Aspects of Predicative Algebraic Set Theory III: Sheaves

Benno van den Berg* & Ieke Moerdijk†

September 29, 2011

Abstract

This is the third installment in a series of papers on algebraic set theory. In it, we develop a uniform approach to sheaf models of constructive set theories based on ideas from categorical logic. The key notion is that of a "predicative category with small maps" which axiomatises the idea of a category of classes and class morphisms, together with a selected class of maps whose fibres are sets (in some axiomatic set theory). The main result of the present paper is that such predicative categories with small maps are stable under internal sheaves. We discuss the sheaf models of constructive set theory this leads to, as well as ideas for future work.[1]

1 Introduction

This is the third in a series of papers on algebraic set theory, the aim of which is to develop a categorical semantics for constructive set theories, including predicative ones, based on the notion of a "predicative category with small maps".[2] In the first paper in this series [9] we discussed how these predicative categories with small maps provide a sound and complete semantics for constructive set theory. In the second one [12], we explained how realizability extensions of such predicative categories with small maps can be constructed. The purpose of the present paper is to do the same for sheaf-theoretic extensions. This program was summarised in [11], where we announced the results that we will present and prove here.

*Mathematisch Instituut, Universiteit Utrecht, PO. Box 80010, 3508 TA Utrecht, the Netherlands. Email address: B.vandenBerg1@uu.nl. Supported by Netherlands Organisation for Scientific Research (NWO project "The Model Theory of Constructive Proofs").

†Corresponding author. Institute for Mathematics, Astrophysics and Particle Physics, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, the Netherlands. Email address: i.moerdijk@math.ru.nl.

[1]MCS: 18F20; 03F50; 03E70.

[2]Accessible and well-written introductions to algebraic set theory are [5, 6, 32].

For the convenience of the reader, and also to allow a comparison with the work by other researchers, we outline the main features of our approach. As said, the central concept in our theory is that of a predicative category with small maps. It axiomatises the idea of a category whose objects are classes and whose morphisms are functions between classes, and which is moreover equipped with a designated class of maps. The maps in the designated class are called small, and the intuitive idea is that the fibres of these maps are sets (in a certain axiomatic set theory). Such categories are in many ways like toposes, and to a large extent the purpose of our series of papers is to develop a topos theory for these categories. Indeed, like toposes, predicative categories with small maps turn out to be closed under realizability and sheaves.

On the other hand, where toposes can be seen as models of a typed version of (constructive) higher-order arithmetic, predicative categories with small maps provide models of (constructive) *set theories*. Furthermore, the notion of a predicative category with small maps is proof-theoretically rather weak: this allows us to model set theories which are proof-theoretically weaker than higher-order arithmetic, such as Aczel's set theory **CZF** (see [1]). But at the same time, the notion of a predicative category with small maps can also be strengthened, so that it leads to models of set theories proof-theoretically stronger than higher-order arithmetic, like **IZF**. The reason for this is that one can impose additional axioms on the class of small maps. This added flexibility is an important feature of algebraic set theory.

A central result in algebraic set theory says that the semantics provided by predicative categories with small maps is complete. More precisely, every predicative category with small maps contains an object ("the initial ZF-algebra" in the terminology of [24], or "the initial $\mathcal{P}_s$-algebra" in the terminology of [9][3]) which carries the structure of a model of set theory. Which set-theoretic axioms hold in this model depends on the properties of the class of small maps and on the logic of the underlying category: in different situations, this initial ZF-algebra can be a model of **CZF**, of **IZF**, or of ordinary **ZF**. (The axioms of the constructive set theores **CZF** and **IZF** are recalled in Section 2 below.) The completeness referred to above results from the fact that from the syntax of **CZF** (or **(I)ZF**), we can build a predicative category with small maps with the property that in the initial ZF-algebra in this category, precisely those sentences are valid which are derivable from the axiom of **CZF** (see [9]). (Completeness theorems of this kind go back to [31, 6]. One should also mention that one can obtain a predicative category with small maps from the syntax of Martin-Löf type theory: Aczel's interpretation of **CZF** in Martin-Löf type theory goes precisely via the initial ZF-algebra in this category. In fact, our proof of the existence of the initial ZF-algebra in any predicative category with small maps in [9] was modelled on Aczel's interpretation, as it was in [29].)

In algebraic set theory we approach the construction of realizability cate-

[3]Appendix A in [24] contains a proof of the fact that both these terms refer to the same object. In the sequel we will use these terms interchangably.

gories and of categories of sheaves in a topos-theoretic spirit; that is, we regard these realizability and sheaf constructions as *closure properties* of predicative categories with small maps. For realizability this means that starting from any predicative category with small maps $(\mathcal{E}, \mathcal{S})$ one can build a predicative realizability category with small maps $(\mathcal{E}ff_{\mathcal{E}}, \mathcal{E}ff_{\mathcal{S}})$ over it. Inside both of these categories, we have models of constructive set theory (**CZF** say), as shown in the following picture. Here, the vertical arrows are two instances of the same construction of the initial ZF-algebra, applied to different predicative categories with small maps:

$$\begin{tikzcd}
(\mathcal{E}, \mathcal{S}) \arrow[r] \arrow[d] & (\mathcal{E}ff_{\mathcal{E}}, \mathcal{E}ff_{\mathcal{S}}) \arrow[d] \\
\text{model of } \mathbf{CZF} \arrow[r] & \text{realizability model of } \mathbf{CZF}
\end{tikzcd}$$

Traditional treatments of realizability either regard it as a model-theoretic construction (which would correspond to the lower edge of the diagram), or as a proof-theoretic interpretation (defining a realizability model of **CZF** inside **CZF**, as in [30], for instance): the latter would correspond to the left-hand vertical arrow in the special case where $\mathcal{E}$ is the syntactic category associated to **CZF**. So in a way our treatment captures both constructions in a uniform way.

That realizability is indeed a closure property of predicative categories with small maps was the principal result of [9]. The main result of the present paper is that the same is true for sheaves, leading to an analogous diagram:

$$\begin{tikzcd}
(\mathcal{E}, \mathcal{S}) \arrow[r] \arrow[d] & (\mathrm{Sh}_{\mathcal{E}}, \mathrm{Sh}_{\mathcal{S}}) \arrow[d] \\
\text{model of } \mathbf{CZF} \arrow[r] & \text{sheaf model of } \mathbf{CZF}
\end{tikzcd}$$

The main technical difficulty in showing that predicative categories with small maps are closed under sheaves lies in showing that the axioms concerning inductive types (W-types) and an axiom called "fullness" (needed to model the subset collection axiom of **CZF**) are inherited by sheaf models. The proofs of these facts are quite long and involved, and take up a large part of this paper (the situation for realizability was very similar).

To summarise, in our approach there is one uniform construction of a model out of a predicative category with small maps $(\mathcal{E}, \mathcal{S})$, which one can apply to different kinds of such categories, constructed using syntax, using realizability, using sheaves, or any iteration or combination of these techniques.

We proceed to compare our results with those of other authors. Early work on categorical semantics of set theory (for example, [16] and [15]) was concerned with sheaf and realizability toposes defined over $\mathcal{S}ets$. The same applies to the book which introduced algebraic set theory [24]. In particular, to the best of our knowledge, before our work a systematic account was lacking of iterations and combinations of realizability and sheaf interpretations. In addition these earlier papers were concerned exclusively with impredicative set theories, such as **ZF** or **IZF**: the only exception seems to have been an early paper [21] by Grayson, treating models of predicative set theory in the context of what would now be called formal topology.

The first paper extending the methods of algebraic set theory to predicative systems was [29]. The authors of this paper showed how categorical models of Martin-Löf type theory (with universes) lead to models of **CZF** extended with a choice principle, which they dubbed the Axiom of Multiple Choice **(AMC)**. They established how such categorical models of type theory are closed under sheaves, hence leading to sheaf models of a strengthening of **CZF**. They did not develop a semantics for **CZF** *per se* and relied on a technical notion of a collection site, which we manage to avoid here (moreover, there was a mistake in their treatment of W-types of sheaves; we correct this in Section 4.4 below, see also [10]).

Two accounts of presheaf models in the context of algebraic set theory have been written by Gambino [18] and Warren [34]. In [18] Gambino shows how an earlier (unpublished) construction of a model of constructive set theory by Dana Scott can be regarded as an initial ZF-algebra in a category of presheaves, and that one can perform the construction in a predicative metatheory as well. Warren shows in [34] that many of the axioms that we will discuss are inherited by categories of coalgebras for a Cartesian comonad, a construction which includes presheaf models as a special case. But note that neither of these authors discusses the technically complicated axioms concerning W-types and fullness, as we will do in Sections 3 and 4 below.

In his PhD thesis [17], Gambino gave a systematic account of Heyting-valued models for **CZF** (see also [19]). This work was in the context of formal topology (essentially, sites whose underlying categories are posets). He has subsequently worked on generalising this to arbitrary sites and on putting this in the context of algebraic set theory. In [20], he took the first step in constructing the sheafification functor and in [7], written together with Awodey, Lumsdaine and Warren, he checks that the basic axioms for small maps are inherited by categories of sheaves in the general setting of sheaves for a Lawvere-Tierney topology. We will extend these results by proving that for sites which have a presentation (for a definition, see Definition 4.1 below), the axioms for W-types and for fullness are stable under taking sheaf extensions. Note that for proof-theoretic reasons, fullness cannot be stable under taking more general kinds of sheaves such as those for a site which does not have a presentation, or for a Lawvere-Tierney topology. The point is that **CZF** extended with the Law of Excluded Middle gives **ZF**, a much stronger system proof-theoretically, and therefore a double-negation interpretation of **CZF** in itself must fail. The culprit turns out to be the fullness axiom, which can therefore not be stable under taking sheaves for the double-negation topology or sheaves for an arbitrary site (see [19] and [21]).

We conclude this introduction by outlining the organisation of our paper. In Section 2 we recall the main definitions from [11, 9]. We will introduce the axioms for a class of small maps necessary to obtain models of **CZF** and **IZF**. Among these necessary axioms, we will discuss the fullness axiom, the axioms concerning W-types and the axiom of multiple choice in detail, as these are the most complicated technically and our main results, which we formulate precisely in Section 2.5, are concerned with these axioms.

In Section 3 we show that predicative categories with small maps are closed under presheaves and that all the axioms that we have listed in Section 2 are inherited by such presheaf models. An important part of our treatment is that we distinguish between two classes of small maps: the "pointwise" and "locally" small ones. It turns out that for certain axioms it is easier to show that they are inherited by pointwise small maps while for other axioms it is easier to show that they are inherited by locally small maps, and therefore it is an important result that these classes of maps coincide.

We follow a similar strategy in Section 4, where we discuss sheaves: we again distinguish between two classes of maps, where for some axioms it is easier to use one definition, while for other axioms it turns out to be easier to use the other. To show that these two classes coincide we use the fullness axiom and assume that the site has a presentation.[4] This section also contains our main technical results: that sheaf models inherit the fullness axiom, as well as the axioms concerning W-types.[5] Strictly speaking our results for presheaves in Section 3 are special cases of our results in Section 4. We believe, however, that it is useful to give direct proofs of the results for presheaves, and in many cases it is helpful to see how the proof goes in the (easier) presheaf case before embarking on the more involved proofs in the sheaf case.

Finally, in Section 5 we give explicit descriptions of the sheaf models of constructive set theory our results lead to. We also point out the connection to forcing for classical set theories.

This will complete our program for developing an abstract semantics of constructive set theory, in particular of Aczel's **CZF**, as outlined [11]. As a result topos-theoretic insights and categorical methods can now be used in the study of constructive set theories. For instance, one can obtain consistency and independence results using sheaf and realizability models or by a combination of these interpretations. In future work, we will use sheaf-theoretic methods to show that the fan rule as well as certain continuity rules are derived rules for **CZF** and related theories [13].

The main results of this paper were presented by the second author in a

[4] In [11] we claimed that (instead of fullness) the exponentiation axiom would suffice to establish this result, but that might not be correct.

[5] One subtlety arises when we try to show that an axiom saying that certain inductives types are small (axiom **(WS)** to be precise) is inherited by sheaf models: we show this using the axiom of multiple choice. In fact, we suspect that something of this sort is unavoidable and one has to go beyond **CZF** proper to show that its validity is inherited by sheaf models.

tutorial on categorical logic at the Logic Colloquium 2006 in Nijmegen. We are grateful to the organisers of the Logic Colloquium for giving one of the authors this opportunity. The final draft of this paper was completed during a stay of the first author at the Mittag-Leffler Institute in Stockholm. We would like to thank the Institute and the organisers of the program in Mathematical Logic in Fall 2009 for awarding him a grant which enabled him to complete this paper in such excellent working conditions. In addition, we would like to acknowledge the helpful discussions we had with Steve Awodey, Nicola Gambino, Jaap van Oosten, Erik Palmgren, Thomas Streicher, Michael Warren, and especially Peter LeFanu Lumsdaine (see Remark 4.14 below).

2 Preliminaries

2.1 Review of Algebraic Set Theory

In this section we recall the main features of our approach to Algebraic Set Theory from [11, 9].

We will always assume that our ambient category $\mathcal{E}$ is a *positive Heyting category*. That means that $\mathcal{E}$ is

(i) Cartesian, i.e., it has finite limits.

(ii) regular, i.e., morphisms factor in a stable fashion as a cover followed by a monomorphism.[6]

(iii) positive, i.e., it has finite sums, which are disjoint and stable.

(iv) Heyting, i.e., for any morphism $f \colon Y \longrightarrow X$ the induced pullback functor $f^* \colon \mathrm{Sub}(X) \longrightarrow \mathrm{Sub}(Y)$ has a right adjoint $\forall_f$.

This means that $\mathcal{E}$ is rich enough to interpret first-order intuitionistic logic. Such a category $\mathcal{E}$ will be called a *category with small maps*, if it comes equipped with a class of maps $\mathcal{S}$ satisfying a list of axioms. To formulate these, we use the notion of a covering square.

Definition 2.1 A diagram in $\mathcal{E}$ of the form

$$\begin{tikzcd} D \arrow[r] \arrow[d, "f"'] & C \arrow[d, "g"] \\ B \arrow[r, "p"'] & A \end{tikzcd}$$

[6] Recall that a map $f \colon B \to A$ is a cover, if the only subobject of A through which it factors, is the maximal one; and that f is a regular epimorphism if it is the coequalizer of its kernel pair. These two classes coincide in regular categories (see [22, Proposition A1.3.4]).

is called a *quasi-pullback*, when the canonical map $D \longrightarrow B \times_A C$ is a cover. If p is also a cover, the diagram will be called a *covering square*. When f and g fit into a covering square as shown, we say that f *covers* g, or that g *is covered by* f.

Definition 2.2 A class of maps in $\mathcal{E}$ satisfying the following axioms **(A1-9)** will be called a *class of small maps*:

(A1) (Pullback stability) In any pullback square

$$\begin{tikzcd} D \arrow[r] \arrow[d, "g"'] & B \arrow[d, "f"] \\ C \arrow[r, "p"'] & A \end{tikzcd}$$

where $f \in \mathcal{S}$, also $g \in \mathcal{S}$.

(A2) (Descent) If in a pullback square as above p is a cover and $g \in \mathcal{S}$, then also $f \in \mathcal{S}$.

(A3) (Sums) Whenever $X \longrightarrow Y$ and $X' \longrightarrow Y'$ belong to $\mathcal{S}$, so does $X + X' \longrightarrow Y + Y'$.

(A4) (Finiteness) The maps $0 \longrightarrow 1$, $1 \longrightarrow 1$ and $1 + 1 \longrightarrow 1$ belong to $\mathcal{S}$.

(A5) (Composition) $\mathcal{S}$ is closed under composition.

(A6) (Quotients) In a commuting triangle

$$\begin{tikzcd} Z \arrow[rr, two heads, "f"] \arrow[dr, "h"'] & & Y \arrow[dl, "g"] \\ & X, & \end{tikzcd}$$

if f is a cover and h belongs to $\mathcal{S}$, then so does g.

(A7) (Collection) Any two arrows $p: Y \longrightarrow X$ and $f: X \longrightarrow A$ where p is a cover and f belongs to $\mathcal{S}$ fit into a covering square

$$\begin{tikzcd} Z \arrow[r] \arrow[d, "g"'] & Y \arrow[r, two heads, "p"] & X \arrow[d, "f"] \\ B \arrow[rr, two heads, "h"'] & & A, \end{tikzcd}$$

where g belongs to $\mathcal{S}$.

(A8) (Heyting) For any morphism $f: Y \longrightarrow X$ belonging to $\mathcal{S}$, the right adjoint to pullback

$$\forall_f : \mathrm{Sub}(Y) \longrightarrow \mathrm{Sub}(X)$$

sends small monos to small monos.

(A9) (Diagonals) All diagonals $\Delta_X : X \longrightarrow X \times X$ belong to $\mathcal{S}$.

For further discussion of these axioms we refer to [9].

A pair $(\mathcal{E}, \mathcal{S})$ in which $\mathcal{S}$ is a class of small maps in $\mathcal{E}$ will be called a *category with small maps*. In such categories with small maps, objects A will be called *small*, if the unique map from A to the terminal object is small. A subobject $A \subseteq X$ will be called a *small subobject* if A is a small object. If any of its representing monomorphisms $m: A \to X$ is small, they all are and in this case the subobject will be called *bounded*.

Remark 2.3 In the sequel we will often implicitly use that categories with small maps are stable under slicing. By this we mean that for any category with small maps $(\mathcal{E}, \mathcal{S})$ and object X in $\mathcal{E}$, the pair $(\mathcal{E}/X, \mathcal{S}/X)$, with $\mathcal{S}/X$ being defined by

$$f \in \mathcal{S}/X \Leftrightarrow \Sigma_X f \in \mathcal{S},$$

is again a category with small maps (here Σ_X is the forgetful functor $\mathcal{E}/X \to \mathcal{E}$ sending an object $p: A \to X$ in $\mathcal{E}/X$ to A and morphisms to themselves). Moreover, any of the further axioms for classes of small maps to be introduced below are stable under slicing, in the sense that their validity in the slice over 1 implies their validity in every slice.

Remark 2.4 A very useful feature of categories of small maps, and one we will frequently exploit, is that they satisfy an internal form of bounded separation. A precise statement is the following: if $\phi(x)$ is a formula in the internal logic of $\mathcal{E}$ with free variable $x \in X$, all whose basic predicates are interpreted as bounded subobjects (note that this includes all equalities, by **(A9)**), and which contains existential and universal quantifications $\exists_f$ and $\forall_f$ along small maps f only, then

$$A = \{x \in X : \phi(x)\} \subseteq X$$

defines a bounded subobject of X. In particular, smallness of X implies smallness of A.

Definition 2.5 A category with small maps $(\mathcal{E}, \mathcal{S})$ will be called a *predicative category with small maps*, if the following axioms hold:

(ΠE) All morphisms $f \in \mathcal{S}$ are exponentiable.

(WE) For all $f: X \longrightarrow Y \in \mathcal{S}$, the W-type W_f associated to f exists.

(NE) $\mathcal{E}$ has a natural numbers object $\mathbb{N}$.

(NS) Moreover, $\mathbb{N} \longrightarrow 1 \in \mathcal{S}$.

(Representability) There is a small map $\pi: E \longrightarrow U$ (the "universal small map") such that any $f: Y \longrightarrow X \in \mathcal{S}$ fits into a diagram of the form

$$\begin{tikzcd} Y \arrow[d, "f"'] & B \arrow[l, two heads] \arrow[r] \arrow[d] & E \arrow[d, "\pi"] \\ X & A \arrow[l, two heads] \arrow[r] & U, \end{tikzcd}$$

where the left hand square is covering and the right hand square is a pullback.

(Bounded exactness) For any equivalence relation

$$R \rightarrowtail X \times X$$

given by a small mono, a stable quotient X/R exists in $\mathcal{E}$.

(For a detailed discussion of these axioms we refer again to [9]; W-types and the axiom **(WE)** will also be discussed in Section 2.3 below.)

In predicative categories with small maps one can derive the existence of a power class functor, classifying small subobjects:

Definition 2.6 By a *D-indexed family of subobjects* of C, we mean a subobject $R \subseteq C \times D$. It will be called a *D-indexed family of small subobjects*, whenever the composite

$$R \subseteq C \times D \longrightarrow D$$

belongs to $\mathcal{S}$. If it exists, the *power class object* $\mathcal{P}_s X$ is the classifying object for the families of small subobjects of X. This means that it comes equipped with a $\mathcal{P}_s X$-indexed family of small subobjects of X, denoted by $\in_X \subseteq X \times \mathcal{P}_s X$ (or simply $\in$, whenever X is understood), with the property that for any Y-indexed family of small subobjects of X, $R \subseteq X \times Y$ say, there exists a unique map $\rho: Y \longrightarrow \mathcal{P}_s X$ such that the square

$$\begin{tikzcd} R \arrow[r] \arrow[d, tail] & \in_X \arrow[d, tail] \\ X \times Y \arrow[r, "\mathsf{id} \times \rho"'] & X \times \mathcal{P}_s X \end{tikzcd}$$

is a pullback.

