states, that is a function $[\![C]\!]_S : \mathrm{State} \to \mathcal{D}(\mathrm{State})$, where $\mathcal{D}(\mathrm{State})$ is the set of probability distributions over states. Alternatively, but equivalently, it can be viewed as a transformation $[\![C]\!]_T : \mathcal{D}(\mathrm{State}) \to \mathcal{D}(\mathrm{State})$ mapping a probability distribution on initial states to the induced probability distribution on final states [25]. There is also a third related possibility. One can view the program as a transformation $[\![C]\!]_R$ mapping an initial *random state* $\Sigma : \Omega \to \mathrm{State}$, for some sample space $\Omega$, to a final random state $\mathrm{T}$ [22]. However, because the program $C$ may make use of randomness not present in $\Omega$, the sample space for $\mathrm{T}$ has to be, in general, an *extension* of $\Omega$, meaning that $\mathrm{T} : \Omega' \to \mathrm{State}$ for some suitable sample space $\Omega'$ equipped with a probability preserving map $q : \Omega' \to \Omega$. While the idea of modelling programs as random-state transformers is very natural, some careful bookkeeping is required to deal with the change of sample space. Such bookkeeping can be avoided entirely if one uses the alternative approach of defining the random-state-transformer semantics in the atomic sheaf logic of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$. Under this approach $[\![C]\!]_R$ is formulated as a relation $[\![C]\!]_R \subseteq \underline{\mathrm{RV}}(\mathrm{State}) \times \underline{\mathrm{RV}}(\mathrm{State})$ satisfying: for any random initial state $\Sigma$ on which $C$ terminates, there exists a random final state $\mathrm{T}$ such that $\Sigma [\![C]\!]_R \mathrm{T}$ and, for any random state $\mathrm{T}'$, it holds that $\Sigma [\![C]\!]_R \mathrm{T}'$ implies $\Sigma, \mathrm{T} \sim \Sigma, \mathrm{T}'$. The key point here is that no sample spaces need to be specified, because, from the viewpoint of atomic sheaf logic, sample spaces are implicit, and the extension of sample spaces is likewise taken care of implicitly by the semantics of the existential quantifier. Not only is such an implicit-sample-space style of manipulating random variables intuitive, it also avoids the bookkeeping required when dealing with explicit sample space extensions. For example, in [22], a property called *relative tightness* is identified as useful property of probabilistic Hoare-triple-like specifications. Such a specification $\{\Phi\}C\{\Psi\}$ asserts that, if the precondition $\Phi$ holds for a random initial state $\Sigma$, and if $C$ terminates from $\Sigma$, then the postcondition $\Psi$ holds for the induced random final state T. The property of relative tightness asserts that the probabilistic behaviour of the random state T on the variables $\mathsf{FV}(\Psi)$ relevant to determining the truth of $\Psi$, depends only on the value of the initial state $\Sigma$ on $\mathsf{FV}(\Phi)$. This can be formulated in a simple way as the statement about conditional independence on the left below

$$\mathrm{T}_{\mathsf{FV}(\Psi)} \perp \Sigma \mid \Sigma_{\mathsf{FV}(\Phi)} \qquad\qquad \mathrm{T}_{\mathsf{FV}(\Psi)} \perp \Sigma \circ q \mid \Sigma_{\mathsf{FV}(\Phi)} \circ q\,,$$

where $\Sigma_{\mathsf{FV}(\Psi)}$ and $\mathrm{T}_{\mathsf{FV}(\Psi)}$ denote the initial and final random states restricted to the specified variable sets. For contrast, we include on the right above the statement of relative tightness that appears in [22], which shows the need for bookkeeping (in this case, composition with $q$) when the standard mathematical formulation of random variables with explicit sample spaces is used. For a more involved example of the efficiency afforded by the implicit-sample-space approach of atomic sheaf logic, we consider how the while statement on the left below is approximated by iterating the conditional statement on the right.

$$\mathsf{while}\ B\ \mathsf{do}\ C \qquad\qquad \mathsf{if}\ B\ \mathsf{then}\ C\ \mathsf{else\ skip}$$

Working within atomic sheaf logic, suppose the while statement terminates in random final state T from a random initial state $\Sigma$. Then defining $\Sigma_0 = \Sigma$ and letting $\Sigma_{n+1}$ be such that $\Sigma_n [\![C]\!]_R \Sigma_{n+1}$, we obtain a sequence $(\Sigma_n)_{n\geq 0}$ of random states that converges almost surely to the random state $T$. The resulting convergence property $\Sigma_n \to \mathrm{T}$ is used in ongoing work extending [22] to prove the correctness of a partial correctness rule for while loops in a probabilistic program logic. The formulation of the same convergence statement with explicit sample states is unwieldy as it involves a sequence $\Omega_0 \xleftarrow{q_0} \Omega_1 \xleftarrow{q_1} \Omega_2 \xleftarrow{q_2} \cdots$ of sample space extensions for the random states $(\Sigma_n)_n$, as well as a cone (in the category-theoretic sense) $(\Omega_n \xleftarrow{r_n} \Omega')_n$ for the sequence, where $\Omega'$ is the sample space for T. With this scaffolding in place, the convergence property can be stated as $\Sigma_n \circ r_n \to \mathrm{T}$.

We have outlined above how atomic sheaf logic might be applied to formulate a random-state-based operational semantics for imperative probabilistic programs. Another potential application is to the assertion logics of Hoare-like program logics for probabilistic programs, in particular to *probabilistic separation logic (PSL)*. PSL was first introduced in [5] as an approach to verifying probabilistic programs using a version of the *separating conjunction* of separation logic [31, 43] to reason about probabilistic independence. The modular style of reasoning is supported by a version of the *frame rule* of separation logic, which, in the case of probabilistic separation logic, allows certain statements about probabilistic independence to be inferred. The paper [5] presents several applications to the verification of cryptographic protocols. Subsequent work has extended the approach to reason about negative dependencies [3], adapted it to a probabilistic functional language [27] and incorporated conditional independence [2, 27]. In all the aforementioned works, the assertion logic has been given as an instance of the *logic of bunched implications (BI)* with a Kripke-style semantics defined over a partially ordered *resource monoid* [36]. This leads to an intuitionistic but not classical assertion logic. It seems likely that one can obtain a classical assertion logic, by replacing the Kripke-style semantics of BI in a partially ordered resource monoid with a category-based semantics utilising the forcing clauses of atomic sheaf logic.[3]

[3]'one version of such a classical assertion logic appears in [22]. However, the very simple setting of abstract *semantic asserions* with no explicit quantifiers in *op. cit.*, enables the category-theoretic genesis of the logic to be hidden. Its one remaining trace is the set of *footprint variables*, which corresponds to the notion of support in the present paper.

