

# NORMALIZATION FOR MULTIMODAL TYPE THEORY

DANIEL GRATZER

ABSTRACT. We consider the conversion problem for multimodal type theory (MTT) [Gra+20a] by characterizing the normal forms of the type theory and proving normalization (Theorem 8.5). Normalization follows from a novel adaptation of Sterling’s Synthetic Tait Computability [Ste21; SH20] which generalizes the framework to accommodate a type theory with modalities and multiple modes. As a corollary of our main result, we reduce the conversion problem of MTT to the conversion problem of its mode theory (Corollary 8.8) and show the injectivity of type constructors (Corollary 8.9).

## CONTENTS

1. Introduction	1
2. Normalization via Synthetic Tait Computability	5
3. MTT Cosmoi	7
4. Foundations of multimodal Synthetic Tait Computability	10
5. The category of renamings for MTT	15
6. Prerequisites for the normalization model	17
7. The normalization model	20
8. The normalization function	25
Acknowledgments	28
Appendix A. Neutral and normal forms	29
Appendix B. A full definition of an internal MTT cosmos	31
Appendix C. Neutral and normal forms, internally	33
References	36

## 1. INTRODUCTION

The last twenty years of development of type theory has seen many different extensions of type theories to account for different modalities. Most of these type theories are specialized—accounting for a specific collection of modalities—and intended for a particular model. This specialization has allowed for concise and practicable syntax in some cases, but has also created a tremendous amount of churn where even the smallest change to the modal situation requires substantial work to produce a new type theory.

In response to this situation, several “frameworks” for modal type theory have been put forward. A framework for modal type theory should enjoy the following:

- (1) sufficient generality to accommodate a large class of modalities,
- (2) practicable syntax in specific applications,

---

*Date:* Wednesday 1<sup>st</sup> September, 2021.

(3) a wide variety of theorems which hold for instantiations.

There is a degree of tension between these three goals: a framework which is more general often admits more cumbersome syntax, and wider classes of type theories typically enjoy fewer shared properties. Recently, Gratzner, Kavvos, Nuyts, and Birkedal [Gra+20b] introduced MTT, a type theory parameterized by a collection of modalities. MTT aims to support a highly usable syntax for any collection of modalities which behave like (dependent) right adjoints [Bir+20]. This proves to be a rich class of modalities, allowing encodings of type theories for guarded recursion, internal parametricity, axiomatic cohesion, etc.

**1.1. MTT.** We briefly recall some of details of MTT. For a full account, see the extended submission on MTT [Gra+20a]. As mentioned, MTT is parameterized by a collection of modalities. Formally, MTT is defined over a *mode theory* [LS16]: a strict 2-category whose objects represent modes, while 1-cells determine modalities and 2-cells natural transformations between them.

Given a mode theory  $\mathcal{M}$ , each mode  $m : \mathcal{M}$  determines a mode in MTT. Each mode behaves somewhat like its own type theory. There are separate judgments for each mode  $\Gamma \text{ cx } @ m$ ,  $\Gamma \vdash A @ m$ , and  $\Gamma \vdash M : A @ m$  and types in mode  $m$  are closed under the standard connectives in Martin-Löf type theory (dependent products, sums, booleans, an intensional identity type, and a universe).

Different modes only interact through modalities. Given a morphism  $\mu : n \rightarrow m$ , there is an induced modality  $\langle \mu \mid - \rangle$  in MTT which sends types from mode  $n$  to mode  $m$ . This modality behaves like a right adjoint [Bir+20], whose left adjoint is given an action only on contexts  $-\cdot\{\mu\}$ .<sup>1</sup> This operation is a proper functor between context categories, with an induced action on substitutions  $\gamma.\{\mu\}$ . As the left adjoint,  $-\cdot\{\mu\}$  sends context from mode  $m$  to mode  $n$ . This left adjoint is used to give the formation and introduction rules for MTT by transposition:

$$\frac{\Gamma.\{\mu\} \vdash A @ n}{\Gamma \vdash \langle \mu \mid A \rangle @ m} \qquad \frac{\Gamma.\{\mu\} \vdash M : A @ n}{\Gamma \vdash \text{mod}_\mu(M) : \langle \mu \mid A \rangle @ m}$$