Proposition 2.7 [9, Corollary 6.11] *In a predicative category with small maps all power class objects exist.*

Moreover, one can show that the assignment $X \mapsto \mathcal{P}_s X$ is functorial and that this functor has an initial algebra.

Theorem 2.8 [9, Theorem 7.4] *In a predicative category with small maps the $\mathcal{P}_s$-functor has an initial algebra.*

The importance of this result resides in the fact that this initial algebra can be used to model a weak intuitionistic set theory: if V is the initial algebra and $E: V \to \mathcal{P}_s V$ is the inverse of the $\mathcal{P}_s$-algebra map on V (which is an isomorphism, since V is an initial algebra), then one can define a binary predicate ϵ on V by setting

$$x \epsilon y \Leftrightarrow x \in_V E(y),$$

where $\in_V \subseteq V \times \mathcal{P}_s V$ derives from the power class structure on $\mathcal{P}_s V$. The resulting structure (V, ϵ) models a weak intuitionistic set theory, which we have called **RST** (for rudimentary set theory), consisting of the following axioms:

Extensionality: $\forall x \,(\, x \epsilon a \leftrightarrow x \epsilon b \,) \to a = b$.

Empty set: $\exists x \, \forall y \, \neg y \epsilon x$.

Pairing: $\exists x \, \forall y \,(\, y \epsilon x \leftrightarrow y = a \vee y = b \,)$.

Union: $\exists x \, \forall y \,(\, y \epsilon x \leftrightarrow \exists z \epsilon a \, y \epsilon z \,)$.

Set induction: $\forall x \,(\forall y \epsilon x \, \phi(y) \to \phi(x)) \to \forall x \, \phi(x)$.

Bounded separation: $\exists x \, \forall y \,(\, y \epsilon x \leftrightarrow y \epsilon a \wedge \phi(y) \,)$, for any bounded formula ϕ in which a does not occur.

Strong collection: $\forall x \epsilon a \, \exists y \, \phi(x, y) \to \exists b \, \mathrm{B}(x \epsilon a, y \epsilon b) \, \phi$, where $\mathrm{B}(x \epsilon a, y \epsilon b) \, \phi$ abbreviates

$$\forall x \epsilon a \, \exists y \epsilon b \, \phi \wedge \forall y \epsilon b \, \exists x \epsilon a \, \phi.$$

Infinity: $\exists a \,(\, \exists x \, x \epsilon a \,) \wedge (\, \forall x \epsilon a \, \exists y \epsilon a \, x \epsilon y \,)$.

In fact, as shown in [9], the initial $\mathcal{P}_s$-algebras in predicative categories with small maps form a complete semantics for the set theory **RST**. To obtain complete semantics for better known intuitionistic set theories, like **IZF** and **CZF**, one needs further requirements on the class of small maps $\mathcal{S}$. For example, the set theory **IZF** is obtained from **RST** by adding the axioms

Full separation: $\exists x \, \forall y \,(\, y \epsilon x \leftrightarrow y \epsilon a \wedge \phi(y) \,)$, for any formula ϕ in which a does not occur.

Power set: $\exists x \, \forall y \,(\, y \epsilon x \leftrightarrow y \subseteq a \,)$, where $y \subseteq a$ abbreviates $\forall z \,(z \epsilon y \to z \epsilon a)$.

And to obtain a sound and complete semantics for **IZF** one requires of ones predicative category of small maps that it satisfies:

(M) All monomorphisms belong to $\mathcal{S}$.

(PS) For any map $f: Y \longrightarrow X \in \mathcal{S}$, the power class object $\mathcal{P}_s^X(f) \longrightarrow X$ in $\mathcal{E}/X$ belongs to $\mathcal{S}$.

The set theory **CZF**, introduced by Aczel in [1], is obtained by adding to **RST** a weakening of the power set axiom called subset collection:

Subset collection: $\exists c \, \forall z \, (\forall x \epsilon a \, \exists y \epsilon b \, \phi(x, y, z) \rightarrow \exists d \epsilon c \, \mathrm{B}(x \epsilon a, y \epsilon d) \, \phi(x, y, z))$.

For a suitable categorical analogue, see Section 2.3 below.

For the sake of completeness we also list the following two axioms, saying that certain Π-types and W-types are small. (The first therefore corresponds to the exponentiation axiom in set theory; we will say more about the second in Section 2.2 below.)

(ΠS) For any map $f: Y \longrightarrow X \in \mathcal{S}$, a functor

$$\Pi_f : \mathcal{E}/Y \longrightarrow \mathcal{E}/X$$

right adjoint to pullback exists and preserves morphisms in $\mathcal{S}$.

(WS) For all $f: X \longrightarrow Y \in \mathcal{S}$ with Y small, the W-type W_f associated to f is small.

2.2 W-types

In a predicative category with small maps $(\mathcal{E}, \mathcal{S})$ the axiom **(ΠE)** holds and therefore any small map $f: B \longrightarrow A$ is exponentiable. It therefore induces an endofunctor on $\mathcal{E}$, which will be called the *polynomial functor* P_f associated to f. The quickest way to define it is as the following composition:

$$\mathcal{C} \cong \mathcal{C}/1 \xrightarrow{B^*} \mathcal{C}/B \xrightarrow{\Pi_f} \mathcal{C}/A \xrightarrow{\Sigma_A} \mathcal{C}/1 \cong \mathcal{C}.$$

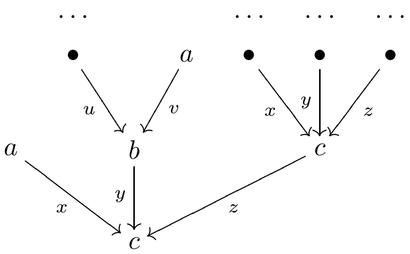
In more set-theoretic terms it could be defined as:

$$P_f(X) = \sum_{a \in A} X^{B_a}.$$

Whenever it exists, the initial algebra for the polynomial functor P_f will be called the *W-type associated to* f.

Intuitively, elements of a W-type are well-founded trees. In the category of sets, all W-types exist, and the W-types have as elements well-founded trees, with an appropriate labelling of its edges and nodes. What is an appropriate labelling is determined by the branching type $f\colon B \longrightarrow A$: nodes should be labelled by elements $a \in A$, edges by elements $b \in B$, in such a way that the edges into a node labelled by a are uniquely enumerated by $f^{-1}(a)$. The following picture hopefully conveys the idea:

$$\begin{array}{rcl} f^{-1}(a) & = & \emptyset \\ f^{-1}(b) & = & \{u, v\} \\ f^{-1}(c) & = & \{x, y, z\} \\ & \cdots & \end{array}$$



This set has the structure of a P_f-algebra: when an element $a \in A$ is given, together with a map $t\colon B_a \longrightarrow W_f$, one can build a new element $\sup_a t \in W_f$, as follows. First take a fresh node, label it by a and draw edges into this node, one for every $b \in B_a$, labelling them accordingly. Then on the edge labelled by $b \in B_a$, stick the tree tb. Clearly, this sup operation is a bijective map. Moreover, since every tree in the W-type is well-founded, it can be thought of as having been generated by a possibly transfinite number of iterations of this sup operation. That is precisely what makes this algebra initial. The trees that can be thought of as having been used in the generation of a certain element $w \in W_f$ are called its subtrees. One could call the trees $tb \in W_f$ the *immediate subtrees* of $\sup_a t$, and $w' \in W_f$ a *subtree* of $w \in W_f$ if it is an immediate subtree, or an immediate subtree of an immediate subtree, or..., etc. Note that with this use of the word subtree, a tree is never a subtree of itself (so proper subtree might have been a better terminology).

We recall that there are two axioms concerning W-types:

(WE) For all $f\colon X \longrightarrow Y \in \mathcal{S}$, the W-type W_f associated to f exists.

(WS) Moreover, if Y is small, also W_f is small.

Maybe it is not too late to point out the following fact, which explains why these axioms play no essential role in the impredicative setting:

Theorem 2.9 *Let* $(\mathcal{E}, \mathcal{S})$ *be a category with small maps satisfying* **(NS)** *and* **(M)**.

1. *If* $\mathcal{S}$ *satisfies* **(PE)**, *then it also satisfies* **(WE)**.
2. *If* $\mathcal{S}$ *satisfies* **(PS)**, *then it also satisfies* **(WS)**.

Proof. Note that in a category with small maps satisfying **(M)** and **(PE)** the object $\mathcal{P}_s(1)$ is a subobject classifier. Therefore the first result can be shown along the lines of Chapter 3 in [24]. For showing the second result, one simply copies the argument why toposes with nno have all W-types from [29]. $\square$

In the sequel we will need the following result. We will write $\mathcal{P}_s^+ X$ for the object of small inhabited subobjects of X:

$$\mathcal{P}_s^+ X = \{A \in \mathcal{P}_s X : \exists x \in X\, (x \in A)\}.$$

Theorem 2.10 *For any small map $f: B \to A$ in a predicative category with small maps $(\mathcal{E}, \mathcal{S})$, the endofunctors on $\mathcal{E}$ defined by*

$$\Phi = P_f \circ \mathcal{P}_s \qquad \textit{and} \qquad \Psi = P_f \circ \mathcal{P}_s^+$$

have initial algebras.

Remark 2.11 Before we sketch the proof of Theorem 2.10, it might be good to explain the intuitive meaning of these initial algebras. In fact, they are variations on the W-types explained above: they are also classes of well-founded trees, but the conditions on the labellings of the nodes and edges are slightly different. It is still the case that nodes are labelled by elements $a \in A$ and edges with elements $b \in B$, in such a way that if $b \in B$ decorates a certain edge, then $f(b)$ decorates the node it points to. But whereas in a W-type, every node in a well-founded tree labelled with $a \in A$ has for every $b \in f^{-1}(a)$ *precisely one* edge into it labelled with b, in the initial algebras for Φ there are *set-many, and possibly none*, and in the initial algebra for Ψ there are *set-many, but at least one*.

Proof. The proof of Theorem 2.10 is a variation on that of Theorem 7.4 in [9] and therefore we will only sketch the argument.

Fix a universal small map $\pi: E \to U$, and write

$$S = \{(a \in A, u \in U, \phi: E_u \to B_a)\}.$$

Let $\mathcal{K}$ be the W-type in $\mathcal{E}$ associated to the map g fitting into the pullback square

$$\begin{tikzcd} R \arrow[r] \arrow[d, "g"'] & E \arrow[d, "\pi"] \\ S \arrow[r, "\mathrm{proj}"'] & U. \end{tikzcd}$$

An element $k \in \mathcal{K}$ is therefore of the form $\sup_{(a,u,\phi)} t$, where $(a, u, \phi) \in S$ is the label of the root of k and $t: E_u \to \mathcal{K}$ is the function that assigns to every element $e \in E_u$ the tree that is attached to the root of k with the edge labelled with e. Define the following equivalence relation on $\mathcal{K}$ by recursion: $\sup_{(a,u,\phi)} t \sim \sup_{(a',u',\phi')} t'$, if $a = a'$ and for all $e \in E_u$ there is an $e' \in E_{u'}$ such that $\phi(e) = \phi'(e')$ and $t(e) \sim t'(e')$, and for all $e' \in E_{u'}$ there is an $e \in E_u$ such that $\phi(e) = \phi'(e')$ and $t(e) \sim t'(e')$.

(The existence of this relation $\sim$ can be justified using the methods of [8] or [9]. See Theorem 7.4 in [9], for instance.) The equivalence relation is bounded (one proves this by induction) and its quotient is the initial algebra for Φ.

The initial algebra for Ψ is constructed in the same way, but with S defined as

$$S = \{(a \in A, u \in U, \phi\colon E_u \to B_a) : \phi \text{ is a cover}\}.$$

$\square$

2.3 Fullness

In order to express the subset collection axiom, introduced by Peter Aczel in [1], in diagrammatic terms, it is helpful to consider an axiom which is equivalent to it called *fullness* (see [4]). In the language of set theory one can formulate fullness using the notion of a *multi-valued section*: a multi-valued section (or *mvs*) of a function $\phi\colon b \longrightarrow a$ is a multi-valued function s from a to b such that $\phi s = \mathsf{id}_a$ (as relations). Identifying s with its image, this is the same as a subset p of b such that $p \subseteq b \longrightarrow a$ is surjective. For us, fullness states that for any such ϕ there is a small family of *mvs*s such that any *mvs* contains one in this family. Written out formally:

Fullness: $\exists z\, (z \subseteq \mathbf{mvs}(\phi) \wedge \forall x \,\epsilon\, \mathbf{mvs}(\phi)\, \exists c \,\epsilon\, z\, (c \subseteq x))$.

Here, $\mathbf{mvs}(\phi)$ is an abbreviation for the class of all multi-valued sections of a function $\phi\colon b \longrightarrow a$, i.e., subsets p of b such that $\forall x \epsilon a\, \exists y \epsilon p\, \phi(y) = x$.

In order to reformulate this diagrammatically, we say that a multi-valued section (*mvs*) for a small map $\phi\colon B \longrightarrow A$, over some object X, is a subobject $P \subseteq B$ such that the composite $P \longrightarrow A$ is a *small* cover. (Smallness of this map is equivalent to P being a bounded subobject of B.) We write

$$\mathrm{mvs}_X(\phi)$$

for the set of all *mvs*s of a map ϕ. This set obviously inherits the structure of a partial order from $\mathrm{Sub}(B)$. Note that any morphism $f\colon Y \longrightarrow X$ induces an order-preserving map

$$f^*\colon \mathrm{mvs}_X(\phi) \longrightarrow \mathrm{mvs}_Y(f^*\phi),$$

obtained by pulling back along f. To avoid overburdening the notation, we will frequently talk about the map ϕ over Y, when we actually mean the map $f^*\phi$ over Y, the map f always being understood.

The categorical fullness axiom now reads:

(F) For any $\phi: B \longrightarrow A \in \mathcal{S}$ over some X with $A \longrightarrow X \in \mathcal{S}$, there is a cover $q: X' \longrightarrow X$ and a map $y: Y \longrightarrow X'$ belonging to $\mathcal{S}$, together with an *mvs* P of ϕ over Y, with the following "generic" property: if $z: Z \longrightarrow X'$ is any map and Q any *mvs* of ϕ over Z, then there is a map $k: U \longrightarrow Y$ and a cover $l: U \longrightarrow Z$ with $yk = zl$ such that $k^*P \leq l^*Q$ as *mvs*s of ϕ over U.

It is easy to see that in a set-theoretic context fullness is a consequence of the powerset axiom (because then the collection of *all* multi-valued sections of a map $\phi: b \to a$ forms a set) and implies the exponentiation axiom (because if z is a set of *mvs*s of the projection $p: a \times b \to a$ such that any *mvs* is refined by one is this set, then the set of functions from a to b can be constructed from z by selecting the univalued elements, i.e., those elements that are really functions). Showing that in a categorical context **(F)** follows from **(PS)** and implies **(ΠS)** is not much harder and we will therefore not write out a formal proof.

In the sequel we will use the following two lemmas concerning the fullness axiom:

Lemma 2.12 *Suppose we have the following diagram*

$$\begin{tikzcd}
Y_2 \arrow[r, "\beta"] \arrow[d, "f_2"'] & Y_1 \arrow[d, "f_1"] \\
X_2 \arrow[r] \arrow[dr, two heads] & X_1 \arrow[d, two heads] \\
& X,
\end{tikzcd}$$

in which the square is a quasi-pullback and f_1 and f_2 are small. When P is a "generic" mvs *for a map $\phi: B \longrightarrow A$ over X living over Y_1 ("generic" as in the statement of the fullness axiom), then β^*P is also a generic* mvs *for ϕ, living over Y_2.*

Proof. A simple diagram chase. □

Lemma 2.13 *Suppose we are given a diagram of the form*

$$\begin{tikzcd}
B_0 \arrow[r, two heads] \arrow[d, "\psi"'] & B \arrow[d, "\phi"] \\
A_0 \arrow[r, two heads] \arrow[d, "i"'] & A \arrow[d, "j"] \\
X_0 \arrow[r, two heads, "p"'] & X,
\end{tikzcd}$$

in which both squares are covering and all the vertical arrows are small. If a generic mvs *for ψ exists over X_0, then also a generic* mvs *for ϕ exists over X.*

Proof. This was Lemma 6.23 in [9]. □

2.4 Axiom of multiple choice

The axiom of multiple choice was introduced by Moerdijk and Palmgren in [29]. Their motivation was to have a choice principle which is implied by the existence of enough projectives ("the presentation axiom" in Aczel's terminology) and is stable under taking sheaves (unlike the existence of enough projectives). We will use it in Section 4.4 to show that the axiom **(WS)** is stable under taking sheaves.

One can give a succinct formulation of the axiom of multiple choice using the notion of a collection span (see [9, Definition 6.14]).[7]

Definition 2.14 A span (g, h) in $\mathcal{E}$

$$C \xleftarrow{g} D \xrightarrow{h} B$$

is called a *collection span*, when, in the internal logic, it holds that for any map $f: E \longrightarrow D_c$ covering some fibre of g, there is a fibre $D_{c'}$ of g and a map $p: D_{c'} \longrightarrow E$ such that fp is a cover over B. A collection span is $\mathcal{E}/A$ will be called a *collection span over A*.

Diagrammatically, we can express this by asking that for any map $e: E \longrightarrow C$ and any epi $F \longrightarrow E \times_C D$ there is a diagram of the form

$$\begin{tikzcd}
 & & B & & \\
D \arrow[urr, bend left, "h"] \arrow[d, "g"'] & E' \times_C D \arrow[l] \arrow[r] \arrow[d] & F \arrow[r, two heads] & E \times_C D \arrow[r] \arrow[d] & D \arrow[ull, bend right, "h"'] \arrow[d, "g"] \\
C & E' \arrow[l] \arrow[rr, two heads] & & E \arrow[r, "c"'] & C
\end{tikzcd}$$

where the middle square is a covering square, involving the given map $F \longrightarrow D \times_C E$, while the other two squares are pullbacks.

(AMC) (Axiom of multiple choice) For any small map $f: Y \longrightarrow X$, there is a

[7] The way we formulate the Axiom of Multiple Choice here is slightly different from how it was stated in [29]. Both formulations are equivalent, however; see [14].

cover $q\colon A \longrightarrow X$ and a diagram

$$\begin{tikzcd} D \arrow[r, two heads, "h"] \arrow[d, "g"'] & q * Y \arrow[r, two heads] \arrow[d, "q^* f"'] & Y \arrow[d, "f"] \\ C \arrow[r, two heads, "r"'] & A \arrow[r, two heads, "q"'] & X, \end{tikzcd}$$

in which the right square is a pullback and the left square a covering square in which all maps are small and in which (g, h) is a collection span over A.

In the internal logic **(AMC)** is often applied in the following form:

Lemma 2.15 *In a predicative category with small maps in which* **(AMC)** *holds, the following principle holds in the internal logic: any small map $f\colon B \longrightarrow A$ between small objects fits into a covering square*

$$\begin{tikzcd} D \arrow[r, "q"] \arrow[d, "g"'] & B \arrow[d, "f"] \\ C \arrow[r, "p"'] & A \end{tikzcd}$$

in which all maps and objects are small and (g, q) is a collection span over A.

Proof. This is proved exactly as Proposition 4.6 in [29]. □

The following result was proved in [29] as well. Recall from [2, 4] that the existence of many inductively defined sets within **CZF** can be guaranteed, in a predicatively acceptable way, by extending **CZF** with Aczel's Regular Extension Axiom.

Proposition 2.16 *If $(\mathcal{E}, \mathcal{S})$ is a predicative category with small maps satisfying the axioms* **(AMC)***,* **(ΠS)** *and* **(WS)***, then Aczel's Regular Extension Axiom holds in the initial $\mathcal{P}_s$-algebra in this category.*

In addition, we will need:

Proposition 2.17 *Let $(\mathcal{E}, \mathcal{S})$ be a predicative category with small maps. If $\mathcal{S}$ satisfies the axioms* **(AMC)** *and* **(ΠS)***, then it satisfies the axiom* **(F)** *as well.*

Proof. We argue internally and use Lemma 2.15. So suppose that **(AMC)** holds and $f\colon B \to A$ is a small map between small objects. We need to find a small collection of *mvss* $\{P_y : y \in Y\}$ such that any *mvs* of f is refined by one in this family.