Another connection with the logic of bunched implications comes from a fact that we have not developed in the present paper: every category $\mathbb{C}$ with independent pullbacks and terminal object is symmetric monoidal, and its category $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$ of atomic sheaves carries, in addition to its cartesian closed structure, a second symmetric monoidal closed structure $\otimes_{\mathrm{Sh}}$ derived, using the methods of [9], from the modoidal structure of $\mathbb{C}$. Categories with two such closed structures are category-theoretic models of BI [32]. In the case of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{C})$, the monoidal structure is furthermore *affine*, hence it has projections $\underline{A} \leftarrow \underline{A} \otimes_{\mathrm{Sh}} \underline{B} \rightarrow \underline{B}$. In the case that $\underline{A}$ and $\underline{B}$ have supports, then the projections are jointly monic and the resulting monomorphism

$$\underline{A} \otimes_{\mathrm{Sh}} \underline{B} \rightarrowtail \underline{A} \times \underline{B}$$

is in fact isomorphic to $\perp\!\!\!\perp_{\underline{A},\underline{B}} \subseteq \underline{A} \times \underline{B}$ given by the unconditional version of (20) (i.e., in which $C$ is a terminal object). That is, the unconditional independence relation of the present paper is recovered as an instance of monoidal structure. This connection will be elaborated in a follow-up paper, where also the relationship with the monoidal category setting of [38] will be discussed. Indeed the notion of *local independence structure with local independent products* in *op. cit.* is equivalent to the independent pullback structure of Section 6, but with a much more involved axiomatisation in terms of fibred monoidal structure. The monoidal structure of $\mathbb{C}$ provides another connection between the work of the present paper and varieties of separation logic including probabilistic separation logic, as elaborated by Li *et. al.* [28]. In their work, the Day monoidal product on presheaves [9] is used to model the separation of state into independent segments, whose probabilistic independence can be superimposed using a probability monad. As in our work, the notion of sheaf with supports, which was introduced independently in [28], plays a crucial role.

The category **Nom** of nominal sets of Gabbay and Pitts [14, 35] has found applications to reasoning about *names* in computer science. The monograph [35] presents many examples of such applications, together with pointers to the literature. One prominent application area is reasoning about abstract syntax for languages involving operators that bind variables.

As mentioned in Section 9, the category **Nom** is equivalent to the Schanuel topos, and so the relations of equivalence and conditional independence defined in Section 9 can be transferred to **Nom**. In **Nom**, the atomic equivalence relation of Proposition 9.4 is the equivalence relation of being in the same orbit. The special case of Proposition 9.5 corresponding to the relation of unconditional independence $x \perp\!\!\!\perp y$ is the relation of *separatedness* ($\mathrm{supp}(x) \cap \mathrm{supp}(y) = \emptyset$), which is a central relation of interest in the literature on nominal sets. The full conditional independence relation $x \perp\!\!\!\perp y \mid z$ is then a *relative* notion of separatedness ($\mathrm{supp}(x) \cap \mathrm{supp}(y) \subseteq \mathrm{supp}(z)$), which first appeared in [38]. We believe that the atomic logic of equivalence and conditional independence developed in the present paper may, when transported to **Nom**, provide a convenient setting for reasoning about syntax with variable binding. Let us illustrate this using the untyped $\lambda$-calculus as an example.

There are several approaches to reasoning about syntax with variable binding. The first is to reason about *raw terms*, in which, for example, $\lambda x.\, x$ is distinguished from $\lambda y.\, y$ because the variable name differs. This leads to an awkward definition of substitution $M[x := N]$ that involves a non-canonical choice of bound-variable renaming, and does not provide a good foundation on which to base structured reasoning principles. Some abitrariness can be avoided by imposing canonicity on bound-variables names, for example using de Bruijn indices. However, syntactic manipulations then involve arithmetic operations on indices, which means that proofs of syntactic properties are entangled with arithmetic proofs that are an artefact of the choice of representation and have no intrinsic connection to the syntactic properties being proved. An alternative, favoured in many informal expositions of syntax, is to work with *equivalence classes* of terms modulo $\alpha$-equivalence instead of raw terms. This leads to a canonical definition of substitution, but has two drawbacks that are particularly significant if one wishes to formalise proofs. The first drawback is that all term manipulations need to be proved compatible with the equivalence relation. Such proofs are often omitted from informal expositions, but of course need to be given in a formal setting. The second drawback is that one loses the structural-induction principle on terms that is derived from the inductive definition of raw terms.. These two issues can be given a very elegant solution by defining syntax in the category of nominal sets. There is a functor called *name abstraction* that can be used to give a direct inductive definition of the nominal set of terms modulo $\alpha$-equivalence. This definition comes with an associated principle of structural induction for reasoning about terms modulo $\alpha$-equivalence, and a principle of structural recursion that allows one to define functions that are automatically well-defined on $\alpha$-equivalence classes. This approach is more fully described in the monograph [35], which also contains pointers to the wider literature. It seems fair to say, however, that this approach does not solve all the practical difficulties of reasoning about binding operators. For example, the structural induction and recursion principles can be cumbersome to work with, due to their side conditions involving concepts such as separatedness and freshness.

We propose here an alternative approach to reasoning about syntax with binding operators in the category of nominal sets. The idea is to reason directly about raw terms rather than about $\alpha$-equivalence classes of terms, but to use properties of the atomic-sheaf-logic equivalence and conditional independence relations to enable definitions and reasoning to be carried out in an elegant structural way. To illustrate the proposal, let us consider untyped $\lambda$-terms presented in the form $\Gamma \vdash M$, where $\Gamma$ is a finite sequence of distinct names that are treated as free variables in term $M$. The rules for generating such terms are:

$$\frac{}{\Gamma \vdash a}\ a \in \Gamma \qquad \frac{\Gamma \vdash M \quad \Gamma \vdash N}{\Gamma \vdash MN} \qquad \frac{\Gamma, a \vdash M}{\Gamma \vdash \lambda a.\, M}\ a \notin \Gamma\,.$$

Then the $\Gamma$-indexed relation $\{\equiv_\Gamma \subseteq \mathsf{Term}_\Gamma \times \mathsf{Term}_\Gamma\}_\Gamma$ of $\alpha$-equivalence, where

$$\mathsf{Term}_\Gamma := \{\Gamma \vdash M \mid \Gamma \vdash M \text{ is a term}\}\,,$$

can be defined as the smallest $\Gamma$-indexed congruence relation containing atomic equivalence $\{\sim_\Gamma \subseteq \mathsf{Term}_\Gamma \times \mathsf{Term}_\Gamma\}_\Gamma$ (i.e., orbit equality).[4] Substitution $\Gamma \vdash M[a := N]$ can be specified as a function defined on any pair of terms $\Gamma, a \vdash M$ and $\Gamma \vdash N$ for which the conditional independence (i.e., relative separation property)

$$\Gamma, a \vdash M \ \perp\ \Gamma \vdash N \ \mid\ \Gamma$$

holds, by simple structural recursion on the structure of the raw term $\Gamma, a \vdash M$. Of course, one would like substitution to be defined on *all* suitable terms, not just on sufficiently separated ones. This is achieved, by defining substitution as a ternary *relation* $\mathsf{Sub}_{\Gamma,a} \subseteq \mathsf{Term}_{\Gamma,a} \times \mathsf{Term}_\Gamma \times \mathsf{Term}_\Gamma$, by specifying that

$$\mathsf{Sub}_{\Gamma,a}(\Gamma, a \vdash M,\ \Gamma \vdash N,\ \Gamma \vdash L)$$

holds precisely when there exists $\Gamma \vdash N'$ such that $\Gamma \vdash N' \sim \Gamma \vdash N$ and $\Gamma, \{a\} \vdash M \perp \Gamma \vdash N' \mid \Gamma$ and $L = M[a := N']$. By the independent existence principle (30), this relation is total in the sense that, for any $M, N$ (for brevity we omit the contexts) there exists $L$ such that $\mathsf{Sub}(M, N, L)$. The relation is also single-valued up to equivalence: if $\mathsf{Sub}(M, N, L)$ and $\mathsf{Sub}(M, N, L')$ then it holds that $\Gamma \vdash L \sim \Gamma \vdash L'$. Preservation of $\alpha$-equivalence, then follows in the form: if $M \equiv_{\Gamma,a} M'$ and $N \equiv_\Gamma N'$ and $\mathsf{Sub}(M, N, L)$ and $\mathsf{Sub}(M', N', L')$ then $L \equiv_\Gamma L'$, which can be established elegantly and abstractly using the characterisation of $\alpha$-equivalence given above.