Two questions remain: how does one eliminate an element of  $\langle \mu \mid A \rangle$  and how can one use a context of the form  $\Gamma.\{\mu\}$ . While related modal type theories solve both issues through defining the elimination rule to be transposition in the opposite direction, this causes issues for the substitution property. Instead, MTT employs an alternative approach. First, variables in the context is annotated with a modality:

$$\frac{\Gamma \text{ cx } @ m \quad \Gamma.\{\mu\} \vdash A @ n}{\Gamma.(\mu \mid A) \text{ cx } @ m}$$

These modal annotations are used to determine when a variable can be accessed; a variable annotated with  $\mu$  can be accessed precisely when it is behind  $-\cdot\{\mu\}$ :

$$\frac{}{\Gamma.(\mu \mid A).\{\mu\} \vdash \mathbf{v}_0 : A[\uparrow.\{\mu\}] @ m}$$

Next, the elimination rule for  $\langle \mu \mid - \rangle$  is used to bridge the gap between a variable annotated with  $\nu \circ \mu$  of type  $A$  and a variable annotated with  $\nu$  of type  $\langle \mu \mid A \rangle$ . There is a canonical substitution  $\Gamma.(\nu \circ \mu \mid A) \rightarrow \Gamma.(\nu \mid \langle \mu \mid A \rangle)$ , but a priori no map in the reverse direction. While we do not provide an inverse outright, the elimination

<sup>1</sup>Other presentations of MTT and Fitch-style modal type theories write  $-\cdot\mathbf{a}_\mu$  for this action, but we find  $-\cdot\{\mu\}$  more uniform and less cumbersome.

rule ensures that these two contexts are isomorphic “from the perspective of a type”. Semantically, one ensures that display maps are weakly right orthogonal to this map, but more concretely we add the following elimination principle:

$$\frac{\begin{array}{c} \nu : o \rightarrow n \quad \mu : n \rightarrow m \\ \Gamma \text{ cx } @ m \quad \Gamma.\{\mu\}.\{\nu\} \vdash A @ o \quad \Gamma.(\mu \mid \langle \nu \mid A \rangle) \vdash B @ m \\ \Gamma.\{\mu\} \vdash M_0 : \langle \nu \mid A \rangle @ n \quad \Gamma.(\mu \circ \nu \mid A) \vdash M_1 : B[\uparrow.\text{mod}_\nu(\mathbf{v}_0)] @ m \end{array}}{\Gamma \vdash \text{let}_\mu \text{ mod}_\nu(\_) \leftarrow M_0 \text{ in } M_1 : B[\text{id}.M_0] @ m}$$

Finally, the interactions between modalities are codified by organizing the collection of functors  $-\cdot\{-\}$  into a single 2-functor out of  $\mathcal{M}^{\text{coop}}$ . Concretely, this ensures that there is a substitution  $\Gamma.\{\mu\} \vdash \{\alpha\}_\Gamma : \Gamma.\{\nu\} @ m$  for each 2-cell  $\alpha : \nu \rightarrow \mu$  and that a variety of functoriality equations hold e.g.:

$$\Gamma.\{\mu \circ \nu\} = \Gamma.\{\mu\}.\{\nu\} \text{ cx } @ o \quad \Gamma.\{\text{id}_m\} = \Gamma \text{ cx } @ m$$

2-functoriality is sufficient to ensure that the modal types organize into a weak 2-functor, with type-theoretic equivalences between e.g.,  $\langle \text{id}_m \mid A \rangle$  and  $A$ .

**1.2. Normalization for MTT.** Returning to the criterion for a framework for modal dependent type theories, Gratzer, Kavvos, Nuyts, and Birkedal [Gra+20a] showed that MTT enjoys substantial metatheoretic properties. In particular, every instantiation of MTT is sound and enjoys a canonicity result. Both of these results are proven semantically, taking advantage of the fact that an instantiation of MTT comes equipped with a well-behaved class of models in which syntax is initial. Notably, however, they stopped short of showing that MTT enjoys a normalization result, and the lack of such a result precludes a general approach to implementing instantiations of MTT.