We apply Lemma 2.15 to $A \to 1$ to obtain a covering square of the form

$$\begin{tikzcd} D \arrow[r, "h"] \arrow[d] & A \arrow[d] \\ C \arrow[r] & 1, \end{tikzcd}$$

such that for any cover $p\colon E \to D_c$ we find a $c' \in C$ and a map $t\colon D_{c'} \to E$ such that pt is a cover over A. Let Y be the collection of all pairs (c, s) with c in C and s a map $D_c \to B$ such that $fs = h_c$, and let P_y be the image of the map $s\colon D_c \to B$. Then P_y is an *mvs*, because the h_c are epi, and Y is small, because **(ΠS)** holds.

Now suppose $n\colon Q \to B$ is any mono such that $fn\colon Q \to B \to A$ is a cover. Pick a $c \in C$ and pull back fn along h_c to obtain a cover $q\colon E \to D_c$, as in:

$$\begin{tikzcd} E \arrow[r, "p"] \arrow[d, two heads, "q"'] & Q \arrow[d, two heads, "fn"] \\ D_c \arrow[r, "h_c"'] & A. \end{tikzcd}$$

It follows that there exists an element $c' \in C$ and a map $g\colon D_{c'} \to E$ such that qg is a cover over A. Set $s = npg$ and $y = (c', s)$. Then $P_y = \mathrm{Im}(g)$ is contained in Q. □

2.5 Main results

After all these definitions, we can formulate our main result. Let $\mathcal{A}$ be either $\{(\mathbf{F})\}$, or $\{(\mathbf{AMC}), (\mathbf{\Pi S}), (\mathbf{WS})\}$, or $\{(\mathbf{M}), (\mathbf{PS})\}$.

Theorem 2.18 *Let $(\mathcal{E}, \mathcal{S})$ be a predicative category with small maps for which all the axioms in $\mathcal{A}$ hold and let $(\mathcal{C}, \mathrm{Cov})$ be an internal Grothendieck site in $\mathcal{E}$, such that the codomain map $\mathcal{C}_1 \to \mathcal{C}_0$ is small and a presentation for the topology exists. Then in the category of internal sheaves $\mathrm{Sh}_{\mathcal{E}}(\mathcal{C})$ one can identify a class of maps making it into a predicative category with small maps for which the axioms in $\mathcal{A}$ holds as well.*

In combination with Theorem 2.8 this result can be used to prove the existence of sheaf models of various constructive set theories:

Corollary 2.19 *Suppose that $(\mathcal{E}, \mathcal{S})$ is a predicative category with small maps satisfying the axiom* $(\mathbf{F})$ *and suppose that $(\mathcal{C}, \mathrm{Cov})$ is an internal Grothendieck site in $\mathcal{E}$, such that the codomain map $\mathcal{C}_1 \to \mathcal{C}_0$ is small and a presentation for the topology exists. Then the initial $\mathcal{P}_s$-algebra in $\mathrm{Sh}_{\mathcal{E}}(\mathcal{C})$ exists and is a model of* **CZF**. *If, moreover,*

1. *the axioms* (**AMC**) *and* (**WS**) *hold in* $\mathcal{E}$*, then the initial* $\mathcal{P}_s$*-algebra in* $\mathrm{Sh}_{\mathcal{E}}(\mathcal{C})$ *also models Aczel's Regular Extension Axiom.*
2. *the axioms* (**M**) *and* (**PS**) *hold in* $\mathcal{E}$*, then the initial* $\mathcal{P}_s$*-algebra in* $\mathrm{Sh}_{\mathcal{E}}(\mathcal{C})$ *is a model of* **IZF**.

3 Presheaves

In this section we show that predicative categories with small maps are closed under presheaves. More precisely, we show that if $(\mathcal{E}, \mathcal{S})$ is a predicative category with small maps and $\mathcal{C}$ is an internal category in $\mathcal{E}$, then inside the category $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ of internal presheaves one can identify a class of maps such that $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ becomes a predicative category with small maps. Our argument proceeds in two steps. First, we need to identify a suitable class of maps in a category of internal presheaves. We take what we will call the pointwise small maps of presheaves. To prove that these pointwise small maps satisfy axioms **(A1-9)**, we need to assume that the codomain map of $\mathcal{C}$ is small (note that the same assumption was made in [34]). Subsequently, we show that the validity in the category with small maps $(\mathcal{E}, \mathcal{S})$ of any of the axioms introduced in the previous section implies its validity in any category of internal presheaves over $(\mathcal{E}, \mathcal{S})$. To avoid repeating the convoluted expression "the validity of axiom **(X)** in a predicative category with small maps implies its validity in any category of internal presheaves over it", we will write "**(X)** is inherited by presheaf models" or "**(X)** is stable under presheaf extensions" to express this.

The main result of this section is that the fullness axiom **(F)** is stable under presheaf extensions. Most of the other stability results in this section are not really new and can in one form or another already be found in [24, 28, 29, 18, 34]. Nevertheless, for several reasons, we have decided to include their proofs here. First of all, none of the references we mentioned uses conditions on the ambient category which are exactly the same as ours (in particular, we assume only bounded exactness). Secondly, these papers use different definitions of the class of small maps in presheaves, which we will compare in Section 3.2 below. And, thirdly, including them will make our presentation self-contained.

3.1 Pointwise small maps in presheaves

Throughout this section, we work in a predicative category with small maps $(\mathcal{E}, \mathcal{S})$ in which we are given an internal category $\mathcal{C}$, whose codomain map

$$\mathrm{cod}\colon \mathcal{C}_1 \longrightarrow \mathcal{C}_0$$

is small. Here we have written $\mathcal{C}_0$ for the object of objects of $\mathcal{C}$ and $\mathcal{C}_1$ for its object of arrows. In addition, we will write $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ for the category of internal presheaves, and π^* for the forgetful functor:

$$\pi^*: \mathrm{Psh}_{\mathcal{E}}(\mathcal{C}) \longrightarrow \mathcal{E}/\mathcal{C}_0.$$

In the sequel, we will use capital letters for presheaves and morphisms of presheaves, and lower case letters for objects and morphisms in $\mathcal{C}$.

We will also employ the following piece of notation. For any map of presheaves $F: Y \to X$ and element $x \in X(a)$, we set

$$Y_x^M := \{ (f \in \mathcal{C}_1, y \in Y(\mathrm{dom} f)) : \mathrm{cod}(f) = a, F(y) = x \cdot f \}.$$

(The capital letter M stands for the maximal sieve on b: for this reason, this piece of notation is consistent with the one to be introduced in Section 4.4.) Occasionally, we will regard Y_x^M as a presheaf: in that case, its fibre at $b \in \mathcal{C}_0$ is

$$Y_x^M(b) = \{ (f: b \to a \in \mathcal{C}_1, y \in Y(b)) : F_b(y) = x \cdot f \},$$

and the restriction of an element $(f, y) \in Y_x^M(b)$ along $g: c \to b$ is given by

$$(f, y) \cdot g = (fg, y \cdot g).$$

A map of presheaves $F: Y \to X$ will be called *pointwise small*, if $\pi^* F$ belongs to $\mathcal{S}/\mathcal{C}_0$ in $\mathcal{E}/\mathcal{C}_0$. Note that for any such pointwise small map of presheaves and for any $x \in X(a)$ with $a \in \mathcal{C}_0$ the object Y_x^M will be small. This is an immediate consequence of the fact that the codomain map is assumed to be small.

Theorem 3.1 *The pointwise small maps make* $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ *into a category with small maps.*

Proof. Observe that finite limits, images and sums of presheaves are computed "pointwise", that is, as in $\mathcal{E}/\mathcal{C}_0$. The universal quantification of $A \subseteq Y$ along $F: Y \longrightarrow X$ is given by the following formula: for any $a \in \mathcal{C}_0$,

$$\forall_F(A)(a) \quad = \quad \{ x \in X(a) : \forall (f, y) \in Y_x^M \, (y \in A) \} \tag{1}$$

This shows that $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ is a positive Heyting category. To complete the proof, we need to check that the pointwise small maps in presheaves satisfy axioms **(A1-9)**. We postpone the proof of the collection axiom **(A7)** (it will be Proposition 3.9). The remaining axioms follow easily, as all we need to do is verify them pointwise. For verifying axiom **(A8)**, one observes that the universal quantifier in (1) ranges over a small object. $\square$

For most of the axioms that we introduced in Section 2, it is relatively straightforward to check that they are inherited by presheaf models. The exceptions are the representability, collection and fullness axioms: verifying these requires an alternative characterisation of the small maps in presheaves and they will therefore be discussed in a separate section.

Proposition 3.2 *The following axioms are inherited by presheaf models:* **(M)**, *bounded exactness,* **(NE)** *and* **(NS)**, *as well as* $(\Pi\mathbf{E})$, $(\Pi\mathbf{S})$ *and* **(PS)**.

Proof. The monomorphisms in presheaves are precisely those maps which are pointwise monic and therefore the axiom **(M)** will be inherited by presheaf models. Similarly, presheaf models inherit bounded exactness, because quotients of equivalence relations are computed pointwise. Since the natural numbers objects in presheaves has that of the base category $\mathcal{E}$ in every fibre, both **(NE)** and **(NS)** are inherited by presheaf models.

Finally, consider the following diagram in presheaves, in which F is small:

$$\begin{tikzcd} B \arrow[d, "G"'] & \\ Y \arrow[r, "F"'] & X. \end{tikzcd}$$

The object $P = \Pi_F(G)$ over an element $x \in X(a)$ is given by the formula:

$$P_x := \{\, s \in \Pi_{(f,y)\in Y_x^M} G^{-1}(y) : s \text{ is natural}\,\}.$$

This shows that $(\Pi\mathbf{E})$ is inherited by presheaf extensions. It also shows that $(\Pi\mathbf{S})$ is inherited, because the formula

$$\forall (f,y) \in Y_x^M(b)\, \forall g{:}\, c \to b\, (s(f,y)\cdot g = s(fg, y\cdot g))$$

expressing the naturality of s is bounded.

To see that **(PS)** is inherited, we first need a description of the $\mathcal{P}_s$-functor in the category of internal presheaves. This was first given by Gambino in [18] and works as follows. If X is a presheaf and $\mathbf{y}c$ is the representable presheaf on $c \in \mathcal{C}_0$, then

$$\mathcal{P}_s(X)(c) = \{\, A \subseteq \mathbf{y}c \times X : A \text{ is a small subpresheaf}\,\},$$

with restriction along $f{:}\, d \to c$ on an element $A \in \mathcal{P}_s(X)(c)$ defined by

$$(A \cdot f)(e) = \{(g{:}\, e \to d, x \in X(e)) : (fg, x) \in A\}.$$

The membership relation $\in_X \subseteq X \times \mathcal{P}_s X$ is defined on an object $c \in \mathcal{C}$ by: for all $x \in X(c)$ and $A \in \mathcal{P}_s(X)(c)$,

$$x \in_X A \Longleftrightarrow (\mathsf{id}_c, x) \in A.$$

This shows that the axiom **(PS)** is inherited, because the formula

$$\forall (f{:}\, b \to c, x) \in A\, \forall g{:}\, a \to b\, [\, (fg, x\cdot g) \in A\,]$$

expressing that A is a subpresheaf is bounded. $\square$

Theorem 3.3 *The axioms* **(WE)** *and* **(WS)** *are inherited by presheaf extensions.*

Proof. For this proof we need to recall the construction of polynomial functors and W-types in presheaves from [28]. For a morphism of presheaves $F: Y \to X$ and a presheaf Z, the value of

$$P_F(Z) = \sum_{x \in X} Z^{Y_x}$$

on an object a of $\mathcal{C}_0$ is given by

$$P_F(Z)(a) = \{ (x \in X(a), t: Y_x^M \to Z)\},$$

where t is supposed to be a morphism of presheaves. The restriction of an element (x, t) along a map $f: b \to a$ is given by $(x \cdot f, f^*t)$, where

$$(f^*t)(g, y) = t(fg, y).$$

The presheaf morphism F induces a map

$$\phi: \sum_{a \in \mathcal{C}_0} \sum_{x \in X(a)} Y_x^M \longrightarrow \sum_{a \in \mathcal{C}_0} X(a)$$

in $\mathcal{E}$ whose fibre over $x \in X(a)$ is Y_x^M and which is therefore small. The W-type in presheaves will be constructed from the W-type V associated to ϕ in $\mathcal{E}$.

A typical element $v \in V$ is a tree of the form

$$v = \sup_x t$$

where x is an element of some $X(a)$ and t is a function $Y_x^M \to V$. For any such v, one defines its root $\rho(v)$ to be a. If one writes $V(a)$ for the set of trees v such that $\rho(v) = a$, the object V will carry the structure of a presheaf, with the restriction of an element $v \in V(a)$ along a map $f: b \longrightarrow a$ given by

$$v \cdot f = \sup_{x \cdot f} f^*t.$$

The W-type associated to F in presheaves is obtained by selecting the right trees from V, the right trees being those all whose subtrees are (in the terminology of [28]) composable and natural. A tree $v = \sup_x(t)$ is called *composable* if for all $(f, y) \in Y_x^M$,

$$\rho(t(f, y)) = \mathrm{dom}(f).$$

A tree $v = \sup_x(t)$ is *natural*, if it is composable and for any $(f, y) \in Y_x^M(a)$ and any $g: b \to a$, we have

$$t(f, y) \cdot g = t(fg, y \cdot g)$$

(so t is actually a natural transformation). A tree will be called *hereditarily natural*, if all its subtrees (including the tree itself) are natural.

In [28, Lemma 5.5] it was shown that for any hereditarily natural tree v rooted in a and map $f: b \to a$ in $\mathcal{C}$, the tree $v \cdot f$ is also hereditarily natural. So when $W(a) \subseteq V(a)$ is the collection of hereditarily natural trees rooted in a, W is a subpresheaf of V.

A proof that W is the W-type for F can be found in the sources mentioned above. Presently, the crucial point is that the construction can be imitated in our setting, so that **(WE)** is stable under presheaves. The same applies to **(WS)**, essentially because W was obtained from V using bounded separation (in this connection it is essential that the object of all subtrees of a particular tree v is small, see [9, Theorem 6.13]). □

3.2 Locally small maps in presheaves

For showing that the representability, collection and fullness axioms are inherited by presheaf models, we use a different characterisation of the small maps in presheaves: we introduce the *locally small maps* and show that these coincide with the pointwise small maps. To define these locally small maps, we have to set up some notation.

Remark 3.4 The functor $\pi^*: \mathrm{Psh}_{\mathcal{E}}(\mathcal{C}) \longrightarrow \mathcal{E}/\mathcal{C}_0$ has a left adjoint, which is computed as follows: to any object $(X, \sigma_X: X \to \mathcal{C}_0)$ and $a \in \mathcal{C}_0$ one associates

$$\pi_!(X)(a) = \{(x \in X, f: a \to b) : \sigma_X(x) = b\},$$

which is a presheaf with restriction given by

$$(x, f) \cdot g = (x, fg).$$

This means that $\pi^* \pi_! X$ fits into the pullback square

$$\begin{tikzcd}
\pi^* \pi_! X \arrow[r] \arrow[d] & \mathcal{C}_1 \arrow[d, "\mathrm{cod}"] \\
X \arrow[r, "\sigma_X"'] & \mathcal{C}_0.
\end{tikzcd}$$

From this one immediately sees that $\pi_!$ preserves smallness. Furthermore, the component maps of the counit $\pi_! \pi^* \to 1$ are small covers (they are covers, because under π^* they become split epis in $\mathcal{E}/\mathcal{C}_0$; that they are also small is another consequence of the fact that the codomain map is assumed to be small).

In what follows, natural transformations of the form

$$\pi_! B \to \pi_! A$$

will play a crucial rôle and therefore it will be worthwhile to analyse them more closely. First, due to the adjunction, they correspond to maps in $\mathcal{E}/\mathcal{C}_0$ of the form

$$B \to \pi^*\pi_!A.$$

Such a map is determined by two pieces of data: a map $r: B \to A$ in $\mathcal{E}$, and, for any $b \in B$, a morphism $s_b: \sigma_B(b) \to \sigma_A(rb)$ in $\mathcal{C}$, as depicted in the following diagram:

$$\begin{tikzcd} B \arrow[r, "s"'] \arrow[rr, bend left, "\sigma_B"] \arrow[d, "r"'] & \mathcal{C}_1 \arrow[r, "\mathrm{dom}"'] \arrow[d, "\mathrm{cod}"] & \mathcal{C}_0 \\ A \arrow[r, "\sigma_A"'] & \mathcal{C}_0. \end{tikzcd} \tag{2}$$

(Note that we do not have $\sigma_A r = \sigma_B$ in general, so that it is best to consider r as a map in $\mathcal{E}$.) We will use the expression (r, s) for the map $B \to \pi^*\pi_!A$ and $(r, s)_!$ for the natural transformation $\pi_!B \to \pi_!A$ determined by a diagram as in (2).

In the following lemma, we collect the important properties of the operation $(-, -)_!$.

Lemma 3.5 *1. Assume r and s are as in diagram (2). Then $(r, s)_!: \pi_!B \to \pi_!A$ is a pointwise small map of presheaves iff $r: B \to A$ is small in $\mathcal{E}$.*

2. *Assume $r: B \to A$ is a cover and $\sigma_A: A \to \mathcal{C}_0$ is an arbitrary map. If we set $\sigma_B = \sigma_A r$ and $s_b = \mathsf{id}_{\sigma_B b}$ for every $b \in B$, then $(r, s)_!: \pi_!B \to \pi_!A$ is a cover.*

3. *If $(r, s): B \to \pi^*\pi_!A$ is a cover and $\sigma_B(b) = \mathrm{dom}(s_b)$ for all $b \in B$, then also $(r, s)_!: \pi_!B \to \pi_!A$ is a cover.*

4. *If $(r, s)_!: \pi_!B \to \pi_!A$ is a natural transformation determined by a diagram as in (2) and we are given a commuting diagram*

$$\begin{tikzcd} V \arrow[r, "p"] \arrow[d, "h"'] & B \arrow[d, "r"] \\ W \arrow[r, "q"'] & A \end{tikzcd}$$

in $\mathcal{E}$, then these data induce a commuting square of presheaves

$$\begin{tikzcd} \pi_!V \arrow[r, "\pi_!p"] \arrow[d, "(h,sp)_!"'] & \pi_!B \arrow[d, "(r,s)_!"] \\ \pi_!W \arrow[r, "\pi_!q"'] & \pi_!A, \end{tikzcd}$$

with $\sigma_V = \sigma_B p$ and $\sigma_W = \sigma_A q$. Moreover, if the original diagram is a pullback (resp. a quasi-pullback or a covering square), then so is the induced diagram.

5. *If $(r,s)_!\colon \pi_! A \to \pi_! X$ and $(u,v)_!\colon \pi_! B \to \pi_! X$ are natural transformations with the same codomain and for every $x \in X$ and every pair $(a,b) \in A \times_X B$ with $x = ra = ub$ there is a pullback square*

$$\begin{tikzcd} k_{(a,b)} \arrow[r, "q_{(a,b)}"] \arrow[d, "p_{(a,b)}"'] & \sigma_B(b) \arrow[d, "v_b"] \\ \sigma_A(a) \arrow[r, "s_a"'] & \sigma_X(x) \end{tikzcd}$$

in $\mathcal{C}$, then $\pi_!$ applied to the object $\sigma_{A\times_X B}\colon A \times_X B \to \mathcal{C}_0$ in $\mathcal{E}/\mathcal{C}_0$ obtained by sending $(a,b) \in A \times_X B$ to $k_{(a,b)}$ is the pullback of $(r,s)_!$ along $(u,v)_!$ in $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$:

$$\begin{tikzcd} \pi_!(A \times_X B) \arrow[r, "(\pi_2,q)_!"] \arrow[d, "(\pi_1,p)_!"'] & \pi_! B \arrow[d, "(u,v)_!"] \\ \pi_! A \arrow[r, "(r,s)_!"'] & \pi_! X. \end{tikzcd}$$

Proof. By direct inspection. □

Using the notation we have set up, we can list the two notions of a small map of presheaves.

1. The pointwise definition (as in the previous section): a map $F\colon B \longrightarrow A$ of presheaves is *pointwise small*, when $\pi^* F$ is a small map in $\mathcal{E}/\mathcal{C}_0$.
2. The local definition (as in [24]): a map $F\colon B \longrightarrow A$ of presheaves is *locally small*, when F is covered by a map of the form $(r,s)_!$ in which r is small in $\mathcal{E}$.