[4]This characterisation depends on the use of terms with explicit contexts and on the restriction to contexts in which all names are distinct.

A high-level summary of the above outlined approach is that one reasons with raw terms, making use of atomic sheaf logic and its equivalence and conditional independence relations to systematically subsume the necessary renaming of bound variables as instances of general logical principles.

The fact that atomic sheaf logic applies both to nominal sets (via the equivalence with $\mathrm{Sh}_{\mathrm{at}}(\mathbb{I}^{\mathrm{op}})$) and to probability (via $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$) means that one can compare the two approaches to nominal syntax, the standard one in which terms are $\alpha$-equivalence classes and the proposed one using raw terms, using an analogy with probability theory. When terms (with explicit context) are considered as $\alpha$-equivalence classes, they are, in particular, equated up to atomic equivalence (orbit equality). In the probabilistic setting of $\mathrm{Sh}_{\mathrm{at}}(\mathbb{SBP}_0)$, atomic equivalence is equality in distribution. So reasoning with $\alpha$-equivalence classes is analogous to doing probability with probability distributions. In contrast, our proposal to reason with raw terms and make use of the atomic equivalence and conditional independence relations is analogous to, in probability theory, reasoning with random variables and exploiting the relations of equality in law and conditional independence between them. Certainly, in mainstream probability theory, reasoning with random variables is usually considered more convenient than reasoning with probability distributions. It therefore seems worth investigating whether our proposed approach to reasoning about syntax will have similar practical advantages over the $\alpha$-equivalence-class-based approach. It is intended to carry out some case studies in this direction as future research.

### 10.3 Further work

We end the paper with two questions for potential further investigation on the theory side, of which the second was suggested by one of the journal referees. The first is to obtain a completeness theorem for the logic of equivalence and conditional independence valid in atomic toposes. The second is to investigate whether atomic sheaf logic enjoys a similar relationship to second-order logic as that enjoyed by dependence logic [42].

## Acknowledgments

I thank Angus Macintyre for drawing my attention to dependence logic, and André Joyal, Paul-André Melliès, Dario Stein and the anonymous reviewers of both conference and journal versions for helpful suggestions. I also thank Terblanche Delport, Willem Fouché and Paul and Petrus Potgieter for their hospitality in Pretoria in January 2023, where half this paper was written. Paul Taylor's diagram macros were used.

## References

[1] Samson Abramsky and Jouko Väänänen. 2009. From IF to BI: a tale of dependence and separation. *Synthese* 167 (2009), 207–230. https://doi.org/10.1007/s11229-008-9415-6

[2] Jialu Bao, Simon Docherty, Justin Hsu, and Alexandra Silva. 2021. A Bunched Logic for Conditional Independence. In *2021 36th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. 1–14. https://doi.org/10.1109/LICS52264.2021.9470712

[3] Jialu Bao, Marco Gaboardi, Justin Hsu, and Joseph Tassarotti. 2022. A separation logic for negative dependence. *Proc. ACM Program. Lang.* 6, POPL, Article 57 (Jan. 2022), 29 pages. https://doi.org/10.1145/3498719

[4] Timon Barlag, Miika Hannula, Juha Kontinen, Nina Pardal, and Jonni Virtema. 2023. Unified foundations of team semantics via semirings. In *Proceedings of the 20th International Conference on Principles of Knowledge Representation and Reasoning* (Rhodes, Greece) *(KR '23)*. Article 8, 11 pages. https://doi.org/10.24963/kr.2023/8

[5] Gilles Barthe, Justin Hsu, and Kevin Liao. 2019. A probabilistic separation logic. *Proc. ACM Program. Lang.* 4, POPL, Article 55 (dec 2019), 30 pages. https://doi.org/10.1145/3371123

[6] Fredrik Dahlqvist, Vincent Danos, Ilias Garnier, and Ohad Kammar. 2016. Bayesian Inversion by Omega-Complete Cone Duality. In *27th International Conference on Concurrency Theory (CONCUR 2016) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 59)*, Josée Desharnais and Radha Jagadeesan (Eds.). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 1:1–1:15. https://doi.org/10.4230/LIPIcs.CONCUR.2016.1

[7] A. P. Dawid. 1979. Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society. Series B (Methodological)* 41, 1 (1979), 1–31. http://www.jstor.org/stable/2984718

[8] A. P. Dawid. 2001. Separoids: A Mathematical Framework for Conditional Independence and Irrelevance. *Annals of Mathematics and Artificial Intelligence* 32, 1 (2001), 335–372. https://doi.org/10.1023/A:1016734104787

[9] Brian Day. 1970. *Construction of Biclosed Categories*. Ph. D. Dissertation.

[10] C. Dellacherie and P.-A. Meyer. 2011. *Probabilities and Potential, C: Potential Theory for Discrete and Continuous Semigroups*. North-Holland Mathematics Studies, Vol. 161. Elsevier Science.

[11] Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema. 2018. Approximation and dependence via multiteam semantics. *Annals of Mathematics and Artificial Intelligence* 83, 3 (2018), 297–320. https://doi.org/10.1007/s10472-017-9568-4

[12] Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema. 2018. Probabilistic team semantics. In *International Symposium on Foundations of Information and Knowledge Systems*. Springer, 186–206.

[13] Marcelo Fiore and Matías Menni. 2005. Reflective Kleisli subcategories of the category of Eilenberg-Moore algebras for factorization monads. *Theory and Applications of Categories [electronic only]* 15 (2005), 40–65. http://eudml.org/doc/125160

[14] Murdoch J. Gabbay and Andrew M. Pitts. 2002. A New Approach to Abstract Syntax with Variable Binding. *Formal Aspects of Computing* 13, 3 (2002), 341–363. https://doi.org/10.1007/s001650200016

[15] Dan Geiger, Azaria Paz, and Judea Pearl. 1991. Axioms and algorithms for inferences involving probabilistic independence. *Information and Computation* 91, 1 (1991), 128–141. https://doi.org/10.1016/0890-5401(91)90077-F

[16] Dan Geiger and Judea Pearl. 1993. Logical and Algorithmic Properties of Conditional Independence and Graphical Models. *The Annals of Statistics* 21, 4 (1993), 2001 – 2021. https://doi.org/10.1214/aos/1176349407

[17] Erich Grädel and Jouko Väänänen. 2013. Dependence and Independence. *Studia Logica* 101, 2 (2013), 399–410. https://doi.org/10.1007/s11225-013-9479-2