Here, we rectify the situation and show that MTT admits a normalization result. As a corollary, conversion in MTT is decidable when equality of modalities and 2-cells is decidable. This result paves the way for an implementation of MTT which does not necessitate a full reimplementaion for each tweak of the modalities.

Normalization results for dependent type theories are typically involved affairs, requiring the construction of a sophisticated PER model, families of logical relations, and other technical constructions. For this proof, however, we adapt recent gluing arguments for normalization to give a concise and detailed proof [AHS95; AK16; Coq19; Fio02; Ste21].

While the literature contains a large number of modal dependent type theories, very few are proven to enjoy normalization. Of these, the most closely related is MLTT<sub>id</sub> [GSB19a]. This is a type theory extending MLTT with an idempotent comonad. However, compared to the normalization proof of MLTT<sub>id</sub> given by Gratzer, Sterling, and Birkedal [GSB19a], this result is significantly shorter and more conceptual.<sup>2</sup> Furthermore, the result for MTT applies to a wide variety of modal situations, including modal type theory with an idempotent comonad.

**1.3. Proof outline.** In broad strokes, the proof of normalization for MTT proceeds in several stages. First, we introduce the technical device of MTT cosmoi in Section 3. An MTT cosmos is a more flexible version of the models of MTT originally described. In particular, cosmoi are only required to be locally Cartesian closed, so

<sup>2</sup>As a crude measurement, the proof of normalization for MLTT<sub>id</sub> occupies the bulk of the accompanying 90 page technical report [GSB19b]. The proof of normalization for MTT, by contrast, takes approximately 25 pages to present with more detail.

that there is no ‘fiberwise representability’ requirement, and morphisms between models are only required to preserve several structures up to isomorphism. This definition of model is more closely related to the presentations of type theories as *representable map categories* [Uem19] or *locally Cartesian closed categories* [GS20]. This flexibility is essential for the application of synthetic Tait computability and allows us to avoid tedious calculations.

*Remark 1.* We emphasize, however, this change of categories is a technical detail of the proof: the normalization function is defined on the stricter syntax constituting an initial object of the CwF models as originally defined rather than the initial object of the category of cosmoi. While it is more work to obtain a normalization result of this form, it ensures that one does not need to accept the initial MTT cosmos as syntax; such acceptance would require a leap of faith, because there is no adequacy result for MTT cosmoi corresponding those proven by Uemura [Uem19] and Gratzer and Sterling [GS20]. Such an adequacy result is difficult to formulate for MTT cosmoi because of the lack of established modal logical frameworks.

In [Section 4](#) we introduce the necessary mathematics to generalize synthetic Tait computability (STC) from working internally to a single gluing category to applying to a collection of gluing categories interconnected by functors and natural transformations. In particular, we show that one may work internally to this network of categories in MTT, using MTT modalities to pass between gluing categories. We also show that the fibered modalities available in each gluing category commute with modalities already present in MTT, a crucial result for constructing the normalization model.

With this framework in place, [Section 5](#) and [Section 6](#) define the basic objects of the normalization model: renamings, normal and neutral forms, etc. In particular, [Section 6](#) defines the gluing categories used for normalization, along with the functors between them. This portion of the proof proceeds with minimal alternations from a standard proof of normalization for MLTT. The only changes are to account for multiple modalities, and this change is mostly one of book-keeping. One novelty of [Section 5](#) is the usage of a non-free presentation of renamings. By allowing renamings to enjoy non-trivial equations, we vastly simplify the development. As a consequence of this decision, we only must show that the action of a renaming on a normal form respects these equations — a straightforward calculation.

The actual construction of the normalization model, and the heart of the proof, takes place in [Section 7](#). In particular, we define an MTT cosmos of glued categories together with the reify and reflect maps needed to define the normalization function. Those familiar with proofs of normalization by gluing, and in particular with those given by STC will find much of the construction familiar. The central challenge is the construction of modal types in the normalization model ([Lemma 7.6](#)), and in particular the construction of the modal elimination law. Other connectives which do not interact with modalities (booleans, universes, dependent sums, etc.) can be constructed in this multimodal setting with no alteration.