We show that these two classes of maps coincide, so that henceforth we can use the phrase "small map" without any danger of ambiguity.

Proposition 3.6 *A map is pointwise small iff it is locally small.*

Proof. We have already observed that maps of the form $(r,s)_!$ with r small are pointwise small, so all maps covered by one of this form are pointwise small as well. This shows that locally small maps are pointwise small. That all pointwise small maps are also locally small follows from the next lemma and the fact that the counit maps $\pi_! \pi^* Y \to Y$ are covers. □

Lemma 3.7 *For any pointwise small map $F\colon Z \longrightarrow Y$ and any map $L\colon \pi_! B \longrightarrow Y$ there is a quasi-pullback square of presheaves of the form*

$$\begin{tikzcd}
\pi_! C \arrow[r] \arrow[d, "(k,l)_!"'] & Z \arrow[d, "F"] \\
\pi_! B \arrow[r, "L"'] & Y,
\end{tikzcd}$$

with k small in $\mathcal{E}$.

Proof. Let S be the pullback of F along L and cover S using the counit as in:

$$\begin{tikzcd}
\pi_! \pi^* S \arrow[r, two heads] & S \arrow[r] \arrow[d] & \pi_! B \arrow[d, "L"] \\
& Z \arrow[r, "F"'] & Y.
\end{tikzcd}$$

We know the composite along the top is of the form $(k, l)_!$. Because k is the composite along the middle of the following diagram and both squares in this diagram are pullbacks, k is the composite of two small maps and hence small.

$$\begin{tikzcd}
& \mathcal{C}_1 \arrow[r, "\mathrm{cod}"] & \mathcal{C}_0 \\
\pi^* S \arrow[r] \arrow[d] & \pi^* \pi_! B \arrow[r] \arrow[u] \arrow[d, "\pi^* L"] & B \arrow[u] \\
\pi^* Z \arrow[r, "\pi^* F"'] & \pi^* Y &
\end{tikzcd}$$

$\square$

Corollary 3.8 *Every pointwise small map is covered by one of the form $(r, s)_!$ in which r is small. In fact, every composable pair (G, F) of pointwise small maps of presheaves fits into a double covering square of the form*

$$\begin{tikzcd}
\pi_! C \arrow[r, two heads] \arrow[d, "(k,l)_!"'] & Z \arrow[d, "G"] \\
\pi_! B \arrow[r, two heads] \arrow[d, "(r,s)_!"'] & Y \arrow[d, "F"] \\
\pi_! A \arrow[r, two heads] & X,
\end{tikzcd}$$

in which k and r are small in $\mathcal{E}$.

Proof. We have just shown that every pointwise small map is covered by one of the form $(r, s)_!$ in which r is small, which is the first statement. The second statement follows immediately from this and the previous lemma. $\square$

Using this alternative characterisation, we can quickly show that the collection axiom is inherited by presheaf models, as promised.

Proposition 3.9 *The collection axiom* **(A7)** *is inherited by presheaf models.*

Proof. Let $F: M \longrightarrow N$ be a small map and $Q: E \longrightarrow M$ be a cover. Without loss of generality, we may assume that F is of the form $(k, l)_!$ for some small map $k: X \longrightarrow Y$ in $\mathcal{E}$.

Let n be the map obtained by pullback in $\mathcal{E}/\mathcal{C}_0$:

$$\begin{tikzcd} T \arrow[r, two heads, "n"] \arrow[d] & X \arrow[d] \\ \pi^* E \arrow[r, two heads, "\pi^* Q"'] & \pi^* \pi_! X. \end{tikzcd}$$

Then use collection in $\mathcal{E}$ to obtain a covering square as follows:

$$\begin{tikzcd} B \arrow[r, "m"] \arrow[d, "d"'] & T \arrow[r, two heads, "n"] & X \arrow[d, "k"] \\ A \arrow[rr, two heads, "p"'] & & Y. \end{tikzcd}$$

Using Lemma 3.5.4 this leads to a covering square in the category of presheaves

$$\begin{tikzcd} \pi_! B \arrow[r, "\pi_! m"] \arrow[d, "(d, lnm)_!"'] & \pi_! T \arrow[r] \arrow[rr, bend left, "\pi_! n"] & E \arrow[r, two heads, "Q"'] & \pi_! X \arrow[d, "(k,l)_!"] \\ \pi_! A \arrow[rrr, two heads, "\pi_! p"'] & & & \pi_! Y, \end{tikzcd}$$

thus completing the proof. $\square$

Proposition 3.10 *The representability axiom is inherited by presheaf models.*

Proof. Let $\pi: E \longrightarrow U$ be a universal small map in $\mathcal{E}$, and define the following two objects in $\mathcal{E}/\mathcal{C}_0$:

$$\begin{aligned}
U' &= \{(u \in U, c \in \mathcal{C}_0, p: E_u \to \mathcal{C}_1) : \forall e \in E_u\, (\mathrm{cod}(pe) = c)\}, \\
\sigma_{U'}(u, c, p) &= c, \\
E' &= \{(u, c, p, e) : (u, c, p) \in U', e \in E_u\}, \\
\sigma_{E'}(u, c, p, e) &= \mathrm{dom}(pe).
\end{aligned}$$

If $r: E' \to U'$ is the obvious projection and $s: E' \to \mathcal{C}_1$ is the map sending (u, c, p, e) to pe, then r and s fit into a commuting square as shown:

$$\begin{tikzcd}
E' \arrow[r, "s"'] \arrow[d, "r"'] \arrow[rr, bend left, "\sigma_{E'}"] & \mathcal{C}_1 \arrow[r, "\mathrm{dom}"'] \arrow[d, "\mathrm{cod}"] & \mathcal{C}_0 \\
U' \arrow[r, "\sigma_{U'}"'] & \mathcal{C}_0.
\end{tikzcd}$$

We claim that the induced map $(r, s)_!$ in the category of presheaves is a universal small map. To show this, we need to prove that any small map F can be covered by a pullback of $(r, s)_!$. Without loss of generality, we may assume that $F = (k, l)_!$ for some small map $k: X \longrightarrow Y$ in $\mathcal{E}$.

Since π is a universal small map, there exists a diagram of the form

$$\begin{tikzcd}
E \arrow[d, "\pi"'] & V \arrow[l, "m"'] \arrow[r, "i"] \arrow[d, "h"'] & X \arrow[r, "l"] \arrow[d, "k"] & \mathcal{C}_1 \arrow[d, "\mathrm{cod}"] \\
U & W \arrow[l, "n"] \arrow[r, "j"'] & Y \arrow[r, "\sigma_Y"'] & \mathcal{C}_0,
\end{tikzcd}$$

in which the left square is a pullback and the middle one a covering square. From this, we obtain a commuting diagram of the form

$$\begin{tikzcd}
 & \mathcal{C}_0 & \\
V \arrow[ur, "\sigma_V"] \arrow[r, "m'"'] \arrow[d, "h"'] \arrow[rr, bend left, "li"] & E' \arrow[r, "s"'] \arrow[d, "r"] & \mathcal{C}_1 \arrow[ul, "\mathrm{dom}"'] \arrow[d, "\mathrm{cod}"] \\
W \arrow[r, "n'"] \arrow[rr, bend right, "\sigma_W"'] & U' \arrow[r, "\sigma_U"] & \mathcal{C}_0
\end{tikzcd}$$

by putting

$$\begin{array}{rcl}\sigma_W & = & \sigma_Y j,\\ n'w & = & (nw, \sigma_W w, \lambda e \in E_{nw}.l_{im^{-1}e}),\\ \sigma_V & = & \sigma_X i,\\ m'v & = & (n'hv, mv).\end{array}$$

Together these two commuting diagrams determine a diagram in the category of internal presheaves

$$\begin{tikzcd} \pi_! E' \arrow[d, "(r,s)_!"'] & \pi_! V \arrow[l, "\pi_! m'"'] \arrow[r, "\pi_! i"] \arrow[d, "(h,li)_!"] & \pi_! X \arrow[d, "(k,l)_!"] \\ \pi_! U' & \pi_! W \arrow[l, "\pi_! n'"] \arrow[r, "\pi_! j"'] & \pi_! Y, \end{tikzcd}$$

in which the left square is a pullback and the right one a covering square (by Lemma 3.5.4). □

Theorem 3.11 *(Assuming $\mathcal{C}$ has chosen pullbacks.) The fullness axiom* **(F)** *is inherited by presheaf models.*

Proof. In view of Lemma 2.13 and Corollary 3.8, we only need to build generic *mvs*s for maps of the form $(k, l)_! : \pi_! B \to \pi_! A$ in which k is small, where $\pi_! A$ lies over some object of the form $\pi_! X$ via a map of the form $(r, s)_!$ in which r is small. To construct this generic *mvs*, we have to apply fullness in $\mathcal{E}$. For this purpose, consider the object

$$B_0 \quad = \quad \{(b \in B, f \in \mathcal{C}_1, g \in \mathcal{C}_1) : \sigma_X(rkb) = \mathrm{cod}(f), (f^* l_b)g = \mathsf{id}\}.$$

Here $f^* l_b$ is understood to be the map fitting, for any $b \in B$ and $f : d \to c$ with $c = \sigma_X(rkb)$, in the double pullback diagram

$$\begin{tikzcd} f^*\sigma_B(b) \arrow[r] \arrow[d, "f^* l_b"'] & \sigma_B(b) \arrow[d, "l_b"] \\ f^*\sigma_A(kb) \arrow[r] \arrow[d, "f^* s_{kb}"'] & \sigma_A(kb) \arrow[d, "s_{kb}"] \\ d \arrow[r, "f"'] & c \end{tikzcd}$$

in $\mathcal{C}$. If we write $k_0 : B_0 \to A$ for the map sending (b, f, g) to $k(b)$, then this map is small, so we can use fullness in $\mathcal{E}$ to find a cover $n : W \to X$ and a small map

$m_0 \colon Z_0 \longrightarrow W$, together with a generic *mvs* P_0 for k_0 over Z_0, as depicted in the following diagram.

$$\begin{tikzcd}
P_0 \arrow[r, tail] \arrow[dr, bend right, two heads] & Z_0 \times_X B_0 \arrow[rr] \arrow[d] & & B_0 \arrow[d, "k_0"] \\
& Z_0 \times_X A \arrow[rr] \arrow[d] & & A \arrow[d, "r"] \\
& Z_0 \arrow[r, "m_0"'] & W \arrow[r, "n"'] & X
\end{tikzcd}$$

Now we make a number of definitions:

$$\begin{aligned}
Z &= \{ (z_0 \in Z_0, f \in \mathcal{C}_1) \colon \mathrm{cod}(f) = nm_0(z_0) \text{ and} \\
&\quad (\forall a \in A_{nm_0(z_0)}) \, (\exists b \in B_a) \, (\exists g \in \mathcal{C}_1) \, (z_0, b, f, g) \in P_0 \}, \\
\sigma_Z(z_0, f) &= \mathrm{dom}(f), \\
\mu(z_0, f) &= f, \\
\sigma_W(w) &= \sigma_X(nw).
\end{aligned}$$

In addition, we write $m_1 \colon Z \to Z_0$ for the obvious (small) projection and $m = m_0 m_1$. Then we obtain the following diagram of presheaves, in which both rectangles are pullbacks computed using Lemma 3.5.5:

$$\begin{tikzcd}
\pi_!(Z \times_X B) \arrow[rr] \arrow[d] & & \pi_! B \arrow[d, "(k,l)_!"] \\
\pi_!(Z \times_X A) \arrow[rr] \arrow[d] & & \pi_! A \arrow[d, "(r,s)_!"] \\
\pi_! Z \arrow[r, "(m,\mu)_!"'] & \pi_! W \arrow[r, "\pi_! n"'] & \pi_! X.
\end{tikzcd}$$

We wish to define a subpresheaf of $\pi_!(Z \times_X B)$ and prove that it is the generic *mvs* of $(k, l)_!$. We can do this by saying:

$$(z_0, f, b, h) \in P \text{ if } h \text{ factors through a map } g \text{ with } (z_0, b, f, g) \in P_0.$$

The inclusion of P in $\pi_!(Z \times_X B)$ is bounded, because P is defined by a bounded formula (using that the codomain map is small). Furthermore, the induced map from P to $\pi_!(Z \times_X A)$ is a cover by definition of Z. Thus it remains to verify genericity.

To verify this, let $E \to \pi_! W$ be any map and Q be an *mvs* of $(k, l)_!$ over E. Without loss of generality, we may assume that E is of the form $\pi_! Y$ (since E

can always be covered using the counit). This leads to the following diagram of presheaves in which the rectangles are pullbacks:

$$\begin{tikzcd}
Q \arrow[r, tail] \arrow[dr, bend right, two heads] & \bullet \arrow[rr] \arrow[d] & & \pi_! B \arrow[d, "(k,l)_!"] \\
& \bullet \arrow[rr] \arrow[d] & & \pi_! A \arrow[d, "(r,s)_!"] \\
& \pi_! Y \arrow[r, "(d,\delta)_!"'] & \pi_! W \arrow[r, "\pi_! n"'] & \pi_! X.
\end{tikzcd}$$

Of course, we will assume that the pullbacks are computed using Lemma 3.5.5, so that they are $\pi_!(Y \times_X B)$ and $\pi_!(Y \times_X A)$, respectively. It follows that in $\mathcal{E}$ we have an *mvs* Q_0 for k_0 over Y, as in

$$\begin{tikzcd}
Q_0 \arrow[r, tail] \arrow[dr, bend right, two heads] & Y \times_X B_0 \arrow[rr] \arrow[d] & & B_0 \arrow[d, "k_0"] \\
& Y \times_X A \arrow[rr] \arrow[d] & & A \arrow[d, "r"] \\
& Y \arrow[r, "d"'] & W \arrow[r, "n"'] & X,
\end{tikzcd}$$

given by

$$(y, b, f, g) \in Q_0 \text{ if } f = \delta_y \text{ and } (y, b, g) \in Q.$$

Therefore, by the genericity of P_0, there is a cover $e: U \longrightarrow Y$ and a map $c: U \longrightarrow Z_0$ with $de = m_0 c$ and

$$c^* P_0 \subseteq e^* Q_0. \tag{3}$$

Claim: If we put $t(u) = (c(u), \delta_{e(u)})$, then $t(u) \in Z$ for every $u \in U$. *Proof:* Suppose $a \in A_{nm_0c(u)}$. We know that there are $b \in B_a, f, g \in \mathcal{C}_1$ such that $(c(u), b, f, g) \in P_0$, because $P_0 \to Z_0 \times_X A$ is surjective, but the question is: do we have $f = \delta_{e(u)}$? The answer is *yes*, because if $(c(u), b, f, g) \in P_0$, then $(e(u), b, f, g) \in Q_0$ by (3). So we have $f = \delta_{e(u)}$ by definition of Q_0. $\square$

It follows that if we put

$$\sigma_U(u) := \mathrm{dom}(\delta_{e(u)}) = \sigma_Y(e(u)) = \sigma_Z(t(u)),$$

then we have the following diagram of presheaves:

$$\begin{tikzcd}
\pi_! U \arrow[r, "\pi_! t"] \arrow[d, "\pi_! e"', two heads] & \pi_!(Z) \arrow[d, "(m,\mu)_!"] \\
\pi_! Y \arrow[r, "(d,\delta)_!"'] & \pi_! W.
\end{tikzcd}$$

To see that this square commutes, we need to chase an element from $\pi_! U$ along the two sides and it suffices to this for an element of the form (u, id).

$$\begin{aligned}
(m,\mu)_! \pi_! t(u,\mathsf{id}) &= (m,\mu)_!(t(u),\mathsf{id}) \\
&= (m,\mu)_!(c(u),\delta_{e(u)},\mathsf{id}) \\
&= (m_0 c(u), \delta_{e(u)}) \\
&= (de(u), \delta_{e(u)}) \\
&= (d,\delta)_!(e(u),\mathsf{id}) \\
&= (d,\delta)_! \pi_! e(u,\mathsf{id}).
\end{aligned}$$

Therefore the proof will be finished, once we show that $(\pi_! t)^* P \subseteq (\pi_! e)^* Q$.

To show this, consider an element $(u,b,h) \in (\pi_! t)^* P$. We then have $(t(u),b,h) \in P$, which, by definition of P, means that h factors through a map g such that $(c(u),b,\delta_{e(u)},g) \in P_0$. From (3) it follows that $(e(u),b,\delta_{e(u)},g) \in Q_0$ and hence $(e(u),b,g) \in Q$, by definition of Q_0. Since Q is a presheaf, we also have $(e(u),b,h) \in Q$, whence $(u,b,h) \in (\pi_! e)^* Q$, as desired. $\square$

Remark 3.12 Diagrammatic proofs as the one we just gave are hard to read and motivate. One can give a more understandable proof using the internal logic of the category of presheaves: for those who are familiar with its intricacies, we present such a proof below.

Theorem 3.13 *(Assuming $\mathcal{C}$ has chosen finite products.) The fullness axiom* **(F)** *is inherited by presheaf models.*

Proof. In view of Lemma 2.13 and Corollary 3.8 we only need to find generic *mvs*s for small map $(k,l)_! \colon \pi_! B \to \pi_! A$, where $\pi_! A$ is fibred by a small map $(r,s)_! \colon \pi_! A \to \pi_! X$ over $\pi_! X$. Then, by replacing $\mathcal{C}$ by $\mathcal{C}/\pi_! X = \sum_{x \in X} \mathcal{C}/\sigma_X(x)$, we may even assume that $X = 1 = \{*\}$ and $\sigma_X(*) = 1$.

Internal universal quantification over *mvs*s of $\pi_! B$ in the category $\mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$ amounts to $\mathcal{E}$-internal quantification over certain subpresheaves P of $\pi_!(B) \times \mathcal{C}(-,c)$, namely those which are *mvs*s over $\pi_!(A) \times \mathcal{C}(-,c)$. Such a subpresheaf satisfies

$$\begin{gathered}
(\forall a \in A)\,(\forall h \colon d \to \sigma_A(a))\,(\forall f \colon d \to c)\,(\exists b \in B)\,(\exists g \colon d \to \sigma_B(b)) \\
(k(b), l_b \circ g) = (a,h) \text{ and } (b,g,f) \in P \subseteq \pi_!(B) \times \mathcal{C}(-,c) \text{ at } d,
\end{gathered}$$

which is equivalent to

$$\begin{gathered}
(\forall a \in A)\,(\exists b \in B)\,(\exists g \colon \sigma_A(a) \times c \to \sigma_B(b)) \\
(k(b), l_b \circ g) = (a,\pi_1) \text{ and } (b,g,\pi_2) \in P \subseteq \pi_!(B) \times \mathcal{C}(-,c) \text{ at } \sigma_A(a) \times c,
\end{gathered}$$

or

$$\begin{gathered}
(\forall a \in A)\,(\exists b \in B)\,(\exists g \colon \sigma_A(a) \times c \to \sigma_B(b)) \\
k(b) = a, l_b \circ g = \pi_1 \text{ and } (b,g,\pi_2) \in P(\sigma_A(a) \times c).
\end{gathered} \tag{4}$$

We use fullness in $\mathcal{E}$ to obtain a small family of subobjects $\{Q_i : i \in I\}$ which form a generic family of *mvs*s for

$$\{(b \in B, c \in \mathcal{C}_0, g\colon \sigma_A(kb) \times c \to \sigma_B(b)) : l_b \circ g = \pi_1\} \to A\colon (b, c, g) \mapsto k(b).$$

From these Q_i we now construct an internal small family of small presheaves of $\pi_!(B)$. Such a family is generated by subpresheaves of $\pi_!(B) \times \mathcal{C}(-, c)$ for varying $c \in \mathcal{C}_0$. For any such $c \in \mathcal{C}_0$, we take the presheaves

$$\begin{aligned} \hat{Q}_i^{(c)}(d) \quad = \quad & \{(b \in B, gh, \pi_2 h) : g\colon \sigma_A(kb) \times c \to \sigma_B(b), h\colon d \to \sigma_A(kb) \times c \\ & \text{and } (b, c, g) \in Q_i\}, \end{aligned}$$

provided i and c make the map $\hat{Q}_i^{(c)} \to \pi_!(A) \times \mathcal{C}(-, c)$ surjective.