[18] Miika Hannula and Juha Kontinen. 2016. A finite axiomatization of conditional independence and inclusion dependencies. *Information and Computation* 249 (2016), 121–137. https://doi.org/10.1016/j.ic.2016.04.001

[19] Jaakko Hintikka and Gabriel Sandu. 1989. Informational Independence as a Semantical Phenomenon. In *Logic, Methodology and Philosophy of Science VIII*, Jens Erik Fenstad, Ivan T. Frolov, and Risto Hilpinen (Eds.). Studies in Logic and the Foundations of Mathematics, Vol. 126. Elsevier, 571–589. https://doi.org/10.1016/S0049-237X(08)70066-1

[20] W. Hodges. 1997. Compositional Semantics for a Language of Imperfect Information. *Logic Journal of the IGPL* 5, 4 (1997), 539–563. https://doi.org/10.1093/jigpal/5.4.539

[21] Tapani Hyttinen, Gianluca Paolini, and Jouko Väänänen. 2017. A logic for arguing about probabilities in measure teams. *Archive for Mathematical Logic* 56, 5 (2017), 475–489. https://doi.org/10.1007/s00153-017-0535-x

[22] Janez Ignacij Jereb and Alex Simpson. 2025. Safety, Relative Tightness and the Probabilistic Frame Rule. *Electronic Notes in Theoretical Informatics and Computer Science* Volume 5 - Proceedings of MFPS XLI, Article 12 (Dec 2025). https://doi.org/10.46298/entics.16743

[23] Peter T Johnstone. 2002. *Sketches of an elephant: a Topos theory compendium*. Oxford Univ. Press, New York, NY. https://cds.cern.ch/record/592033

[24] Olav Kallenberg. 1997. *Foundations of Modern probability Theory* (first edition ed.). Springer Cham.

[25] Dexter Kozen. 1981. Semantics of probabilistic programs. *J. Comput. System Sci.* 22, 3 (1981), 328–350. https://doi.org/10.1016/0022-0000(81)90036-2

[26] F. William Lawvere. 1969. Adjointness in Foundations. *Dialectica* 23, 3-4 (1969), 281–296. https://doi.org/10.1111/j.1746-8361.1969.tb01194.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1746-8361.1969.tb01194.x

[27] John M. Li, Amal Ahmed, and Steven Holtzen. 2023. Lilac: A Modal Separation Logic for Conditional Probability. *Proc. ACM Program. Lang.* 7, PLDI, Article 112 (jun 2023), 24 pages. https://doi.org/10.1145/3591226

[28] John M. Li, Jon Aytac, Philip Johnson-Freyd, Amal Ahmed, and Steven Holtzen. 2024. A Nominal Approach to Probabilistic Separation Logic. In *39th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. ACM, New York, NY, USA. https://doi.org/10.1145/3661814.3662135

[29] Saunders Mac Lane and Ieke Moerdijk. 1992. *Sheaves in Geometry and Logic a First Introduction to Topos Theory*. Springer New York, New York, NY. http://link.springer.com/book/10.1007/978-1-4612-0927-0

[30] Matías Menni. 2003. About И-quantifiers. *Applied Categorical Structures* 11, 5 (2003), 421–445. https://doi.org/10.1023/A:1025750816098

[31] Peter O'Hearn, John Reynolds, and Hongseok Yang. 2001. Local Reasoning about Programs that Alter Data Structures. In *Computer Science Logic*, Laurent Fribourg (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–19.

[32] Peter W. O'Hearn and David J. Pym. 1999. The Logic of Bunched Implications. *The Bulletin of Symbolic Logic* 5, 2 (1999), 215–244. http://www.jstor.org/stable/421090

[33] Judea Pearl and Azaria Paz. 1986. GRAPHOIDS: Graph-Based Logic for Reasoning about Relevance Relations OrWhen Would x Tell You More about y If You Already Know z? *Probabilistic and Causal Inference* (1986). https://api.semanticscholar.org/CorpusID:262089256

[34] Andrew M Pitts. 2001. Categorical logic. In *Handbook of Logic in Computer Science: Volume 5. Algebraic and Logical Structures*. Oxford University Press. https://doi.org/10.1093/oso/9780198537816.003.0002 arXiv:https://academic.oup.com/book/0/chapter/348287300/chapter-pdf/43264150/isbn-9780198537816-book-part-5.pdf

[35] Andrew M. Pitts. 2013. *Nominal Sets: Names and Symmetry in Computer Science*. Cambridge University Press.

[36] David J. Pym, Peter W. O'Hearn, and Hongseok Yang. 2004. Possible worlds and resources: the semantics of BI. *Theoretical Computer Science* 315, 1 (2004), 257–305. https://doi.org/10.1016/j.tcs.2003.11.020 Mathematical Foundations of Programming Semantics.

[37] Alex Simpson. 2017. Probability Sheaves and the Giry Monad. https://api.semanticscholar.org/CorpusID:11927690

[38] Alex Simpson. 2018. Category-theoretic Structure for Independence and Conditional Independence. *Electronic Notes in Theoretical Computer Science* 336 (2018), 281–297. https://doi.org/10.1016/j.entcs.2018.03.028 The Thirty-third Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIII).

[39] Alex Simpson. 2024. Equivalence and Conditional Independence in Atomic Sheaf Logic. In *Proceedings of the 39th Annual ACM/IEEE Symposium on Logic in Computer Science* (Tallinn, Estonia) *(LICS '24)*. Association for Computing Machinery, New York, NY, USA, Article 70, 14 pages. https://doi.org/10.1145/3661814.3662132

[40] Wolfgang Spohn. 1980. Stochastic independence, causal independence, and shieldability. *Journal of Philosophical Logic* 9, 1 (1980), 73–99. https://doi.org/10.1007/BF00258078

[41] Sam Staton. 2009. Two Cotensors in One: Presentations of Algebraic Theories for Local State and Fresh Names. *Electronic Notes in Theoretical Computer Science* 249 (2009), 471–490. https://doi.org/10.1016/j.entcs.2009.07.103 Proceedings of the 25th Conference on Mathematical Foundations of Programming Semantics (MFPS 2009).

[42] Jouko Väänänen. 2007. *Dependence Logic: A New Approach to Independence Friendly Logic*. Cambridge University Press.

[43] Hongseok Yang and Peter O'Hearn. 2002. A Semantic Basis for Local Reasoning. In *Foundations of Software Science and Computation Structures*, Mogens Nielsen and Uffe Engberg (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 402–416.

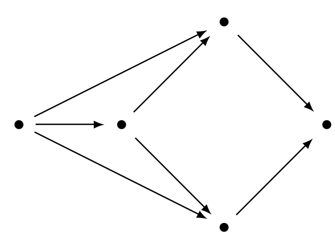
## A Proof of Theorem 7.7

Recall that Theorem 7.7 states that every sheaf in $\mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$ has supports. The main tool needed to prove this is Theorem A.1 below.

Theorem A.1. *Every atomic sheaf $\underline{P} \in \mathrm{Sh}_{\mathrm{at}}(\mathbb{Sur})$ maps pushouts in $\mathbb{Sur}$ to pullbacks in **Set**.*

The proof of Theorem A.1 given below builds on Theorem 6.6. When I discussed this work with André Joyal, he told me that he already knew Theorem A.1, and he kindly showed me his own proof, which is somewhat different in structure from the argument given below.