Now we show these are generic. Take $c_0 \in \mathcal{C}$ and P a *mvs* over c_0, as in

$$\begin{tikzcd} P \arrow[r, tail] \arrow[dr, two heads] & \pi_!(B) \times \mathcal{C}(-, c_0) \arrow[d, "(k,l)_! \times \mathsf{id}"] \\ & \pi_!(A) \times \mathcal{C}(-, c_0). \end{tikzcd}$$

This P satisfies (4) for c_0, so there is a Q_i contained in

$$\{(b \in B, c_0, g\colon \sigma_A(kb) \times c_0 \to \sigma_B(b)) : (b, g, \pi_2) \in P, l_b \circ g = \pi_1\}.$$

We claim that the map $\hat{Q}_i^{(c_0)} \to \pi_!(A) \times \mathcal{C}(-, c_0)$ is surjective, and to show this it suffices to prove that elements of the form $(a, \pi_1\colon \sigma_A(a) \times c_0 \to c_0, \pi_2\colon \sigma_A(a) \times c_0 \to \sigma_A(a))$ are hit by this map. Since Q_i is a *mvs*, we know that there are $(b, c, g) \in Q_i$ with $k(b) = a$ and $l_b \circ g = \pi_1$. Since $Q_i \subseteq P$, we have $c = c_0$ and hence $(b, g, \pi_2) \in \hat{Q}_i^{(c_0)}$ and $((k, l)_! \times \mathsf{id})(-, c_0)(b, g, \pi_2) = (a, \pi_1, \pi_2)$.

So it remains to check $\hat{Q}_i^{(c_0)} \subseteq P$. But if $(b, gh, \pi_2 h) \in Q_i^{(c_0)}(d)$ for some maps $h\colon d \to \sigma_A(kb) \times c$ and $g\colon \sigma_A(kb) \times c \to \sigma_B(b)$ with $(b, c_0, g) \in Q_i$, then $(b, g, \pi_2) \in P(\sigma_A(kb) \times c)$ and hence $(b, gh, \pi_2 h) = (b, g, \pi_2) \cdot h \in P(d)$. $\square$

4 Sheaves

In this section we continue to work in the setting of a predicative category with small maps $(\mathcal{E}, \mathcal{S})$ together with an internal category $\mathcal{C}$ in $\mathcal{E}$ whose codomain map is small. To define a category of internal sheaves, we have to assume that the category $\mathcal{C}$ comes equipped with a Grothendieck topology, so as to become a Grothendieck site. There are different formulations of the notion of a site, all essentially equivalent ([23] provides an excellent discussion of this point), but for our purposes we find the following ("sifted") formulation the most useful.

Definition 4.1 Let $\mathcal{C}$ be an internal category whose codomain map in small. A *sieve* S on an object $a \in \mathcal{C}_0$ is a *small* collection of arrows in $\mathcal{C}$ all having codomain a and closed under precomposition (i.e., if $f: b \to a$ and $g: c \to b$ are arrows in $\mathcal{C}$ and f belongs to S, then so does fg). Since we insist that sieves are small, there is an object of sieves (a subobject of $\mathcal{P}_s\mathcal{C}_1$).

We call the set M_a of all arrows into a the *maximal sieve* on a (it is a sieve, since we are assuming that the codomain map is small). If S is a sieve on a and $f: b \to a$ is any map in $\mathcal{C}$, we write f^*S for the sieve $\{g: c \to b : fg \in S\}$ on b. In case f belongs to S, we have $f^*S = M_b$.

A *(Grothendieck) topology* Cov on $\mathcal{C}$ is given by assigning to every object $a \in \mathcal{C}$ a collection of sieves $\mathrm{Cov}(a)$ such that the following axioms are satisfied:

(Maximality) The maximal sieve M_a belongs to $\mathrm{Cov}(a)$;

(Stability) If $f: b \to a$ is any map and S belongs to $\mathrm{Cov}(a)$, then f^*S belongs to $\mathrm{Cov}(b)$;

(Local character) If S is a sieve on a and $R \in \mathrm{Cov}(a)$ is such that for all $f: b \to a \in R$ the sieve f^*S belongs to $\mathrm{Cov}(b)$, then S belongs to $\mathrm{Cov}(a)$.

A pair $(\mathcal{C}, \mathrm{Cov})$ consisting of a category $\mathcal{C}$ and a topology Cov on it is called a *site*. If a site $(\mathcal{C}, \mathrm{Cov})$ has been fixed, we call the sieves belonging to some $\mathrm{Cov}(a)$ *covering sieves*. If S belongs to $\mathrm{Cov}(a)$ we say that S is a *sieve covering* a, or that a *is covered by* S.

Finally, a *presentation* for a site $(\mathcal{C}, \mathrm{Cov})$ is a function BCov which yields, for every $a \in \mathcal{C}_0$, a *small* collection of *basic covering sieves* $\mathrm{BCov}(a)$ such that:

$$S \in \mathrm{Cov}(a) \Leftrightarrow \exists R \in \mathrm{BCov}(a): R \subseteq S.$$

A site for which such a presentation exists will be called *presentable*.[8]

Our first goal in this section is prove that any category of internal sheaves over a predicative category with small maps $(\mathcal{E}, \mathcal{S})$ is a positive Heyting category. The proof of this relies on the existence of a sheafification functor (a left adjoint to the inclusion of sheaves in presheaves), and since this functor is built by taking a quotient, we use the bounded exactness of $(\mathcal{E}, \mathcal{S})$. To ensure that the equivalence relation by which we quotient is bounded, we will have to assume that the site is presentable. Next, we have to identify a class of small maps in any category of internal sheaves over $(\mathcal{E}, \mathcal{S})$. We will define pointwise small and locally small maps of sheaves and we will insist that these should again coincide (as happened in presheaves). For this to work out, we again seem to need the assumption that the site is presentable; moreover, we will assume that the fullness axiom holds in $\mathcal{E}$ (note that similar assumptions were made in [21]).

[8] This is supposed to be reminiscent of Aczel's notion of a set-presentable formal space (see [3]). Note that in **IZF** every site is presentable.

So, in effect, we will work in a predicative category with small maps $(\mathcal{E}, \mathcal{S})$ equipped with a Grothendieck site $(\mathcal{C}, \mathrm{Cov})$ such that:

1. The fullness axiom **(F)** holds in $\mathcal{E}$.
2. The codomain map $\mathrm{cod}: \mathcal{C}_1 \to \mathcal{C}_0$ is small.
3. The site is presentable.

After we have shown that a category of sheaves can be given the structure of a category with small maps, we prove that the validity of any of the axioms introduced in Section 2 in $(\mathcal{E}, \mathcal{S})$ implies its validity in any category of internal sheaves over it (Theorems 4.8–4.11 and Theorem 4.17): we will say that the axiom is "inherited by sheaf models". There is one exception to this, however: we will not be able to show that the axiom **(WS)** is inherited by sheaf models. We will discuss the problem and provide a solution based on the axiom of multiple choice in Section 4.4 below (see Theorem 4.20 and Theorem 4.21).

The main results of this section are that we establish the stability of fullness **(F)** under sheaves and we correct the treatment of W-types in [29]. In addition, we show that the two different notions of a class of small maps that occur in the literature coincide in our setting. As far as the basic axioms are concerned, their stability can in one form or another already be found in the literature (see [24, 29, 20, 7]). In particular, we should point out that [7] establishes the more general result that they are stable under sheaves for a Lawvere-Tierney topology. Nevertheless, it is not quite true that our results are a special case of theirs, because, to achieve this generality, they work in a setting which has full (not just bounded) exactness. In addition, as we already mentioned in the introduction, it is not true that the fullness axiom **(F)** is stable under sheaves for a Lawvere-Tierney topology.

4.1 Sheafification

Our next theorem shows the existence of a sheafification functor, a Cartesian left adjoint to the inclusion of sheaves in presheaves. The proof relies in an essential way on the assumption of bounded exactness and on the fact that our site is presentable.

Theorem 4.2 *The inclusion*

$$i_*: \mathrm{Sh}_{\mathcal{E}}(\mathcal{C}) \rightarrowtail \mathrm{Psh}_{\mathcal{E}}(\mathcal{C})$$

has a Cartesian left adjoint i^* (a "sheafification functor").

Proof. We verify that it is possible to imitate the standard construction (see [27, Section III.5]).

Let P be a presheaf. A pair (R, x) will be called a *compatible family* on $a \in \mathcal{C}_0$, if R is a covering sieve on a, and x specifies for every $f\colon b \longrightarrow a \in R$ an element $x_f \in P(b)$, such that for any $g\colon c \longrightarrow b$ the equality $(x_f) \cdot g = x_{fg}$ holds. Because **(ΠE)** holds and sieves are small, by definition, there is an object of compatible families. Actually, the compatible families form a presheaf $\mathrm{Comp}(P)$ with restriction given by

$$(R, x) \cdot f = (f^* R, x \cdot f),$$

where $(x \cdot f)_g = x_{fg}$.

We define an equivalence relation on $\mathrm{Comp}(P)$ by declaring two compatible families (R, x) and (T, y) on a equivalent, when there is a covering sieve $S \subseteq R \cap T$ on a with $x_f = y_f$ for all $f \in S$. Since the site is assumed to be presentable, this quantification over the (large) collection of covering sieves S on a, can be replaced with a quantification over the small collection of basic covering sieves on a. Therefore the equivalence relation is bounded and has a quotient P^+. This object P^+ is easily seen to carry a presheaf structure in such a way that the quotient map $\mathrm{Comp}(P) \to P^+$ is a morphism of presheaves.

First claim: P^+ is separated. Proof: Suppose two elements $[R, x]$ and $[S, y]$ of $\mathcal{P}^+(a)$ agree on a cover T. Pick representatives (R, x) and (S, y), and define:

$$Q = \{f\colon b \longrightarrow a \in R \cap S : x_f = y_f\}.$$

Once we show that Q is covering, we are done. But this follows immediately from the local character axiom for sites: for any $f \in T$, the sieve $f^* Q$ is covering, by assumption.

Second claim: when P is separated, P^+ is a sheaf. Proof: Let R be a covering sieve on a, and let compatible elements $p_f \in P^+(b)$ be given for every $f\colon b \longrightarrow a \in R$. Using the collection axiom, we find for every $f \in R$ a family of representatives $(R^{(f,i)}, x^{(f,i)})$ of p_f, with the variable i running through some inhabited and *small* index set I_f. Therefore

$$S = \{\, f \circ g : f \in R, i \in I_f, g \in R^{(f,i)} \,\}$$

is small; in fact, it is a covering sieve, by local character.

We now prove that for any two triples $(f \in R, i \in I_f, g \in R^{(f,i)})$ and $(f' \in R, i' \in I_{f'}, g' \in R^{(f',i')})$ with $fg = f'g'$, we must have $x_g^{(f,i)} = x_{g'}^{(f',i')}$. Since the elements p_f are assumed to be compatible, the equality

$$[R^{(f,i)}, x^{(f,i)}] \cdot g = p_{fg} = p_{f'g'} = [R^{(f',i')}, x^{(f',i')}] \cdot g'$$

holds. Hence the elements $x_g^{(f,i)}$ and $x_{g'}^{(f',i')}$ agree on a covering sieve. Since P is assumed to be separated, this implies that the elements $x_g^{(f,i)}$ and $x_{g'}^{(f',i')}$ are in fact identical.

This argument shows that the definition $z_{fg} = x_g^{(f,i)}$ is unambiguous for $fg \in S$, and also that (S, z) is a compatible family. As its equivalence class $[S, z]$ is the glueing of the family p_f we started with, the second claim is proved.

From the construction it is clear that for any presheaf P the sheaf P^{++} has to be its sheafification. So we have shown that the construction of the sheafification functor carries through in the setting we are working in; that this assignment is moreover functorial as well as Cartesian is proved in the usual manner. $\square$

Theorem 4.3 $\mathrm{Sh}_{\mathcal{E}}(\mathcal{C})$ *is a positive Heyting category.*

Proof. The category of sheaves has finite limits, because these are computed pointwise, as in presheaves. Using the following description of images and covers in categories of sheaves, one can easily show these categories have to be regular: the image of a map $F: Y \longrightarrow X$ of sheaves consists of those $x \in X(a)$ that are "locally" hit by F, i.e., for which there is a sieve S covering a such that for any $f: b \longrightarrow a \in S$ there is an element $y \in Y(b)$ with $F(y) = x \cdot f$. Therefore a map $F: Y \longrightarrow X$ is a cover, if for every $x \in X(a)$ there is a sieve S covering a and for any $f: b \to a \in S$ an element $y \in Y(b)$ such that $F(y) = x \cdot f$ (such maps are also called *locally surjective*).

The Heyting structure in sheaves is the same as in presheaves, so the universal quantification of $A \subseteq Y$ along $F: Y \longrightarrow X$ is given by the formula (1). Indeed, from this description is readily seen that belonging to $\forall_F(A)$ is a local property.

The sums in sheaves are obtained by sheafifying the sums in presheaves. They are still disjoint and stable, because the sheafification functor is Cartesian. $\square$

4.2 Small maps in sheaves

We will now define two classes of maps in the categories of sheaves, those which are pointwise small and those which are locally small. Using that **(F)** holds in $\mathcal{E}$ and the fact that the site is presentable, we will then show that they coincide. But before we define these two classes of maps, note that we have the following diagram of functors:

$$\begin{tikzcd} \mathcal{E}/\mathcal{C}_0 \arrow[rr, shift left, "\pi_!"] \arrow[dr, shift left, "\rho_!"] & & \mathrm{Psh}_{\mathcal{E}}(\mathcal{C}) \arrow[ll, shift left, "\pi^*"] \arrow[dl, shift left, "i^*"] \\ & \mathrm{Sh}_{\mathcal{E}}(\mathcal{C}), \arrow[ul, shift left, "\rho^*"] \arrow[ur, shift left, "i_*"] & \end{tikzcd}$$

where the maps ρ^* and $\rho_!$ are defined as the composites of π and i via the diagram. So ρ^* is the forgetful functor, $\rho_!$ is defined as

$$\rho_! X = i^* \pi_! X,$$

and they are adjoint. It follows immediately from the maximality axiom for sites that the components of the counit $\rho_!\rho^* \longrightarrow 1$ are covers.

One final remark before we give the definitions. We have seen that any pair of maps (r, s) in $\mathcal{E}$ making

$$\begin{tikzcd} B \arrow[r, "s"'] \arrow[rr, bend left, "\sigma_B"] \arrow[d, "r"'] & \mathcal{C}_1 \arrow[r, "\mathrm{dom}"'] \arrow[d, "\mathrm{cod}"] & \mathcal{C}_0 \\ A \arrow[r, "\sigma_A"'] & \mathcal{C}_0. & \end{tikzcd}$$

commute determines a map $(r, s)_! \colon \pi_! B \to \pi_! B$ of presheaves. Therefore it also determines a map $i^*(r, s)_! \colon \rho_! B \to \rho_! A$ of sheaves, but note that now not all maps $\rho_! B \to \rho_! A$ will be of this form, in contrast to what happened in the presheaf case.

Finally, the two classes of maps are defined as:

1. The pointwise definition: a morphism $F \colon B \longrightarrow A$ of sheaves is *pointwise small*, when $\rho^* F$ is a small map in $\mathcal{E}/\mathcal{C}_0$.
2. The local definition (as in [24]): a morphism $F \colon B \longrightarrow A$ of sheaves is *locally small* in case it is covered by a map of the form $i^*(r, s)_!$ where r is a small map in $\mathcal{E}$.

That these two classes of maps coincide will follow from the next two propositions, both whose proofs use the fullness axiom.

Proposition 4.4 *The sheafification functor i^* preserves pointwise smallness: if F is a (pointwise) small map of presheaves, then $i^* F$ is a pointwise small map of sheaves.*

Proof. To prove the proposition, it suffices to show that the $(-)^+$-construction preserves smallness. So let $F \colon P \longrightarrow Q$ be a (pointwise) small morphism of presheaves and q be an element of $Q^+(a)$, i.e. $q = [R, x]$ where R is a sieve and $(x_f)_{f \in R}$ is a family of compatible elements. The fibre of F^+ over q consists of equivalence classes of all those compatible families (S, y) on a such that $(S, F(y))$ and (R, q) are equivalent (by $F(y)$ we of course mean the family given by $F(y)_f = F(y_f)$). Because every such equivalence class is represented by a compatible family (S, y) where S is a *basic* covering sieve contained in R and $F(y_f) = x_f$ for all $f \in S$, the fibre of F over q is covered by the object:

$$\sum_{S \in \mathrm{BCov}(a), S \subseteq R} \prod_{f \in S} F^{-1}(x_f).$$

It follows from the fullness axiom in $\mathcal{E}$ that this object is small (actually, the exponentiation axiom **(ΠS)** would suffice for this purpose) and then it follows from the quotient axiom **(A6)** that the fibre of F over q is small as well. $\Box$

Proposition 4.5 *The pointwise small maps in sheaves are closed under covered maps: if*

$$\begin{tikzcd} X \arrow[r] \arrow[d, "F"'] & A \arrow[d, "G"] \\ Y \arrow[r, two heads, "P"'] & B \end{tikzcd}$$

is a covering square of sheaves (i.e., P and the induced map $X \to Y \times_B A$ are locally surjective) and F is pointwise small, then also G is pointwise small.

Proof. To make the proof more perspicuous, we will split the argument in two: first we show closure of pointwise small maps under quotients and then under descent.

So suppose first that we have a commuting triangle of sheaves

$$\begin{tikzcd} Y \arrow[rr, two heads, "G"] \arrow[dr, "F"'] & & X \arrow[dl, "H"] \\ & B, & \end{tikzcd}$$

with F pointwise small and G locally surjective. Fix an element $b \in B(c)$. The fullness axiom in $\mathcal{E}$ implies that for any basic covering sieve $S \in \mathrm{BCov}(c)$ there is a small generic family P_b^S of *mvs*s of the obvious (small) projection map

$$p_b^S : Y_b^S = \{(f : d \to c \in S, y \in Y(d)) : F_d(y) = b \cdot f\} \longrightarrow S,$$

such that any *mvs* of this map is refined by one in P_b^S (recall that an *mvs* of p_b^S would be a subobject $L \subseteq Y_b^S$ such that the composite $L \subseteq Y_b^S \to S$ is a small cover). Strictly speaking, the fullness axiom says that for every $S \in \mathrm{BCov}(c)$ such a generic *mvs* exists, not necessarily as a function of S. This does follow, however, using the collection axiom: for this axiom tells us that there is a small family $\{P_i : i \in I_b^S\}$ of such *mvs*s for every S. So we can set $P_b^S = \bigcup_{i \in I_b^S} P_i$ to get a generic *mvs* of p_b^S as a function of S.

Call an element $L \in P_b^S$ *compatible after* G, if for any pair of elements $(f\colon d \to c, y)$ and $(f'\colon d' \to c, y')$ in L we have

$$\forall g\colon e \to d, g'\colon e \to d' \,(\, fg = f'g' \Rightarrow G_d(y) \cdot g = G_{d'}(y') \cdot g' \,).$$

Note that there is a map

$$q\colon \sum_{S \in \mathrm{BCov}(c)} \{L \in P_b^S : L \text{ compatible after } G\,\} \to H_c^{-1}(b),$$

which one obtains by sending (S, L) to the glueing of the elements $\{G_d(y) : (f\colon d \to c, y) \in L\}$ in X. The domain of this map q is small, so the desired result will follow, once we show that this map is a cover. For this we use the local surjectivity of G.

Local surjectivity of G means that for every $x \in X(c)$ in the fibre over $b \in B(c)$, there is a basic covering sieve $S \in \mathrm{BCov}(c)$ such that

$$\forall f\colon d \to c \in S\, \exists y \in Y(d) : G_d(y) = x \cdot f.$$

But $G_d(y) = x \cdot f$ implies that $F_d(y) = b \cdot f$, so

$$\{(f\colon d \to c, y \in Y(d)) : G_d(y) = x \cdot f\}$$

is an *mvs* of p_b^S and therefore it is refined by an element of P_b^S. Since this element must be compatible after G, we have shown that q is a cover.

Second, suppose we have a pullback square of sheaves

$$\begin{tikzcd} X \arrow[r, "Q"] \arrow[d, "F"'] & A \arrow[d, "G"] \\ Y \arrow[r, two heads, "P"'] & B, \end{tikzcd}$$

where F is pointwise small and P and Q are locally surjective. Again, for any $b \in B(c)$ and basic covering sieve S of c, let p_b^S be the map

$$p_b^S\colon Y_b^S = \{(f\colon d \to c \in S, y \in Y(d)) : P_d(y) = b \cdot f\} \longrightarrow S,$$

as above. Furthermore, let $\mathrm{mvs}(p_b^S)$ be the object of *mvs*s of p_b^S and set

$$\begin{aligned} Y'(c) &= \sum_{b \in B(c)} \sum_{S \in \mathrm{BCov}(c)} \mathrm{mvs}(p_b^S), \\ X'(c) &= \sum_{(b,S,L) \in Y'(c)} \{\, k \in \prod_{(f:d \to c, y) \in L} F_d^{-1}(y) : k \text{ compatible after } Q\,\}, \end{aligned}$$

where we call $k \in \prod_{(f:d\to c,y)\in L} F_d^{-1}(y)$ *compatible after* Q, if for any $(f\colon d \to c, y)$ and $(f'\colon d' \to c, y')$ in L we have

$$\forall g\colon e \to d, g'\colon e \to d' \in \mathcal{C}\, (fg = f'g' \Rightarrow Q_d(k_{(f,y)}) \cdot g = Q_{d'}(k_{(f',y')}) \cdot g').$$

This leads to a commuting square in $\mathcal{E}/\mathcal{C}_0$

$$\begin{tikzcd}
X'(c) \arrow[r, two heads, "Q'_c"] \arrow[d, "F'_c"'] & A(c) \arrow[d, "G_c"] \\
Y'(c) \arrow[r, two heads, "P'_c"'] & B(c),
\end{tikzcd}$$

in which P' and F' are the obvious projections and Q' sends (b, S, L, k) to the glueing of $\{Q_d(k_{(f,y)}) \colon (f, y) \in L\}$. The square is a pullback in which the map P' is a cover (this uses the collection axiom) and F' is small, so that G_c is a small map by descent **(A2)** in $\mathcal{E}/\mathcal{C}_0$. This completes the proof. $\square$

Theorem 4.6 *The pointwise small maps and locally small maps of sheaves coincide.*

Proof. That all locally small maps of sheaves are also pointwise small follows from the previous two propositions. To prove that all pointwise small maps are also locally small we use that the pointwise and locally small maps coincide in presheaves.

So consider a pointwise small map $F \colon B \longrightarrow A$ of sheaves. Since i_*F is a pointwise small map of presheaves, there is a small map of presheaves $(k, l)_!$ with k small in $\mathcal{E}$ such that

$$\begin{tikzcd}
\pi_! X \arrow[r] \arrow[d, "(k,l)_!"'] & i_* B \arrow[d, "i_* F"] \\
\pi_! Y \arrow[r] & i_* A
\end{tikzcd}$$

is a covering square in presheaves. Applying sheafification i^* and using that $i^* i_* \cong 1$, we obtain a diagram of the desired form. $\square$

Corollary 4.7 *Any pointwise small map is covered by one of the form $i^*(r, s)_!$ with r small in $\mathcal{E}$. In fact, every composable pair (G, F) of pointwise small maps of sheaves fits into a double covering square of the form*

$$\begin{tikzcd}
\rho_! C \arrow[r, two heads] \arrow[d, "i^*(k,l)_!"'] & Z \arrow[d, "G"] \\
\rho_! B \arrow[r, two heads] \arrow[d, "i^*(r,s)_!"'] & Y \arrow[d, "F"] \\
\rho_! A \arrow[r, two heads] & X,
\end{tikzcd}$$

in which k and r are small in $\mathcal{E}$.

Proof. Immediate from the previous theorem and the corresponding fact for presheaves (Corollary 3.8). □

Henceforth we can therefore use the term "small map" without danger of ambiguity. The first thing to do now is to show that the small maps in sheaves really satisfy the axioms for a class of small maps.

Theorem 4.8 *The small maps in sheaves satisfy axioms* **(A1-9)**.

Proof. Again, we postpone the proof of the collection axiom **(A7)** (it will be Theorem 4.10). Because limits in sheaves are computed as in presheaves, **(A1)** and **(A9)** are inherited from presheaves. Colimits in sheaves are computed by sheafifying the result in presheaves, hence the axioms **(A3)** and **(A4)** follow from Proposition 4.4. That pointwise small maps are closed under covered maps was Proposition 4.5: this disposes of **(A2)** and **(A6)**. Pointwise small maps are closed under composition, so **(A5)** holds as well. Finally, since universal quantification in sheaves is computed as in presheaves, the axiom **(A8)** holds in sheaves, because it holds in presheaves. □

Theorem 4.9 *The following axioms are inherited by sheaf models: bounded exactness, representability,* **(NE)**, **(NS)**, **(ΠE)**, **(ΠS)***,* **(M)** *and* **(PS)**.

Proof. Bounded exactness is inherited by sheaf models, since one can sheafify the quotient in presheaves. Representability is inherited for the same reason: one sheafifies the universal small maps in presheaves. Also the natural numbers object in sheaves is obtained by sheafifying the natural numbers object in presheaves, so **(NE)** and **(NS)** are inherited by sheaf models. Since Π-types in presheaves are computed as in sheaves and **(ΠE)** and **(ΠS)** are inherited by presheaf models, they will also be inherited by sheaf models. Finally, since monos in sheaves are pointwise, **(M)** is inherited as well.

The $\mathcal{P}_s$-functor in sheaves is obtained by quotienting the $\mathcal{P}_s$-functor in presheaves (see Proposition 3.2) by the following equivalence relation (basically, bisimulation understood as in sheaves): if $A, A' \subseteq \mathbf{y}c \times X$, then $A \sim A'$ if for all $(f\colon b \to c, x) \in A(c)$, the sieve

$$\{\, g\colon a \to b : (fg, x \cdot g) \in A' \,\}$$

covers b, and for all $(f'\colon b' \to c, x') \in A'(c)$ the sieve

$$\{\, g'\colon a' \to b' : (f'g', x' \cdot g') \in A \,\}$$

covers b'.

One easily verifies that this defines an equivalence relation in presheaves; moreover, it is bounded, since the site is assumed to be presentable. Its quotient

P has the structure of a sheaf (as we have seen several times, to construct the glueing one uses the collection axiom to select small collections of representatives from each equivalence class). One defines the relation $\in_X \subseteq X \times P$ on an object $c \in \mathcal{C}$ by putting for any $x \in X(c)$ and $A \in \mathcal{P}_s(X)(c)$,

$$x \in [A] \Longleftrightarrow \text{the sieve } \{f\colon d \to c\colon (f, x \cdot f) \in A\} \text{ covers } c.$$

A straightforward verification establishes that this is indeed the power class object of X in sheaves. Hence the axiom **(PS)** is inherited by sheaf models. $\square$

In the coming two subsections we will discuss the collection and fullness axioms and W-types in sheaf categories.

4.3 Collection and fullness in sheaves

Theorem 4.10 *The collection axiom* **(A7)** *is inherited by sheaf models.*

Proof. Let $F\colon M \longrightarrow N$ be small map and $E\colon Y \longrightarrow M$ be a cover in sheaves (i.e. E is locally surjective). Without loss of generality we may assume that F is of the form $i^*(k,l)_!\colon \rho_! B \to \rho_! A$.

If the map $Q\colon X \to \pi_! B$ of presheaves is obtained by pulling back the map $i_* E$ along the component of the unit $1 \to i_* i^*$ at $\pi_! B$ as in

$$\begin{tikzcd} X \arrow[r] \arrow[d, "Q"'] & i_* Y \arrow[d, "i_* E"] \\ \pi_! B \arrow[r] & i_* \rho_! B = i_* i^* \pi_! B, \end{tikzcd}$$

then this map Q also has to be locally surjective. This means that for the following object in $\mathcal{E}$

$$\begin{aligned} C \quad = \quad & \{\, (b \in B, S \in \mathrm{Cov}(c))\colon \sigma_B(b) = c \text{ and} \\ & (\forall f\colon d \to c \in S)\,(\exists x \in X(d))\,(Q(x) = (b, f))\}, \end{aligned}$$

the obvious projection $s_0\colon C \to B$ is a cover. Therefore we can apply the collection axiom in $\mathcal{E}$ to obtain a covering square of the form:

$$\begin{tikzcd} V \arrow[r, "s_1"] \arrow[d, "l"'] & C \arrow[r, two heads, "s_0"] & B \arrow[d, "k"] \\ U \arrow[rr, "r_0"'] & & A, \end{tikzcd} \tag{5}$$

with l small in $\mathcal{E}$. We wish to apply the collection axiom again. For this purpose, define the following two objects in $\mathcal{E}$:

$$
\begin{aligned}
V' &= \{ (v \in V, f \in \mathcal{C}) \colon \text{if } s_1(v) = (b, S), \text{then } f \in S \}, \\
W &= \{ (v \in V, f \colon d \to c, x \in X(d)) \colon \text{if } s_1(d) = (b, S), \\
&\quad \text{then } f \in S \text{ and } Q_d(x) = (b, f) \},
\end{aligned}
$$

and let $s_3 \colon W \to V'$ and $s_2 \colon V' \to V$ be the obvious projections. s_3 is a cover (essentially by definition of C), and the composite $l' = ls_2$ is small. So we can apply collection to obtain a covering square in $\mathcal{E}$

$$
\begin{tikzcd}
J \arrow[r, "s_4"] \arrow[d, "m"'] & W \arrow[r, two heads, "s_3"] & V' \arrow[d, "l'"] \\
I \arrow[rr, "r_1"'] & & U,
\end{tikzcd}
\tag{6}
$$

in which m is small. Writing $r = r_0 r_1$ and $s = s_0 s_1 s_2 s_3 s_4$, we obtain a commuting square

$$
\begin{tikzcd}
J \arrow[r, "s"] \arrow[d, "m"'] & B \arrow[d, "k"] \\
I \arrow[r, "r"'] & A,
\end{tikzcd}
$$

with every $j \in J$ determining an element $b \in B$, a sieve S on $\sigma_B(b)$, an arrow $f \in S$ and an element $x \in X(\mathrm{dom}\ f)$ such that $Q(x) = (b, f) \in \pi_! B$. If for such an element $j \in J$ we put $\rho_J(j) = \mathrm{dom}(f), t_j = f, n_j = l_b \circ f$ and for every $i \in I$ we define $\sigma_I(i) = \sigma_A(ri)$, then we obtain a square of presheaves:

$$
\begin{tikzcd}
\pi_! J \arrow[r, "(s,t)_!"] \arrow[d, "(m,n)_!"'] & \pi_! B \arrow[d, "(k,l)_!"] \\
\pi_! I \arrow[r, "\pi_! r"'] & \pi_! A.
\end{tikzcd}
$$

To see that it commutes, we chase an element around the two sides of the diagram and it suffices to do that for an element of the form (j, id). So

$$
\pi_! r (m, n)_! (j, \mathsf{id}) = \pi_! r (mj, l_b \circ f) = (rmj, l_b \circ f),
$$

and $(k, l)_! (s, t)_! (j, \mathsf{id}) = (k, l)_! (sj, f) = (ksj, l_b \circ f)$.

We claim that sheafifying the square gives a covering square. Since r is a cover and $\rho_!$ preserves these, this means that we have to show that the map from $\pi_! J$ to the pullback of the above square is locally surjective. Lemma 3.5.4 tells us that we may assume that the pullback is of the form $\pi_!(I \times_A B)$ with $\sigma_{I \times_A B}(i, b) = \sigma_B(b)$. The induced map $K \colon \pi_! J \longrightarrow \pi_!(I \times_A B)$ sends (j, g) to $((mj, sj), f \circ g)$, where f is the element in $\mathcal{C}_1$ determined by $j \in J$ as above. To show that this map is locally surjective, it suffices to prove that every element

$((i,b),\mathsf{id}) \in \pi_!(I \times_A B)$ is locally hit by K. The element $i \in I$ determines an element $r_1 i \in U$, and since (5) is a covering square, we find a $v \in V$ with $lv = r_1 i$ and $s_0 s_1 v = b$, hence a covering sieve S on $\rho_B(b)$. Moreover, since (6) is a covering square, we find for every $f \in S$ an element $j \in J$ such that $m(j) = i$ and $s(j) = b$. Then $K(j,\mathsf{id}) = ((i,b),f) = ((i,b),\mathsf{id}) \cdot f$, which proves that K is locally surjective.

To complete the proof, we need to show that $(s,t)_!\colon \pi_! J \longrightarrow \pi_! B$ factors through $Q\colon X \longrightarrow \pi_! B$. There is a map $(p,q)\colon J \longrightarrow \pi^* X$ which sends every $j \in J$ to the $x \in X(\mathrm{dom}\, f)$ that it determines. Its transpose $(p,q)_!$ sends (j,id) to $x \in X$ which in turn is sent by Q to $Q(x) = (sj, f) = (s,t)_!(j,\mathsf{id})$. Therefore $(s,t)_! = Q(p,q)_!$. □

Theorem 4.11 *(Assuming $\mathcal{C}$ has chosen pullbacks.) The fullness axiom* **(F)** *is inherited by sheaf models.*

Proof. In view of Lemma 2.13 and Corollary 4.7, it will suffice to show that there exists a generic *mvs* for any map of the form $i^*(k,\kappa)_!\colon \rho_! B \longrightarrow \rho_! A$, living over some object of the form $\rho_! X$ via some map $i^*(r,\rho)_!\colon \rho_! A \longrightarrow \rho_! X$, with k and r small.

We first construct the generic *mvs* P. To this end, define:

$$\begin{aligned}
S_0 &= \{(a \in A, \alpha\colon d \to c, S \in \mathrm{BCov}(\alpha^*\sigma_A(a))) : \sigma_X(ra) = c\} \\
M_0 &= \{(a \in A, \alpha\colon d \to c, S \in \mathrm{BCov}(\alpha^*\sigma_A(a)), \beta \in S) : \sigma_X(ra) = c\} \\
B_0 &= \{(b \in B, \alpha\colon d \to c, S \in \mathrm{BCov}(\alpha^*\sigma_A(kb)), \beta \in S, \gamma \in \mathcal{C}_1 : \\
&\qquad \sigma_X(rkb) = c, \alpha^*\kappa_b \circ \gamma = \beta\}
\end{aligned}$$

(In the definition of S_0 and M_0 we have used that any pair consisting of a map $\alpha\colon d \to c \in \mathcal{C}$ and element $a \in A$ with $\sigma_X(ra) = c$ determines a pullback diagram

$$\begin{tikzcd}
\alpha^*\sigma_A(a) \arrow[r] \arrow[d, "\alpha^*\rho_a"'] & \sigma_A(a) \arrow[d, "\rho_a"] \\
d \arrow[r, "\alpha"'] & c
\end{tikzcd}$$

in $\mathcal{C}$; in the definition of B_0 we have used that any pair consisting of a map $\alpha\colon d \to c \in \mathcal{C}$ and element $b \in B$ with $\sigma_X(rkb) = c$ determines a double pullback diagram

$$\begin{tikzcd}
\alpha^*\sigma_B(b) \arrow[r] \arrow[d, "\alpha^*\kappa_b"'] & \sigma_B(b) \arrow[d, "\kappa_b"] \\
\alpha^*\sigma_A(kb) \arrow[r] \arrow[d, "\alpha^*\rho_{kb}"'] & \sigma_A(kb) \arrow[d, "\rho_{kb}"] \\
d \arrow[r, "\alpha"'] & c
\end{tikzcd}$$

in $\mathcal{C}$.) One easily checks that all the projections in the chain

$$B_0 \longrightarrow M_0 \longrightarrow S_0 \longrightarrow A \longrightarrow X$$

are small.

For the construction of P, we first build a generic *mvs* for $S_0 \longrightarrow A$ over X. This means we have a cover $n\colon W \longrightarrow X$ and a small map $m_0\colon Z_1 \longrightarrow W$, together with a generic *mvs* P_1 for $S_0 \longrightarrow A$ over Z_1, as in the diagram

$$\begin{tikzcd}
P_1 \arrow[r, tail] \arrow[dr, two heads, bend right] & S_1 \arrow[rr] \arrow[d] & & S_0 \arrow[d] \\
& A_1 \arrow[rr] \arrow[d] & & A \arrow[d] \\
& Z_1 \arrow[r, "m_0"'] & W \arrow[r, two heads, "n"'] & X,
\end{tikzcd}$$

where the rectangles are understood to be pullbacks. Next, we pull $B_0 \to M_0 \to S_0$ back along $P_1 \to S_0$ and obtain the diagram

$$\begin{tikzcd}
B_1 \arrow[r] \arrow[d] & B_0 \arrow[d] \\
M_1 \arrow[r] \arrow[d] & M_0 \arrow[d] \\
P_1 \arrow[r] & S_0.
\end{tikzcd}$$

Then we build a generic *mvs* for $B_1 \to M_1$ over Z_1. This we obtain over an object Z_2 via a small map $Z_2 \longrightarrow W'$ and a cover $W' \longrightarrow Z_1$. Without loss of generality, we may assume that the latter map $W' \to Z_1$ is the identity. (Proof: apply the collection axiom to the small map $Z_1 \longrightarrow W$ and the cover $W' \longrightarrow Z_1$ to obtain a small map $S \longrightarrow R$ covering the morphism $Z_1 \longrightarrow W$. Lemma 2.12 tells us that there lives a generic *mvs* for $S_0 \longrightarrow A$ over S as well. By another application of Lemma 2.12, there lives a generic *mvs* for $B_1 \to M_1$ over T, if $T \to S$ is the pullback of $Z \longrightarrow W'$ along the map $S \longrightarrow W'$.) So we may assume there is a small map $m_1\colon Z_2 \longrightarrow Z_1$, such that over Z_2 there is a generic *mvs* P_2 for $B_1 \to M_1$, as in the following diagram

$$\begin{tikzcd}
P_2 \arrow[r, tail] \arrow[dr, two heads, bend right] & B_2 \arrow[r] \arrow[d] & B_1 \arrow[rr] \arrow[d] & & B_0 \arrow[d] \\
& M_2 \arrow[r] \arrow[d] & M_1 \arrow[rr] \arrow[d] & & M_0 \arrow[d] \\
& Z_2 \arrow[r, "m_1"] \arrow[rrr, two heads, bend right, "t"'] & Z_1 \arrow[r, "m_0"] & W \arrow[r, two heads, "n"] & X,
\end{tikzcd}$$

where all the rectangles are supposed to be pullbacks. Fo convenience, write $t = nm_0m_1$.

We make some definitions. First of all, let

$$\begin{aligned} Z &= \{(z_2 \in Z_2, \delta \colon d \to c) : \sigma_X(t(z_2)) = c \text{ and} \\ &\quad (\forall a \in A_{t(z_2)})\, (\exists S \in \mathrm{BCov}(\delta^*\sigma_A(a))\, (m_1(z), a, \delta, S) \in P_1\}. \end{aligned}$$

Furthermore, we write $m_2 \colon Z \to Z_2$ for the obvious projection and put $m = m_0m_1m_2$. Finally, we let P_3 be the pullback of P_2 along m_2.

We wish to construct a diagram of presheaves of the form:

$$\begin{tikzcd} P \arrow[r, tail] \arrow[dr, bend right] & \pi_!(Z \times_X B) \arrow[rr] \arrow[d] & & \pi_! B \arrow[d, "(k,\kappa)_!"] \\ & \pi_!(Z \times_X A) \arrow[rr] \arrow[d] & & \pi_! A \arrow[d, "(r,\rho)_!"] \\ & \pi_! Z \arrow[r, "(m,\mu)_!"'] & \pi_! W \arrow[r, "\pi_! n"'] & \pi_! X, \end{tikzcd}$$

which we can do by putting $\sigma_Z(z_2, \delta) = \mathrm{cod}(\delta)$ and $\mu_{(z_2,\delta)} = \delta$. Note that $\pi_! n$ is a cover and $(m, \mu)_!$ is small. In addition, P is defined by saying that an element $(z \in Z, b \in B, \eta \colon c \to d) \in \pi_!(Z \times_X B)(c)$ belongs to $P(c)$ if

> there is a sieve $S \in \mathrm{BCov}(\mu_z^*\sigma_A(kb))$, a map $\beta \in S$ and a map $\gamma \in \mathcal{C}_1$ such that $(z, b, \mu_z, S, \beta, \gamma)$ belongs to P_3 and η factors through γ.

By construction, the map $P \longrightarrow \pi_!(Z \times_X A)$ is locally surjective. By sheafifying the whole diagram, we therefore obtain an *mvs* i^*P for $i^*(k, \kappa)_!$ over $\rho_! Z$ in the category of sheaves. The remainder of the proof will show it is generic.