Observe first that the category $\mathbb{Sur}$ has pushouts, and that these are defined as in **Set**. Observe also that, in any commuting diagram in $\mathbb{Sur}$ of the form below, the outer kite is a pushout if and only if the right-hand square is a pushout (because all maps in $\mathbb{Sur}$ are epimorphic).



Lemma A.2. *Suppose we have a commuting diagram as above in $\mathbb{Sur}$. Let $P \in \mathrm{Psh}(\mathbb{Sur})$ be a separated presheaf. Then $P$ maps the right-hand square to a pullback in **Set** if and only if it maps the outer kite to a pullback in **Set**.*

Proof. Easy. □

Proof of Theorem A.1. A relation $R \subseteq \Omega_X \times \Omega_Y$ is said to be *bitotal* if:

$$\forall \omega_X \in \Omega_X.\, \exists \omega_Y \in \Omega_Y,\ \omega_X R \omega_Y \text{ and } \forall \omega_Y \in \Omega_Y.\, \exists \omega_X \in \Omega_X,\ xRy\ .$$

Let $R \subseteq \Omega_X \times \Omega_Y$ be a bitotal relation. Then the projections $\Omega_X \xleftarrow{r} R \xrightarrow{r'} \Omega_Y$ form a span in $\mathbb{S}\mathrm{ur}$. Construct the pushout

$$\begin{array}{ccc} R & \xrightarrow{r} & \Omega_X \\ {\scriptstyle r'}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ \Omega_Y & \xrightarrow[q]{} & \Omega_Z \end{array} \tag{42}$$

For any $n \geq 0$ define $R_n \subseteq \Omega_X \times \Omega_X$ and $S_n \subseteq \Omega_X \times \Omega_X$ by: $R_0 := R$ and $S_n = R^{-1} \circ R_n$ and $R_{n+1} := R \circ S_n$. Let $r_n \colon R_n \longrightarrow \Omega_X$ and $r'_n \colon R_n \longrightarrow \Omega_Y$ be the first and second projections and similarly for $s_n \colon S_n \longrightarrow \Omega_X$ and $s'_n \colon S_n \longrightarrow \Omega_X$ . Alternatively, we can formulate this in diagrammatic terms, taking pullbacks for both top-left squares below,

$$\begin{array}{ccccc} U_n & \xrightarrow{u'_n} & R & \xrightarrow{r} & \Omega_X \\ {\scriptstyle u_n}\big\downarrow & & \big\downarrow{\scriptstyle r'} & & \vdots{\scriptstyle p} \\ R_n & \xrightarrow{r'_n} & \Omega_Y & \overset{q}{\cdots\to} & \Omega_Z \\ {\scriptstyle r_n}\big\downarrow & & \vdots{\scriptstyle q} & & \vdots{\scriptstyle \mathrm{id}_{\Omega_Z}} \\ \Omega_X & \underset{p}{\cdots\to} & \Omega_Z & \underset{\mathrm{id}_{\Omega_Z}}{\cdots\to} & \Omega_Z \end{array} \qquad \begin{array}{ccccc} T_{n+1} & \xrightarrow{t'_{n+1}} & R & \xrightarrow{r'} & \Omega_Y \\ {\scriptstyle t_{n+1}}\big\downarrow & & \big\downarrow{\scriptstyle r} & & \vdots{\scriptstyle q} \\ S_n & \xrightarrow{s'_n} & \Omega_X & \overset{p}{\cdots\to} & \Omega_Z \\ {\scriptstyle s_n}\big\downarrow & & \vdots{\scriptstyle p} & & \vdots{\scriptstyle \mathrm{id}_{\Omega_Z}} \\ \Omega_X & \underset{p}{\cdots\to} & \Omega_Z & \underset{\mathrm{id}_{\Omega_Z}}{\cdots\to} & \Omega_Z \end{array}$$

and defining the relations $(s_n, s'_n) : S_n \rightarrowtail \Omega_X \times \Omega_X$ and $(r_{n+1}, r'_{n+1}) : R_n \rightarrowtail \Omega_X \times \Omega_X$ as the following epi-mono factorisations in **Set**

$$U_n \overset{u_n^S}{\twoheadrightarrow} S_n \overset{(s_n,s'_n)}{\rightarrowtail} \Omega_X \times \Omega_X = U_n \xrightarrow{(r_n \circ u_n,\, r \circ u'_n)} \Omega_X \times \Omega_X$$

$$T_{n+1} \overset{t_{n+1}^R}{\twoheadrightarrow} R_n \overset{(r_{n+1},r'_{n+1})}{\rightarrowtail} \Omega_X \times \Omega_Y = T_{n+1} \xrightarrow{(s_n \circ t_{n+1},\, r' \circ t'_{n+1})} \Omega_X \times \Omega_Y$$

We first claim that, for any $n \geq 0$, both diagrams below commute.

$$\begin{array}{ccc} S_n & \xrightarrow{s_n} & \Omega_X \\ {\scriptstyle s'_n}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ \Omega_X & \xrightarrow[p]{} & \Omega_Z \end{array} \qquad \begin{array}{ccc} R_n & \xrightarrow{r_n} & \Omega_X \\ {\scriptstyle r'_n}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ \Omega_Y & \xrightarrow[q]{} & \Omega_Z \end{array} \tag{43}$$

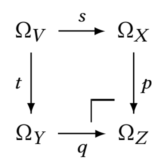
This first claim is proved by a straightforward induction on $n$. For example, one can use the induction hypothesis to complete the diagrams involving $U_n$ and $T_{n+1}$ above with the dotted arrows.

Our second claim is that, for some $n \geq 0$, the right-hand square of (43) is a pullback in **Set**. (The same holds for the left-hand square, but we shall not need this.) This holds because the fibres of the pushout maps $p$ and $q$ from (42) are the connected components in the bipartite graph $R \subseteq \Omega_X \times \Omega_Y$ restricted to $\Omega_X$ and $\Omega_Y$ respectively, and the $R_n$ construction approximates the path relation from below, necessarily reaching a fixed point at some finite $n$.

Our third claim is that every atomic sheaf $\underline{P} \in \mathrm{Sh}_{\mathrm{at}}(\mathbb{S}\mathrm{ur})$ maps the pushout diagram (42) to a pullback in **Set**. For this, let $x \in \underline{P}(\Omega_X)$ and $y \in \underline{P}(\Omega_Y)$ be such that $x \cdot r = y \cdot r'$. We prove, by induction on $n$ that $x \cdot r_n = y \cdot r'_n$ and $x \cdot s_n = x \cdot s'_n$ for all $n$. For $n = 0$, we have $r_0 = r$ and $r'_0 = r'$ so indeed $x \cdot r_0 = y \cdot r'_0$. Next, assuming $x \cdot r_n = y \cdot r'_n$, we show $x \cdot s_n = x \cdot s'_n$. For this, we have $x \cdot s_n \cdot u_n^S = x \cdot r_n \cdot u_n = y \cdot r'_n \cdot u_n = y \cdot r' \cdot u'_n = x \cdot r \cdot u'_n = x \cdot s'_n \cdot u_n^S$; whence by separatedness $x \cdot s_n = x \cdot s'_n$. Similarly, assuming $x \cdot s_n = x \cdot s'_n$, we show $x \cdot r_{n+1} = y \cdot r'_{n+1}$. For this, we have $x \cdot r_{n+1} \cdot t_{n+1}^R = x \cdot s_n \cdot t_{n+1} = x \cdot s'_n \cdot t_{n+1} = x \cdot r \cdot t'_{n+1} = y \cdot r' \cdot t'_{n+1} = y \cdot r'_{n+1} \cdot t_{n+1}^R$; whence by separatedness

$x \cdot r_{n+1} = y \cdot r'_{n+1}$. This completes the argument by induction. The second claim above now gives us $n$ such that the right-hand square of (43) is a pullback, in **Set**, hence an independent square in $\mathbb{S}\mathrm{ur}$. By Theorem 6.6, the square is mapped by $\underline{P}$ to a pullback in **Set**. By the statement proved by induction, $x \cdot r_n = y \cdot r'_n$. So, by the pullback property in **Set**, there exists a unique $z \in \underline{P}(\Omega_Z)$ such that $z \circ p = x$ and $z \circ q = y$, which is what we needed to show to establish the third claim.

We now establish the property asserted by the theorem. Consider any pushout diagram in $\mathbb{S}\mathrm{ur}$.




Define $R \subseteq \Omega_X \times \Omega_Y$ to be the image $(s,t)(\Omega_V)$. Since $s, t$ are surjective, $R$ is bitotal. By the observations at the start of this section, (42) is also a pushout. By the third claim above, $\underline{P}$ maps (42) to a pullback in **Set**. This property transfers to the original pushout, by Lemma A.2. □

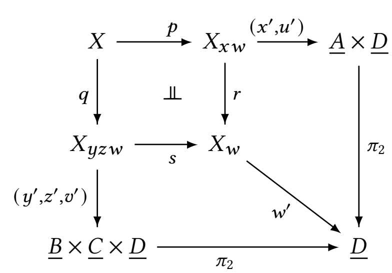
Proof of Theorem 7.7. Given a sheaf $\underline{P} \in \mathrm{Sh}_{\mathrm{at}}(\mathbb{S}\mathrm{ur})$ and element $x \in \underline{P}(\Omega_X)$, the support is obtained by taking a joint pushout in $\mathbb{S}\mathrm{ur}$ of all (inequivalent) representable factorisations of $x$, of which there are only finitely many (because there are only finitely many partitions of $\Omega_X$). By Theorem A.1, this joint pushout is itself a representable factorisation of $x$. □

## B Validity of axioms (28) and (29) from Figure 4

The lemma below establishes the validity of axiom (28).

Lemma B.1. *Suppose* $(x,(y,z),w) \in \perp\!\!\!\perp_{\underline{A},\,\underline{B}\times\underline{C}\,|\,\underline{D}}(X)$ *then* $(x,y,(z,w)) \in \perp\!\!\!\perp_{\underline{A},\underline{B}\,|\,\underline{C}\times\underline{D}}(X)$.

Proof. If $(x,(y,z),w) \in \perp\!\!\!\perp_{\underline{A},\,\underline{B}\times\underline{C}\,|\,\underline{D}}(X)$ then we have a hybrid diagram




where $x' \cdot p = s$ and $y' \cdot q = y$ and $z' \cdot q = z$ and $(X_w,\, r \circ p,\, w')$ is support for $w$ and, without loss of generality, $(X_{xw}, p, (x', u'))$ is support for $(x, z)$ and $(X_{yzw}, q, (y', z', v'))$ is support for $(y, z, w)$.

The independent square in the diagram above can be factorised as a composite of two commuting squares as in the top row below

$$\begin{array}{ccccc} X & \xrightarrow{p'} & X_{xzw} & \xrightarrow{p''} & X_{xw} \\ {\scriptstyle q}\downarrow & & \downarrow{\scriptstyle t} & & \downarrow{\scriptstyle r} \\ X_{yzw} & \xrightarrow[s']{} & X_{zw} & \xrightarrow[s'']{} & X_w \\ {\scriptstyle s'}\downarrow & & \downarrow{\scriptstyle \mathsf{id}} & & \downarrow{\scriptstyle \mathsf{id}} \\ X_{zw} & \xrightarrow[\mathsf{id}]{} & X_{zw} & \xrightarrow[s'']{} & X_w \end{array} \tag{44}$$

where all objects are defined as the supports indicated by their names. For example, $(X_{zw},\ s' \circ q,\ (z'', w''))$ is support for $(z, w)$ and $(X_{xzw},\ p', (x' \cdot p'',\ z'' \cdot t,\ u' \cdot p''))$ is support for $(x, z, w)$. We show that the top-right square is an independent pullback.

To see it is independent, observe that the full composite square (44) above is a composite of an independent top-row rectangle with the two independent squares in the bottom row. So (44) is independent. That is, the square

$$\begin{array}{ccc} X & \xrightarrow{p''\circ p'} & X_{xw} \\ {\scriptstyle t\circ p'}\downarrow & & \downarrow{\scriptstyle r} \\ X_{zw} & \xrightarrow[s'']{} & X_w \end{array}$$

is independent. It thus follows from the descent property that the top-right square in (44) is independent.

For the independent pullback property, consider any independent pullback of $r$ along $s''$

$$\begin{array}{ccc} Y & \xrightarrow{h} & X_{xw} \\ {\scriptstyle k}\downarrow & & \downarrow{\scriptstyle r} \\ X_{zw} & \xrightarrow[s'']{} & X_w \end{array}$$

Since the top-right square of (44) is independent, there exists $j : X_{xzw} \to Y$ such that $k \circ j = t$ and $j \circ h = p''$. This gives us a representable factorisation $(Y,\ j \circ p',\ (x' \cdot h,\ z'' \cdot k,\ u' \cdot h))$ of $(x, z, w)$. Since $(X_{xzw},\ p',\ (x' \cdot p'',\ z'' \cdot t,\ u' \cdot p''))$ is support for $(x, z, w)$, we obtain a map $i : Y \to X_{xzw}$ of representable factorisations. However $j$ is also a map of representable factorisations in the opposite direction, so $i$ and $j$ are mutual inverses. Thus the top-right square in (44) is indeed an independent pullback.

Since the top-row rectangle of (44) is independent and the top-right square an independent pullback it follows that the top-left square is independent. Using this, we form the hybrid diagram

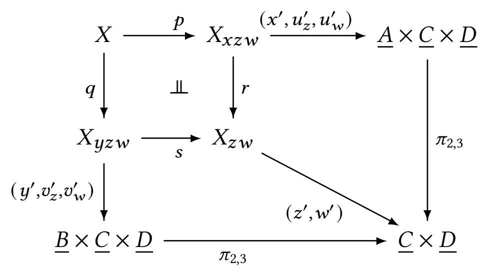
$$\begin{array}{ccccc} X & \xrightarrow{p'} & X_{xzw} & \xrightarrow{(x'\cdot p'',\, z''\cdot t,\, u'\cdot p'')} & \underline{A}\times\underline{C}\times\underline{D} \\ {\scriptstyle q}\downarrow & \perp\!\!\!\perp & \downarrow{\scriptstyle t} & & \downarrow{\scriptstyle \pi_{2,3}} \\ X_{yzw} & \xrightarrow[s']{} & X_{zw} & \xrightarrow{(z'',w'')} & \underline{D} \\ {\scriptstyle (y',z',v')}\downarrow & & & & \\ \underline{B}\times\underline{C}\times\underline{D} & & \xrightarrow[\pi_{2,3}]{} & & \underline{D} \end{array}$$

showing that indeed $(x, y, (z, w)) \in \perp\!\!\!\perp_{\underline{A},\underline{B}\,|\,\underline{C}\times\underline{D}}(X)$. ◻

The lemma below establishes the validity of axiom (29).