To that purpose, let $V \longrightarrow \rho_! W$ be a map of sheaves and Q be an *mvs* for $i^*(k, \kappa)_!$ over V. Let $\overline{Y}$ be the pullback in presheaves of V along the map $\pi_! W \to \rho_! W$ and cover $\overline{Y}$ using the counit $\pi_!\pi^*\overline{Y} \to \overline{Y}$. Writing $Y = \pi^*\overline{Y}$, this means we have a commuting square of presheaves

$$\begin{tikzcd} \pi_! Y \arrow[r, "(l,\lambda)_!"] \arrow[d] & \pi_! W \arrow[d] \\ V \arrow[r] & \rho_! W, \end{tikzcd}$$

in which the vertical arrows are locally surjective and the top arrow is of the form $(l, \lambda)_!$. Finally, let $\overline{Q}$ be the pullback of Q along $\pi_! Y \to V$. This means we have the following diagram of presheaves:

$$\begin{tikzcd}
\overline{Q} \arrow[r, tail] \arrow[dr, bend right] & \pi_!(Y \times_X B) \arrow[rr] \arrow[d] & & \pi_! B \arrow[d, "(k,\kappa)_!"] \\
& \pi_!(Y \times_X A) \arrow[rr] \arrow[d] & & \pi_! A \arrow[d, "(r,\rho)_!"] \\
& \pi_! Y \arrow[r, "(l,\lambda)_!"'] & \pi_! W \arrow[r, "\pi_! n"'] & \pi_! X,
\end{tikzcd},$$

where the rectangles are pullbacks, computed, as usual, using Lemma 3.5.5 (so $\sigma_{Y\times_X A}(y,a) = \lambda_y^*\sigma_A(a)$ and $\sigma_{Y\times_X B}(y,b) = \lambda_y^*\sigma_B(b)$). The map $\overline{Q} \to \pi_!(Y \times_X A)$ is locally surjective, and therefore

$$\begin{aligned}
Q_1 &= \{\, (y, a, \lambda_y, S \in \mathrm{BCov}(\lambda_y^*\sigma_A(a))) : \\
&\qquad (\forall \beta \in S)\,(\exists b \in B_a)\,(\exists \gamma \in \mathcal{C}_1)\,(y,b,\gamma) \in \overline{Q} \text{ and } \lambda_y^*\kappa_b \circ \gamma = \beta \,\} \\
&= \{\, (y, a, \lambda_y, S \in \mathrm{BCov}(\lambda_y^*\sigma_A(a))) : \\
&\qquad (\forall \beta \in S)\,(\exists b \in B_a)\,(\exists \gamma \in \mathcal{C}_1)\,(y,b,\gamma) \in \overline{Q} \text{ and } (b, \lambda_y, S, \beta, \gamma) \in B_0 \,\}
\end{aligned}$$

is an *mvs* of $S_0 \to A$ over Y. By the genericity of P_1 this implies the existence of a map $v_1 : U_1 \to Z_1$ and a cover $w_1 : U_1 \to Y$ such that $m_0 v_1 = l w_1$ and $v_1^* P_1 \leq w_1^* Q_1$ as *mvs*s of $S_0 \to A_0$ over U_1. Note that this means that

$$(v_1(u_1), a, \alpha, S) \in P_1 \Longrightarrow \alpha = \lambda_{w_1(u_1)}. \tag{7}$$

Next, define the subobject $Q_2 \subseteq v_1^* B_1$ by saying for any element $(u_1 \in U_1, b \in B, S \in \mathrm{BCov}(\lambda^*_{w_1(u_1)}\sigma_A(kb)), \beta \in S, \gamma \in \mathcal{C}_1) \in v_1^* B_1$:

$$(u_1, b, S, \beta, \gamma) \in Q_2 \Longleftrightarrow (w_1 u_1, b, \gamma) \in \overline{Q}(\mathrm{dom}(\gamma)).$$

It follows from (7) and the definition of Q_1 that Q_2 is a small *mvs* of $B_1 \to M_1$ over U_1. Therefore there is a map $v_2 : U \longrightarrow Z_2$ and a cover $w_2 : U \longrightarrow U_1$ such that $v_1 w_2 = m_1 v_2$ and $v_2^* P_2 \leq w_2^* Q_2$. Note that (7) implies that v_2 factors through $m_2 : Z \to Z_2$ via a map $v : U \to Z$ given by $v(u) = (v_2(u), \lambda_{w_1 w_2(u)})$.

If we put $w = w_1 w_2$, then $lw = l w_1 w_2 = m_0 v_1 w_2 = m_0 m_1 v_2 = m_0 m_1 m_2 v = mv$. Since for each $u \in U$, $\sigma_Z(vu) = \mathrm{dom}(\lambda_{wu}) = \sigma_Y(wu)$, we may put $\sigma_U(u) = \sigma_Z(vu) = \sigma_Y(wu)$ and then $\pi_! w$ and $\pi_! v$ define maps $\pi_! U \to \pi_! Y$ and $\pi_! U \to \pi_! Z$, respectively, such that $(l,\lambda)_! \pi_! w = (m,\mu)_! \pi_! v$. Because $\pi_! w$ is a cover, the proof will be finished, once we show that $(\pi_! v)^* P \leq (\pi_! w)^* \overline{Q}$.

To show this, consider an element $(u \in U, b \in B, \eta : c \to d \in \mathcal{C}) \in \pi_!(U \times_X B)(c)$ for which we have $(u, b, \eta) \in (\pi_! v)^* P(c)$. This means that $(vu, b, \eta) \in P(c)$ and hence that there is a sieve $S \in \mathrm{BCov}(\mu_{vu}^* \sigma_A(kb))$, a map $\beta \in S$ and a map $\gamma : e \to d \in \mathcal{C}_1$ such that $(vu, b, \mu_{vu}, S, \beta, \gamma) \in P_3$ and η factors through γ. The former means that $(v_2 u, b, \mu_{vu}, S, \beta, \gamma) \in P_2$ and since $v_2^* P_2 \leq w_2^* Q_2$, it follows that $(w_2u, b, S, \beta, \gamma) \in Q_2$. By definition this means that $(wu, b, \gamma) \in \overline{Q}(e)$. Since $\overline{Q}$ is a presheaf, also $(wu, b, \eta) \in \overline{Q}(c)$ and hence $(u, b, \eta) \in (\pi_! w)^* \overline{Q}(c)$. This completes the proof. $\square$

Remark 4.12 Again, one can also prove this result using the internal logic of categories of sheaves. Also to illustrate its power, we give one such proof here.

Theorem 4.13 *(Assuming $\mathcal{C}$ has chosen finite products.) The fullness axiom* **(F)** *is inherited by sheaf models.*

Proof. In view of Lemma 2.13 and Corollary 3.8, we only need to build generic *mvs*s for maps of the form $i^*(k,l)_!: \rho_! B \to \rho_! A$ in which k is small, where $\rho_! A$ lies over some object of the form $\rho_! X$ via a map of the form $i^*(r,s)_!$ in which r is small. Again, by replacing $\mathcal{C}$ by $\mathcal{C}/\rho_!(X)$, we may assume that $X = 1 = \{*\}$ and $\sigma_X(*) = 1$.

Note that for a fixed $c \in \mathcal{C}_0$ an *mvs* of $i^*(k,l)_!$ over $i^*\mathcal{C}(-,c)$ as in

$$\begin{tikzcd} P \arrow[r, tail] \arrow[dr, two heads] & \rho_!(B) \times i^*\mathcal{C}(-,c) \arrow[d, "i^*((k,l)_! \times \mathsf{id})"] \\ & \rho_!(A) \times i^*\mathcal{C}(-,c) \end{tikzcd}$$

satisfies

$$\begin{gathered} (\forall a \in A)\,(\exists S \in \mathrm{BCov}(\sigma_A(a) \times c)) \\ (\forall g\colon d \to \sigma_A(a) \times c \in S)\,(\exists b \in B)\,(\exists f\colon d \to \sigma_B(b))\,(\exists h\colon d \to c) \\ (k(b), l_b \circ f) = (a, \pi_1 \circ g), \pi_2 \circ g = h \text{ and } (b, f, h) \in P(d), \end{gathered}$$

or

$$\begin{gathered} (\forall a \in A)\,(\exists S \in \mathrm{BCov}(\sigma_A(a)) \times c) \\ (\forall g\colon d \to \sigma_A(a) \times c \in S)\,(\exists b \in B)\,(\exists f\colon d \to \sigma_B(b)) \\ k(b) = a, l_b \circ f = \pi_1 \circ g \text{ and } (b, f, \pi_2 \circ g) \in P(d). \end{gathered} \tag{8}$$

We first apply fullness in $\mathcal{E}$ to the map

$$\{(a \in A, c \in \mathcal{C}_0, S \in \mathrm{BCov}(\sigma_A(a) \times c))\} \to A\colon (a, c, S) \to a$$

to obtain a generic small family of *mvss* $(Q_j)_{j \in J}$.

Writing for every $j \in J$

$$\begin{aligned} \bar{Q}_j &:= \{(a, c, S, g) : (a, c, S) \in Q_j, g \in S\}, \\ \tilde{Q}_j &:= \{(a, c, S, g, b, f) : (a, c, S) \in Q_j, g \in S, b \in B, k(b) = a, l_b f = \pi_1 g\}, \end{aligned}$$

we have an obvious projection

$$\tilde{Q}_j \to \bar{Q}_j.$$

Applying fullness and using the collection axiom we obtain generic families $\{P_{ij}\}_{i \in I_j}$ for these maps as well. (The collection axiom is employed here to obtain these generic families as a function of j.)

For fixed $c \in \mathcal{C}, j \in J$ and $i \in I_j$ the object P_{ij} determines a subsheaf

$$\hat{P}_{ij}^{(c)} \subseteq \rho_!(B) \times i^*\mathcal{C}(-, c)$$

generated by those elements $(b, f, \pi_2 \circ g)$ for which there is a basic covering sieve S such that $(k(b), c, S, g, b, f) \in P_{ij}$. Again, we only take those which are *mvs*s, i.e., map in a locally surjective manner to $\rho_!(A) \times i^*\mathcal{C}(-, c)$.

Now suppose c_0 is arbitrary and $R \subseteq \rho_!(B) \times \mathcal{C}(-, c_0)$ is an *mvs*. This means that (8) holds with $c = c_0$. Hence there is a Q_j with

$$\begin{array}{rcl} Q_j & \subseteq & \{ (a, c_0, S) \colon (\forall g \colon d \to \sigma_A(a) \times c \in S)\, (\exists b \in B)\, (\forall f \colon d \to \sigma_B(b)) \\ & & k(b) = a, l_b \circ f = \pi_1 \circ g \text{ and } (b, f, \pi_2 \circ g) \in R(d) \} \end{array} \tag{9}$$

and an $i \in I_j$ with

$$P_{ij} \quad \subseteq \quad \{ (a, c_0, S, g, b, f) \colon (b, f, \pi_2 \circ g) \in R(d) \}.$$

It is clear that $\hat{P}_{ij}^{(c_0)} \subseteq R$, so it remains to verify that $\hat{P}_{ij}^{(c_0)}$ is an *mvs*.

We check (8) for $c = c_0$. So take $a \in A$. We want to show that the generator $(a, \pi_1, \pi_2) \in \rho_!(A) \times i^*\mathcal{C}(-, c_0)$ is locally hit by $i^*((k, l)_! \times \mathsf{id})$. Because Q_j is an *mvs*, there are $e \in \mathcal{C}_0$ and $S \in \mathrm{BCov}(\sigma_A(a) \times e)$ such that $(a, e, S) \in Q_j$; morever, we must have $e = c_0$, because (9) holds. Since P_{ij} is an *mvs* we know that for every $g \in S$ there are $b \in B$ and $f \in \mathcal{C}_1$ with $(a, c_0, S, g, b, f) \in P_{ij}$. In particular $(a, c_0, S, g, b, f) \in \tilde{Q}_j$, so $k(b) = a$ and $l_b \circ f = \pi_1 \circ g$. By construction $(b, f, \pi_2 \circ g) \in \hat{P}_{ij}^{(c_0)}$ and for this element the equation

$$i^*((k, l)_! \times \mathsf{id})(b, f, \pi_2 \circ g) = (k(b), l_b \circ f, \pi_2 \circ g) = (a, \pi_1, \pi_2) \cdot g$$

holds. This concludes the proof. □

4.4 W-types in sheaves

In this final subsection, we show that the axiom **(WE)** is inherited by sheaf models. It turns out that the construction of W-types in categories of sheaves is considerably more involved than in the presheaf case (in [10] we showed that some of the complications can be avoided if the metatheory includes the axiom of choice). We then go on to show that the axiom **(WS)** is inherited as well, if we assume the axiom of multiple choice.

Remark 4.14 In [28] the authors claimed that W-types in categories of sheaves are computed as in presheaves (Proposition 5.7 in *loc.cit.*) and can therefore be described in the same (relatively easy) way. But, unfortunately, this claim is incorrect, as the following counterexample shows. Let $F: 1 \to 1$ be the identity map on the terminal object. The W-type associated to F is the initial object, which, in general, is different in categories of presheaves and sheaves. (This was noticed by Peter Lumsdaine together with the first author.)

We fix a small map $F: Y \to X$ of sheaves. If $x \in X(a)$ and S is a covering sieve on a, then we put

$$Y_x^S := \{(f: b \to a \in S, y \in Y(b)) : F(y) = x \cdot f\}.$$

Observe that Y_x^S is small and write ψ for the obvious projection

$$\psi: \sum_{(S,x)} Y_x^S \to X \times_{\mathcal{C}_0} \mathrm{Cov}.$$

Let $\Psi = P_\psi \circ \mathcal{P}_s^+$ and let $\mathcal{V}$ be its initial algebra (see Theorem 2.10). Elements v of $\mathcal{V}$ are therefore of the form $\sup_{(a,x,S)} t$ with $(a, x, S) \in X \times_{\mathcal{C}_0} \mathrm{Cov}$ and $t: Y_x^S \to \mathcal{P}_s^+ \mathcal{V}$. We will think of such an element v as a labelled well-founded tree, with a root labelled with (a, x, S). To this root is attached, for every $(f, y) \in Y_x^S$ and $w \in t(f, y)$, the tree w with an edge labelled with (f, y). To simplify the notation, we will denote by $v(f, y)$ the *small* collection of all trees that are attached to the root of v with an edge that has the label (f, y).

We now wish to define a presheaf structure on $\mathcal{V}$. We say that a tree $v \in \mathcal{V}$ is *rooted* at an object a in $\mathcal{C}$, if its root has a label whose first component is a. If $v = \sup_{(a,x,S)} t$ is rooted at a and $f: b \to a$ is a map in $\mathcal{C}$, then we can define a tree $v \cdot f$ rooted at b, as follows:

$$v \cdot f = \sup_{(b, x \cdot f, f^* S)} f^* t,$$

with

$$(f^* t)(g, y) = t(fg, y).$$

This clearly gives $\mathcal{V}$ the structure of a presheaf. Note that

$$(v \cdot f)(g, y) = v(fg, y).$$

Next, we define by transfinite recursion a relation on $\mathcal{V}$:

$v \sim v' \quad \Leftrightarrow \quad$ if the root of v is labelled with (a, x, S) and the root of v' with (a', x', S'), then $a = a'$, $x = x'$ and there is a covering sieve $R \subseteq S \cap S'$ such that for every $(f, y) \in Y_x^R$ we have $v(f, y) \sim v'(f, y)$.

Here, the formula $v(f, y) \sim v'(f, y)$ is supposed to mean

$$\forall m \in v(f, y), n \in v'(f, y) : m \sim n.$$

In general, we will write $M \sim N$ for small subobjects M and N of $\mathcal{V}$ to mean

$$\forall m \in M, n \in N : m \sim n.$$

In a similar vein, we will write for such a subobject M,

$$M \cdot f = \{m \cdot f : m \in M\}.$$

That the relation $\sim$ is indeed definable can be shown by the methods of [8] or [9]. By transfinite induction one can show that $\sim$ is symmetric and transitive, and compatible with the presheaf structure ($v \sim w \Rightarrow v \cdot f \sim w \cdot f$).

Next, we define *composability* and *naturality* of trees (as we did in the presheaf case, see Theorem 3.3).

- A tree $v \in \mathcal{V}$ whose root is labelled with (a, x, S) is *composable*, if for any $(f : b \to a, y) \in Y_x^S$ and $w \in v(f, y)$, the tree w is rooted at b.
- A tree $v \in \mathcal{V}$ whose root is labelled with (a, x, S) is *natural*, if it is composable and for any $(f : b \to a, y) \in Y_x^S$ and $g : c \to b$,

$$v(f, y) \cdot g \sim v(fg, y \cdot g).$$

One can show that if v is natural, and $v \sim w$, then also w is natural; moreover, natural trees are stable under restriction. The same applies to the trees that are *hereditarily* natural (i.e. not only are they themselves natural, but the same is true for all their subtrees).

We shall write $\mathcal{W}$ for the object consisting of those trees that are hereditarily natural. The relation $\sim$ defines an equivalence on $\mathcal{W}$, for if a tree $v = \sup_{(a,x,S)} t$ is natural, then for all $(f, y) \in Y_x^S$ one has $v(f, y) \cdot \mathsf{id} \sim v(f \cdot \mathsf{id}, y \cdot \mathsf{id})$, that is, $v(f, y) \sim v(f, y)$, and therefore $v \sim v$. By induction one proves that the equivalence relation $\sim$ on $\mathcal{W}$ is bounded and hence a quotient exists. We denote it by $\overline{\mathcal{W}}$. It follows from what we have said that the quotient $\overline{\mathcal{W}}$ is a presheaf, but more is true: one can actually show that $\overline{\mathcal{W}}$ is a sheaf and, indeed, the W-type associated to F in sheaves.

Lemma 4.15 *Let $w, w' \in \mathcal{W}$ be rooted at $a \in \mathcal{C}$. If T is a sieve covering a and $w \cdot f \sim w' \cdot f$ for all $f \in T$, then $w \sim w'$. In other words, $\overline{\mathcal{W}}$ is separated.*

Proof. If the label of the root of w is of the form (a, x, S) and that of w' is of the form (a, x', S'), then $w \cdot f \sim w' \cdot f$ implies that $x \cdot f = x' \cdot f$ for all $f \in T$. As X is separated, it follows that $x = x'$.

Consider

$$R \quad = \quad \{\, g : b \to a \in (S \cap S') : \forall (h, y) \in Y_{x \cdot g}^{M_b} \,[\, w(gh, y) \sim w'(gh, y) \,]\,\}.$$

R is a sieve, and the statement of the lemma will follow once we have shown that it is covering.

Fix an element $f \in T$. That $w \cdot f \sim w' \cdot f$ holds means that there is a covering sieve $R_f \subseteq f^*S \cap f^*S'$ such that for every $(k, y) \in Y_{x \cdot f}^{R_f}$ we have $w(fk, y) = (w \cdot f)(k, y) \sim (w' \cdot f)(k, y) = w'(fk, y)$. In other words, $R_f \subseteq f^*R$. So R is a covering sieve by local character. $\square$

Lemma 4.16 $\overline{\mathcal{W}}$ *is a sheaf.*

Proof. Let S be a covering sieve on a and suppose we have a compatible family of elements $(\overline{w}_f \in \overline{\mathcal{W}})_{f \in S}$. Using the collection axiom, we know that there must be a span

$$\begin{array}{ccccc} S & \leftarrow & J & \rightarrow & \mathcal{W} \\ f_j & \leftarrow\!\shortmid & j & \mapsto & w_j \end{array}$$

with J small and $[w_j] = \overline{w}_{f_j}$ for all $j \in J$. Every w_j is of form $\sup_{(a_j, x_j, R_j)} t_j$. If $f_j = f_{j'}$, then $w_j \sim w_{j'}$, so $x_j = x_{j'}$. Thus the x_j form a compatible family and, since X is a sheaf, can be glued together to obtain an element $x \in X(a)$. We claim that the desired glueing is $[w]$, where $w = \sup_{(a,x,R)} t \in \mathcal{V}$ is defined by:

$$\begin{array}{rcl} R & = & \{f_j g \colon j \in J, g \in R_j\}, \\ t(h, y) & = & \bigcup_{j \in J} \{t_j(g, y) \colon f_j g = h\} \end{array}$$

For this to make sense, we first need to show that $w \in \mathcal{W}$, i.e., that w is hereditarily natural. In order to do this, we prove the following claim.

Claim. Assume we are given $(h, y) \in Y_x^R$, with $h = f_j g$ for some $j \in J$. Then

$$w(h, y) \sim w_j(g, y).$$

Proof. Since

$$w(h, y) = \bigcup_{j' \in J} \{w_{j'}(g', y) \colon f_{j'} g' = h\},$$

it suffices to show that $w_j(g, y) \sim w_{j'}(g', y)$ if $h = f_{j'} g'$.