Lemma B.2. *Suppose* $(x, y, (z, w)) \in \perp\!\!\!\perp_{\underline{A},\underline{B}\,|\,\underline{C}\times\underline{D}}(X)$ *and* $(x, z, w) \in \perp\!\!\!\perp_{\underline{A},\underline{C}\,|\,\underline{D}}(X)$ *then* $(x, (y, z), w) \in \perp\!\!\!\perp_{\underline{A},\underline{B}\times\underline{C}\,|\,\underline{D}}(X)$.

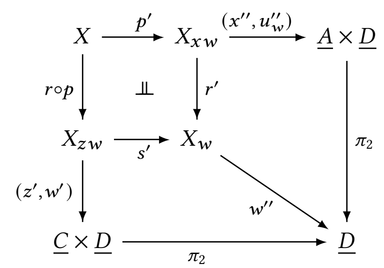
Proof. The assumption $(x, y, (z, w)) \in \perp\!\!\!\perp_{\underline{A},\underline{B}\,|\,\underline{C}\times\underline{D}}(X)$ gives us:



(45)

where $x' \cdot p = x$ and $y' \cdot q = y$ and $(X_{zw}, r \circ p, (z', w'))$ is support for $(z, w)$ and, without loss of generality, $(X_{xzw}, p, (x', u'_z, u'_w))$ is support for $(x, z, w)$ and $(X_{yzw}, q, (y', v'_z, v'_w))$ is support for $(y, z, w)$.

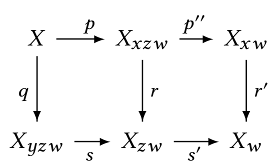
Similarly, the assumption $(x, z, w) \in \perp\!\!\!\perp_{\underline{A},\underline{C}\,|\,\underline{D}}(X)$ gives us:



(46)

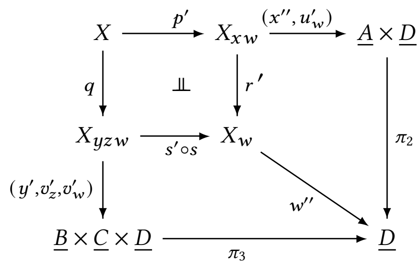
where $x'' \cdot p' = x$ and $z'' \cdot q' = z$ and $(X_w, r' \circ p', w'')$ is support for $w$ and, without loss of generality, $(X_{xw}, p', (x'', u''_w))$ is support for $(x, w)$ and we can use $r \circ p$ because $(X_{zw}, r \circ p, (z', w'))$ is support for $(z, w)$.

Exploiting the support property of $X_{xw}$, we obtain $p''$ in



such that $p'' \circ p = p'$. The left-hand square above is the independent square from (45). Since $p'' \circ p = p'$, the right-hand square is also independent, by descent along $p$ from the independent square in (46). So the composite rectangle is independent.

The composite rectangle provides the independent square in




showing that $(x, (y, z), w) \in ⫫_{\underline{A}, \underline{B}\times\underline{C} \mid \underline{D}}(X)$ as required. □

## C Proof of Proposition 8.12

The goal of the section is to prove Proposition 8.12, which states that Definition 8.11 endows $\mathbb{SBP}_0$ with independent pullback structure satisfying the descent property.

Recall that Definition 8.11 defines a commuting square (37) in $\mathbb{SBP}_0$ to be *independent* if $p ⫫ q \mid r \circ p$ according to Definition 8.7. Since the square is commuting, the property in Definition 8.7 simplifies to: for every $S \in \mathcal{B}_{\Omega_Y}$ and $T \in \mathcal{B}_{\Omega_Z}$, and for $P_{\Omega_W}$-almost all $\omega \in \Omega_W$,

$$P_{(r\circ p)^{-1}(\omega)}(p^{-1}(S) \cap q^{-1}(T)) \;=\; P_{r^{-1}(\omega)}(S) \cdot P_{s^{-1}(\omega)}(T) \;\; . \tag{47}$$

The key proposition below characterises the independence of (37) as being equivalent to $p$, considered as a map on fibre sets $q^{-1}(\omega_Z) \to r^{-1}(s(\omega_Z))$, preserving the disintegration-induced probability measures, for almost all $\omega_Z$.

Proposition C.1. *A commuting square in $\mathbb{SBP}_0$ (37) is independent if and only if, for $P_{\Omega_Z}$-almost-all $\omega_Z \in \Omega_Z$, it holds that $p_*(P_{q^{-1}(\omega_Z)}) = P_{r^{-1}(s(\omega_Z))}$ .*

Proof. We first prove the right-to-left implication. Accordingly, suppose $p_*(P_{q^{-1}(\omega_Z)}) = P_{r^{-1}(s(\omega_Z))}$ holds for $P_{\Omega_Z}$-almost-all $\omega_Z \in \Omega_Z$. For $P_{\Omega_W}$-almost every $\omega \in \Omega_W$, we prove (47) by

$$\begin{aligned}
& P_{(r\circ p)^{-1}(\omega)}(p^{-1}(S) \cap q^{-1}(T)) \\
&\quad = \int P_{q^{-1}(\omega_Z)}(p^{-1}(S) \cap q^{-1}(T))\, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \\
&\quad = \int \mathbb{1}_T(\omega_Z) \cdot P_{q^{-1}(\omega_Z)}(p^{-1}(S))\, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \\
&\quad = \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(s(\omega_Z))}(S)\, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) && \text{because } p_*(P_{q^{-1}(\omega_Z)}) = P_{r^{-1}(s(\omega_Z))} \\
&\quad = \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(\omega)}(S)\, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \\
&\quad = P_{r^{-1}(\omega)}(S) \cdot \int \mathbb{1}_T(\omega_Z)\, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \\
&\quad = P_{r^{-1}(\omega)}(S) \cdot P_{s^{-1}(\omega)}(T) \;\; .
\end{aligned}$$

For the left-to-right implication, suppose (47) holds, for $P_{\Omega_W}$-almost every $\omega \in \Omega_W$. Note that, for any $S \in \mathcal{B}_{\Omega_Y}$ the function

$$T \;\mapsto\; \int \mathbb{1}_T(\omega_Z) \cdot P_{q^{-1}(\omega_Z)}(p^{-1}(S))\, \mathrm{d}P_{\Omega_Z}(\omega_Z)$$

is a measure $\mathcal{B}_{\Omega_Z} \to [0,1]$ with density

$$\omega_Z \mapsto P_{q^{-1}(\omega_Z)}(p^{-1}(S)) \ .$$

Similarly, the function

$$T \mapsto \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(s(\omega_Z))}(S) \, \mathrm{d}P_{\Omega_Z}(\omega_Z)$$

is a measure with density

$$\omega_Z \mapsto P_{r^{-1}(s(\omega_Z))}(S) \ .$$

Below we prove

$$\int \mathbb{1}_T(\omega_Z) \cdot P_{q^{-1}(\omega_Z)}(p^{-1}(S)) \, \mathrm{d}P_{\Omega_Z}(\omega_Z) = \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(s(\omega_Z))}(S) \, \mathrm{d}P_{\Omega_Z}(\omega_Z) \ , \tag{48}$$

which establishes that the two measures are equal, and hence their densities are almost surely equal. That is, for $P_{\Omega_Z}$-almost-all $\omega_Z \in \Omega_Z$, we have $P_{q^{-1}(\omega_Z)}(p^{-1}(S)) = P_{r^{-1}(s(\omega_Z))}(S)$, for all $S \in \mathcal{B}_{\Omega_Y}$. That is, $p_*(P_{q^{-1}(\omega_Z)}) = P_{r^{-1}(s(\omega_Z))}$, as required.