By compatibility of the family $(\overline{w}_f \in \overline{\mathcal{W}})_{f \in S}$ we know that $w_j \cdot g \sim w_{j'} \cdot g' \in \mathcal{W}(c)$. This means that there is a covering sieve $T \subseteq g^* R_j \cap (g')^* R_{j'}$ such that for all $(k, z) \in Y_{x \cdot h}^T$, we have $(w_j \cdot g)(k, z) \sim (w_{j'} \cdot g')(k, z)$. So if $k \colon d \to c \in T$, then

$$\begin{array}{rcl} w_j(g, y) \cdot k & \sim & w_j(gk, y \cdot k) \\ & = & (w_j \cdot g)(k, y \cdot k) \\ & \sim & (w_{j'} \cdot g')(k, y \cdot k) \\ & = & w_{j'}(g'k, y \cdot k) \\ & \sim & w_{j'}(g', y) \cdot k. \end{array}$$

Because $\overline{\mathcal{W}}$ is separated (as was shown in Lemma 4.15), it follows that $w_j(g, y) \sim w_{j'}(g', y)$. This proves the claim. $\square$

Any subtree of w is a subtree of some w_j and therefore natural. Hence we only need to prove of w itself that it is composable and natural. Direct inspection shows that the tree that we have constructed is composable. For verifying that w is also natural, let $(h{:}\, c \to a, y) \in Y_x^R$ and $k{:}\, d \to c$. Since $h \in R$, there are $j \in J$ and $g \in R_j$ such that $h = f_j g$. Then

$$w(h, y) \cdot k \sim w_j(g, y) \cdot k \sim w_j(gk, y \cdot k) \sim w(hk, y \cdot k),$$

by using naturality of w_j and the claim (twice).

It remains to show that $[w]$ is a glueing of all the $\overline{w}_f$, i.e., that $w \cdot f_j \sim w_j$ for all $j \in J$. So let $j \in J$. First of all, $x \cdot f_j = x_j$, by construction. Secondly, for every $g{:}\, c \to b \in R_j = (R_j \cap f_j^* R)$ and $y \in Y(c)$ such that $F(y) = x \cdot f_j g$, we have

$$(w \cdot f_j)(g, y) = w(f_j g, y) \sim w_j(g, y).$$

This completes the proof. $\square$

Lemma 4.17 *$\overline{\mathcal{W}}$ is a P_F-algebra.*

Proof. We have to describe a natural transformation $S{:}\, P_F \overline{\mathcal{W}} \to \overline{\mathcal{W}}$. An element of $P_F \overline{\mathcal{W}}(a)$ is a pair (x, t) consisting of an element $x \in X(a)$ together with a natural transformation $G{:}\, Y_x^{M_a} \to \overline{\mathcal{W}}$. Using collection, there is a map

$$Y_x^{M_a} \xrightarrow{\ t\ } \mathcal{P}_s^+ \mathcal{W} \tag{10}$$

such that $[w] = G(y, f)$, for all $(f, y) \in Y_x^{M_a}$ and $w \in t(f, y)$. We define $S_x G$ to be

$$[\, \sup_{(a, x, M_a)} t\,].$$

One now needs to check that w is hereditarily natural. And then another verification is needed to check that $[w]$ does not depend on the choice of the map in (10). Finally, one needs to check the naturality of S. These verifications are all relatively straightforward and similar to some of the earlier calculations, and therefore we leave all of them to the reader. $\square$

Lemma 4.18 *$\overline{\mathcal{W}}$ is the initial P_F-algebra.*

Proof. We will show that $S{:}\, P_F \overline{\mathcal{W}} \to \overline{\mathcal{W}}$ is monic and that $\overline{\mathcal{W}}$ has no proper P_F-subalgebras; it will then follow from Theorem 26 of [8] (or Theorem 6.13 in [9]) that $\overline{\mathcal{W}}$ is the W-type of F.

We first show that S is monic. So let $(x, G), (x', G') \in P_F X(a)$ be such that $S_x G = S_{x'} G' \in \overline{\mathcal{W}}$. It follows that $x = x'$ and that there is a covering sieve S on a such that for all $(h, y) \in Y_x^S$, we have $G(h, y) = G'(h, y)$. We need to show that $G = G'$, so let $(f, y) \in Y_x^{M_a}$ be arbitrary. For every $g \in f^*S$, we have:

$$G(f, y) \cdot g = G(fg, y \cdot g) = G'(fg, y \cdot g) = G'(f, y) \cdot g.$$

Since f^*S is covering, it follows that $G(f, y) = G'(f, y)$, as desired.

The fact that $\overline{\mathcal{W}}$ has no proper P_F-subalgebras is a consequence of the inductive properties of $\mathcal{V}$ (recall that $\mathcal{V}$ is an initial algebra). Let $\mathcal{A}$ be a sheaf and P_F-subalgebra of $\overline{\mathcal{W}}$. We claim that

$$\mathcal{B} \quad = \quad \{v \in \mathcal{V} \colon \text{if } v \text{ is hereditarily natural, then } [v] \in \mathcal{A}\}$$

is a subalgebra of $\mathcal{V}$. Proof: Suppose v is a tree that is hereditarily natural. Assume moreover that $v = \sup_{(a,x,S)} t$ and for all $(f, y) \in Y_x^S$ and $w \in t(f, y)$, we know that $[w] \in \mathcal{A}$. Our aim is to show that $[v] \in \mathcal{A}$.

For the moment fix an element $f \colon b \to a \in S$. Since $v \cdot f$ has a root labelled by $(b, x \cdot f, M_b)$ and $(v \cdot f)(g, y) = v(fg, y)$ for all $(g, y) \in Y_{x \cdot f}^{M_b}$, we have that $[v] \cdot f = S_{x \cdot f} G$, where $G(g, y) = [v(fg, y)] \in \mathcal{A}$. Because $\mathcal{A}$ is a P_F-subalgebra of $\overline{\mathcal{W}}$ this implies that $[v] \cdot f \in \mathcal{A}$. Since this holds for every $f \in S$, while S is a covering sieve and $\mathcal{A}$ is a subsheaf of $\overline{\mathcal{W}}$, we obtain that $[v] \in \mathcal{A}$, as desired.

We conclude that $\mathcal{B} = \mathcal{V}$ and hence $\mathcal{A} = \overline{\mathcal{W}}$. This completes the proof. □

To wrap up:

Theorem 4.19 *The axiom* **(WE)** *is inherited by sheaf models.*

We believe that one has to make additional assumptions on ones predicative category with small maps $(\mathcal{E}, \mathcal{S})$ to show that the axiom **(WS)** is inherited by sheaf models (the argument above does not establish this, the problem being that the initial algebra $\mathcal{V}$ will be large, even when the codomain of the map $F \colon Y \to X$ we have computed the W-type of is small). We will now show that this problem can be circumvented if we assume that the axiom of multiple choice **(AMC)** holds in $\mathcal{E}$. It is quite likely that one can also solve this problem by using Aczel's Regular Extension Axiom: it implies the axiom **(WS)** and is claimed to be stable under sheaf extensions (but, as far as we are aware, no proof of that claim has been published).

Theorem 4.20 *The axiom* **(AMC)** *is inherited by sheaf models.*

Proof. This was proved in Section 10 of [29]. □

Theorem 4.21 *(Assuming that* **(AMC)** *holds in $\mathcal{E}$.) The axiom* **(WS)** *is inherited by sheaf models.*

Proof. We will continue to use the notation from the proof of the previous theorem. So, again, we assume we have a small map $F: Y \to X$ of sheaves. Moreover, we let ψ be the map in $\mathcal{E}$ and Ψ be the endofunctor on $\mathcal{E}$ defined above, we let $\mathcal{V}$ be its initial algebra and $\sim$ be the symmetric and transitive relation we defined on $\mathcal{V}$, and $\overline{\mathcal{W}}$ the W-type associated to F, obtained by quotienting the hereditarily natural elements in $\mathcal{V}$ by $\sim$.

Assume that X is a small sheaf. Since **(AMC)** holds in $\mathcal{E}$, it is the case that, internally in $\mathcal{E}/\mathcal{C}_0$, the map ψ fits into a covering square as shown

$$\begin{tikzcd} D \arrow[r, "q"] \arrow[d, "g"'] & \sum_{(S,x)} Y_x^S \arrow[d, "\psi"] \\ C \arrow[r, "p"'] & X \times_{\mathcal{C}_0} \mathrm{Cov}, \end{tikzcd}$$

in which all objects and maps are small in $\mathcal{E}/\mathcal{C}_0$ and (g, q) is a collection span over $X \times_{\mathcal{C}_0} \mathrm{Cov}$. The W-type $\mathcal{U} = W_g$ in $\mathcal{E}/\mathcal{C}_0$ is small in $\mathcal{E}/\mathcal{C}_0$, because we are assuming that **(WS)** holds in $\mathcal{E}$ (and hence also in $\mathcal{E}/\mathcal{C}_0$). The idea is to use this to show that $\overline{\mathcal{W}}$ is small as well.

Every element $u = \mathrm{sup}_c s \in \mathcal{U}$ determines an element in $\varphi(u) \in \mathcal{V}$ as follows: first compute $p(c) = (a, x, S)$. Then let for every $(y, f) \in Y_x^S$ the element $t(y, f)$ be defined by

$$t(y, f) = \{(\varphi \circ s)(d) : d \in q_c^{-1}(y, f)\}.$$

Then $\varphi(u) = \mathrm{sup}_{(a,x,S)} t$ (so this is an inductive definition). We claim that for every hereditarily natural tree $v \in \mathcal{W}$ there is an element $u \in \mathcal{U}$ such that $v \sim \varphi(u)$. The desired result follows readily from this claim.

We prove the claim by induction: so let $v = \mathrm{sup}_{(a,x,S)} t$ be a hereditarily natural element of $\mathcal{V}$ and assume the claim holds for all subtrees of v. Since all subtrees of v are hereditarily natural as well, this means that for every $(y, f) \in Y_x^S$ and $w \in t(y, f)$ there is an element $u \in \mathcal{U}$ such that $\varphi(u) = w$. From the fact that (g, q) is a collection span over $X \times_{\mathcal{C}_0} \mathrm{Cov}$, it follows that there is a $c \in C$ with $p(c) = (a, x, S)$ together with two functions: first one picking for every $d \in D_c$ an element $r(d) \in t(y, f)$ (because $t(y, f)$ is non-empty) and a second one picking for every $d \in D$ an element $s(d) \in \mathcal{U}$ such that $\varphi(s(d)) \sim r(d)$. It is not hard to see that $v \sim \varphi(\mathrm{sup}_c s)$, using that v is natural and therefore all elements in $t(y, f)$ are equivalent to each other. $\square$

This completes the proof of our main result, Theorem 2.18.

5 Sheaf models of constructive set theory

Our main result Theorem 2.18 in combination with Theorem 2.8 yields the existence of sheaf models for **CZF** and **IZF** (see Corollary 2.19). For the sake of completeness and in order to allow a comparison with classical forcing, we describe this model in concrete terms. We will not present verifications of the correctness of our descriptions, because they could in principle be obtained by unwinding the existence proofs, and other descriptions which differ only slightly from what we present here can already be found in the literature.

To construct the initial $\mathcal{P}_s$-algebra in a category of internal presheaves over a predicative category with small maps $(\mathcal{E}, \mathcal{S})$, let $\mathcal{W}$ be the initial algebra of the endofunctor $\Phi = P_{\text{cod}} \circ \mathcal{P}_s$ on $\mathcal{E}$ (see Theorem 2.10). Elements of $w \in \mathcal{W}$ are therefore of the form $\sup_c t$, with $c \in \mathcal{C}_0$ and t a function from $\{f \in \mathcal{C}_1 : \text{cod}(f) = c\}$ to $\mathcal{P}_s\mathcal{W}$. We think of such an element w as a well-founded tree, where the root is labelled with c and for every $v \in t(f)$, the tree v is connected to the root of w with an edge labelled with f. The object $\mathcal{W}$ carries the structure of a presheaf, with $\mathcal{W}(c)$ consisting of trees whose root is labelled with c, and with a restriction operation defined by putting for any $w = \sup_c t$ and $f \colon d \to c$,

$$w \cdot f = \sup_d t(f \circ -).$$

The initial $\mathcal{P}_s$-algebra $\mathcal{V}$ in the category of presheaves is constructed from $\mathcal{W}$ by selecting those trees that are hereditarily composable and natural:

- A tree $w = \sup_c(t) \in \mathcal{W}$ is *composable*, if for any $f \colon d \to c$ and $v \in t(f)$, the tree v has a root labelled with d.
- A tree $w = \sup_c(t) \in \mathcal{W}$ is *natural*, if it is composable and for any $f \colon d \to c$, $g \colon e \to d$ and $v \in t(f)$, we have $v \cdot g \in t(fg)$.

The $\mathcal{P}_s$-algebra structure, or, equivalently, the membership relation on $\mathcal{V}$, is given by the formula $(x, \sup_c t \in \mathcal{V})$

$$x \in \sup_c t \Longleftrightarrow x \in t(\mathsf{id}_c).$$

The easiest way to prove the correctness of the description we gave is by appealing to Theorem 1.1 from [25] (or Theorem 7.3 from [9]). This model was first presented in the paper [18] by Gambino, based on unpublished work by Dana Scott.

The initial $\mathcal{P}_s$-algebra in categories of internal sheaves is obtained as a quotient of this object $\mathcal{V}$. Roughly speaking, we quotient by bisimulation in a way which reflects the semantics of a category of sheaves. More precisely, we take $\mathcal{V}$ as defined above and we write: $\sup_c t \sim \sup_c t'$ if for all $f \colon d \to c$ and $v \in t(f)$, the sieve

$$\{\, g \colon e \to d : \exists v' \in t'(fg)\, (\, v \cdot g \sim v' \,)\,\}$$

covers d and for all $f'\colon d \to c$ and $v' \in t'(f')$, the sieve

$$\{\, g\colon e \to d : \exists v \in t(f'g)\,(\, v' \cdot g \sim v\,)\,\}$$

covers d. On the quotient the membership relation is defined by:

$$[v] \in [\sup_c t] \iff \text{the sieve } \{f\colon d \to c : \exists v' \in t(f)\,(\, v \cdot g \sim v'\,)\} \text{ covers } c.$$

To see that this is correct, one should verify that $\sim$ defines a bounded equivalence relation and the quotient is a sheaf. Then one proves that it is the initial $\mathcal{P}_s$-algebra by appealing to Theorem 1.1 from [25] (or Theorem 7.3 from [9]). The reader who wishes to see more details, should consult [33].

Remark 5.1 To see the analogy with classical forcing (as in [26], for example), note that any poset $\mathbb{P}$ determines a site, by declaring that S covers p whenever S is dense below p. In this case, the elements of $\mathcal{V}$ are a particular kind of *names* (as they are traditionally called). One could regard composability and naturality as saturation properties of names (so that, in effect, we only consider nice, saturated names). It is not too hard to show that every name (in the usual sense) is equal in a forcing model to such a saturated name, so that in the case of classical **ZF** the models that we have constructed are not different from standard forcing models.

References

[1] P. Aczel. The type theoretic interpretation of constructive set theory. In *Logic Colloquium '77 (Proc. Conf., Wrocław, 1977)*, volume 96 of *Stud. Logic Foundations Math.*, pages 55–66. North-Holland, Amsterdam, 1978.

[2] P. Aczel. The type theoretic interpretation of constructive set theory: inductive definitions. In *Logic, methodology and philosophy of science, VII (Salzburg, 1983)*, volume 114 of *Stud. Logic Found. Math.*, pages 17–49. North-Holland Publishing Co., Amsterdam, 1986.

[3] P. Aczel. Aspects of general topology in constructive set theory. *Ann. Pure Appl. Logic*, 137(1-3):3–29, 2006.

[4] P. Aczel and M. Rathjen. Notes on constructive set theory. Technical Report No. 40, Institut Mittag-Leffler, 2000/2001.

[5] S. Awodey. A brief introduction to algebraic set theory. *Bull. Symbolic Logic*, 14(3):281–298, 2008.

[6] S. Awodey, C. Butz, A.K. Simpson, and T. Streicher. Relating first-order set theories and elementary toposes. *Bull. Symbolic Logic*, 13(3):340–358, 2007.

[7] S. Awodey, N. Gambino, P.L. Lumsdaine, and M.A. Warren. Lawvere-Tierney sheaves in algebraic set theory. *J. Symbolic Logic*, 74(3):861–890, 2009.

[8] B. van den Berg. Inductive types and exact completion. *Ann. Pure Appl. Logic*, 134:95–121, 2005.

[9] B. van den Berg and I. Moerdijk. Aspects of predicative algebraic set theory I: Exact Completion. *Ann. Pure Appl. Logic*, 156(1), 2008.

[10] B. van den Berg and I. Moerdijk. W-types in sheaves. arXiv:0810.2398, 2008.

[11] B. van den Berg and I. Moerdijk. A unified approach to algebraic set theory. In *Logic Colloquim 2006*, Lecture Notes in Logic, pages 18–37. Cambridge University Press, Cambridge, 2009.

[12] B. van den Berg and I. Moerdijk. Aspects of predicative algebraic set theory, II: Realizability. *Theoret. Comput. Sci.*, 412(20):1916–1940, 2011.

[13] B. van den Berg and I. Moerdijk. Derived rules for predicative set theory: an application of sheaves. Submitted for publication. Available as arXiv:1009.3553, 2011.

[14] B. van den Berg and I. Moerdijk. A note on the Axiom of Multiple Choice. In preparation, 2011.

[15] M.P. Fourman. Sheaf models for set theory. *J. Pure Appl. Algebra*, 19:91–101, 1980.

[16] P.J. Freyd. The axiom of choice. *J. Pure Appl. Algebra*, 19:103–125, 1980.

[17] N. Gambino. *Sheaf interpretations for generalised predicative intuitionistic systems*. PhD thesis, University of Manchester, 2002.

[18] N. Gambino. Presheaf models for constructive set theories. In *From sets and types to topology and analysis*, volume 48 of *Oxford Logic Guides*, pages 62–77. Oxford University Press, Oxford, 2005.

[19] N. Gambino. Heyting-valued interpretations for constructive set theory. *Ann. Pure Appl. Logic*, 137(1-3):164–188, 2006.

[20] N. Gambino. The associated sheaf functor theorem in algebraic set theory. *Ann. Pure Appl. Logic*, 156(1):68–77, 2008.

[21] R.J. Grayson. Forcing in intuitionistic systems without power-set. *J. Symbolic Logic*, 48(3):670–682, 1983.

[22] P.T. Johnstone. *Sketches of an elephant: a topos theory compendium. Volume 1*, volume 43 of *Oxf. Logic Guides.* Oxford University Press, New York, 2002.

[23] P.T. Johnstone. *Sketches of an elephant: a topos theory compendium. Volume 2*, volume 44 of *Oxf. Logic Guides*. Oxford University Press, Oxford, 2002.

[24] A. Joyal and I. Moerdijk. *Algebraic set theory*, volume 220 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1995.

[25] C. Kouwenhoven-Gentil and J. van Oosten. Algebraic set theory and the effective topos. *J. Symbolic Logic*, 70(3):879–890, 2005.

[26] K. Kunen. *Set theory – An introduction to independence proofs*, volume 102 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1980.

[27] S. Mac Lane and I. Moerdijk. *Sheaves in geometry and logic – A first introduction to topos theory*. Universitext. Springer-Verlag, New York, 1992.

[28] I. Moerdijk and E. Palmgren. Wellfounded trees in categories. *Ann. Pure Appl. Logic*, 104(1-3):189–218, 2000.

[29] I. Moerdijk and E. Palmgren. Type theories, toposes and constructive set theory: predicative aspects of AST. *Ann. Pure Appl. Logic*, 114(1-3):155–201, 2002.

[30] M. Rathjen. Realizability for constructive Zermelo-Fraenkel set theory. In *Logic Colloquium '03*, volume 24 of *Lect. Notes Log.*, pages 282–314. Assoc. Symbol. Logic, La Jolla, CA, 2006.

[31] A.K. Simpson. Elementary axioms for categories of classes (extended abstract). In *14th Symposium on Logic in Computer Science (Trento, 1999)*, pages 77–85. IEEE Computer Soc., Los Alamitos, CA, 1999.

[32] A.K. Simpson. Constructive set theories and their category-theoretic models. In *From sets and types to topology and analysis*, volume 48 of *Oxford Logic Guides*, pages 41–61. Oxford University Press, Oxford, 2005.

[33] T. Streicher. Forcing for IZF in sheaf toposes. *Georgian Mathematical Journal*, 16(1):203–209, 2009.

[34] M.A. Warren. Coalgebras in a category of classes. *Ann. Pure Appl. Logic*, 146(1):60–71, 2007.