It remains to prove (48). For this, we calculate

$$\begin{aligned}
&\int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(s(\omega_Z))}(S) \, \mathrm{d}P_{\Omega_Z}(\omega_Z) \\
&= \int \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(s(\omega_Z))}(S) \, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \, \mathrm{d}P_\Omega(\omega) \\
&= \int \int \mathbb{1}_T(\omega_Z) \cdot P_{r^{-1}(\omega)}(S) \, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \, \mathrm{d}P_\Omega(\omega) \\
&= \int P_{r^{-1}(\omega)}(S) \cdot \left( \int \mathbb{1}_T(\omega_Z) \, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \right) \mathrm{d}P_\Omega(\omega) \\
&= \int P_{r^{-1}(\omega)}(S) \cdot P_{s^{-1}(\omega)}(T) \, \mathrm{d}P_\Omega(\omega) \\
&= \int P_{(s \circ q)^{-1}(\omega)}(p^{-1}(S) \cap q^{-1}(T)) \, \mathrm{d}P_\Omega(\omega) && \text{by (47)} \\
&= \int \int P_{q^{-1}(\omega_Z)}(p^{-1}(S) \cap q^{-1}(T)) \, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \, \mathrm{d}P_\Omega(\omega) \\
&= \int \int \mathbb{1}_T(\omega_Z) \cdot P_{q^{-1}(\omega_Z)}(p^{-1}(S)) \, \mathrm{d}P_{s^{-1}(\omega)}(\omega_Z) \, \mathrm{d}P_\Omega(\omega) \\
&= \int \mathbb{1}_T(\omega_Z) \cdot P_{q^{-1}(\omega_Z)}(p^{-1}(S)) \, \mathrm{d}P_{\Omega_Z}(\omega_Z) \ .
\end{aligned}$$

For any fixed $S \in \mathcal{B}_{\Omega_Y}$ the function mapping any $T$ to the left-hand side of (48) is clearly a measure $\mathcal{B}_{\Omega_Z} \to [0,1]$. □

We now verify that independent squares in $\mathbb{SBP}_0$ indeed satisfy the axioms for independent pullback structure. Axioms (IP1) and (IP2) are straightforward. Axiom (IP3) is an easy consequence of Proposition C.1. For Axiom (IP5), it is not difficult to verify that (38) indeed constructs an independent pullback square. The descent property is also straightforward. This leaves us with (IP4), which is established in greater generality by the proposition below.

Proposition C.2. *In a commuting diagram in $\mathbb{SBP}_0$ as below, if both the composite rectangle (AB) and right-hand square (B) are independent and $[q]$, $[t]$ are also jointly monic, then the left-hand square (A) is independent.*

$$\begin{array}{ccccc} \Omega_X & \xrightarrow{[s]} & \Omega_Y & \xrightarrow{[t]} & \Omega_Z \\ {\scriptstyle [p]}\big\downarrow & (A) & \big\downarrow{\scriptstyle [q]} & (B) & \big\downarrow{\scriptstyle [r]} \\ \Omega_U & \xrightarrow[{[u]}]{} & \Omega_V & \xrightarrow[{[v]}]{} & \Omega_W \end{array}$$

Proof. We use Proposition C.1 to prove that (A) is independent. That is, we show that, for $P_{\Omega_U}$-almost-all $\omega_U \in \Omega_U$, and for all $C \in \mathcal{B}_{\Omega_Y}$,

$$P_{p^{-1}(\omega_U)}(s^{-1}(C)) \;=\; P_{q^{-1}(u(\omega_U))}(C) \;\;. \tag{49}$$

We show this first for $C$ of the form $t^{-1}(A) \cap q^{-1}(B)$, where $A \in \mathcal{B}_{\Omega_Z}$ and $B \in \mathcal{B}_{\Omega_V}$. In this case, we have

$$\begin{aligned} & P_{p^{-1}(\omega_U)}(s^{-1}t^{-1}(A) \cap s^{-1}q^{-1}(B)) \\ &\quad = P_{p^{-1}(\omega_U)}(s^{-1}t^{-1}(A) \cap p^{-1}u^{-1}(B)) \\ &\quad = \mathbb{1}_{u^{-1}(B)}(\omega_U) \cdot P_{p^{-1}(\omega_U)}(s^{-1}t^{-1}(A)) \\ &\quad = \mathbb{1}_B(u(\omega_U)) \cdot P_{p^{-1}(\omega_U)}(s^{-1}t^{-1}(A)) \\ &\quad = \mathbb{1}_B(u(\omega_U)) \cdot P_{r^{-1}(v(u(\omega_U)))}(A) && \text{by Proposition C.1 for (AB)} \\ &\quad = \mathbb{1}_B(u(\omega_U)) \cdot P_{q^{-1}(u(\omega_U))}(t^{-1}(A)) && \text{by Proposition C.1 for (B)} \\ &\quad = P_{q^{-1}(u(\omega_U))}(t^{-1}(A) \cap q^{-1}(B)) \;\;. \end{aligned}$$

The joint monicity of $[q]$ and $[t]$ means that the there is a measure 1 set $S \in \mathcal{B}_{\Omega_Y}$ such that the paired function $(t, q) : S \to \Omega_Z \times \Omega_V$ is injective. Since $S \subseteq \Omega_Y$ is Borel, the standard Borel structure on $\Omega_Y$ restricts to $S$, and $(t, q)$ is a measurable embedding of the standard Borel space $S$ into the product standard Borel space $\Omega_Z \times \Omega_V$. Thus every Borel subset of $S$ is the restriction of a Borel subset of $\Omega_Z \times \Omega_V$. Since the $\sigma$-algebra of Borel subsets of $\Omega_Z \times \Omega_V$ is generated by Borel rectangles $A \times B$, it follows that the Borel subsets of $S$ are generated by sets of the form $S \cap (t^{-1}(A) \cap q^{-1}(B))$. Moreover, such sets are closed under finite intersections.

The left-hand and right-hand sides of (49) define measures $C \mapsto P_{p^{-1}(\omega_U)}(s^{-1}(C))$ and $P_{q^{-1}(u(\omega_U))}$ respectively. By the equality we have shown for $C$ of the form $t^{-1}(A) \cap q^{-1}(B)$, these measures agree on a generating set for $\mathcal{B}_{\Omega_Y}$ (restricted to $S$) that is closed under finite intersections. The two measures are therefore equal. This proves (49). ◻