Finally, the normalization function is extracted from this model in [Section 8](#). Taking advantage of the fact that this proof is constructive, we also observe that the normalization function is effective. This uses a novel approach to derive a normalization result for strict syntax, rather than the initial MTT cosmos. Using these results, we establish several other corollaries such as the decidability of equality and the injectivity of type-constructors.

## 2. NORMALIZATION VIA SYNTHETIC TAIT COMPUTABILITY

The proof of normalization for MTT follows from a systematic generalization of Synthetic Tait Computability (STC) to multimodal type theories. While a full introduction of the theory of STC is beyond the scope of this work, given that a comprehensive account has not yet emerged<sup>3</sup> we give a short summary of the key ideas here. We focus in particular on proving normalization via STC. Other descriptions of STC are given by Sterling and Harper [SH20], Sterling and Gratzer [SG20], and Sterling and Angiuli [SA21].

Prior to discussing STC, we remark that in the process of embracing STC we are naturally led to a number of ideas which are not intrinsically categorical but still crucial to the concision of the proof.

- Our normalization proof is *reduction-free*, and
- the algorithm works only over equivalences classes of well-typed terms.

More explicitly, our proof is reduction-free in that it does not proceed by fixing a rewriting system presenting the equational theory which we then prove to be confluent and strongly normalizing. This avoids the thorny issue of finding an rewriting system which faithfully presents the equational theory of MTT—the inclusion of  $\eta$ -laws for dependent sums greatly complicates such a task.

To the second point, working with terms only up to definitional equality obviates the need for a separate proof of *completeness* of the normalization algorithm. More than this, it is also a necessary step to treating MTT categorically; without this quotienting, the category of contexts is not even a category. Similarly, the universal properties of various connectives of MTT are crucially leveraged to simplify aspects of the normalization proof, but these universal properties only come into being after taking terms up to definitional equality.

**2.1. Normalization by gluing.** Proving normalization via STC involves working internally to a particular glued category. This idea branches off the observation that that logical relations arguments could be systematically recast as the construction of a model of type theory in a *Freyd cover* or *Sierpinski cone (scone)* [MS93]. This is a special case of the more general gluing construction:

**Definition 2.1.** Given a functor  $F : \mathcal{C} \rightarrow \mathcal{D}$ , the gluing category  $\mathbf{Gl}(F)$  is the comma category  $(\mathrm{id}_{\mathcal{D}} \downarrow F)$ . Explicitly, objects are triples  $(D, C, f : D \rightarrow F(C))$  and morphisms are commuting squares between them.

**Definition 2.2.** Given a category  $\mathcal{C}$ , The Sierpinski cone  $\mathbf{S}(\mathcal{C})$  is  $\mathbf{Gl}([1_{\mathcal{C}}, -]_{\mathcal{C}})$ .

To a rough approximation, a logical relation can be seen as an assignment of types to predicates on their closed terms. Taking  $\mathbb{T}$  to be the category of contexts and substitutions between them and blurring the distinction between contexts and types as is often done in the simply-typed case, this is more or less the content of an object of  $\mathbf{S}(\mathbb{T})$ ; a triple  $(S, A, f : S \rightarrow [1, A])$  can be viewed as a predicate on the closed elements of  $A$  with  $P(a) = f^{-1}(a)$ . Notice that the correspondence is imperfect because there may be multiple distinct elements in the fiber  $P(a)$ . In fact,  $f^{-1}(-)$  determines a *proof-relevant* predicate on closed elements of  $A$ . This can be fixed by cutting  $\mathbf{S}(\mathbb{T})$  down to consist of objects  $(S, A, f : S \hookrightarrow [1, A])$ , but this is both less categorically natural and, as we shall see in later, counterproductive.

<sup>3</sup>We hope that the forthcoming work of Sterling [Ste21] will serve this role

What makes this shift in perspective so useful is the remarkable ability of  $\mathbf{S}(\mathcal{C})$  to seemingly inherit all the structure of  $\mathcal{C}$  in such a way that the projection  $\mathbf{S}(\mathcal{C}) \rightarrow \mathcal{C}$  preserves this structure. For instance, if  $\mathcal{C}$  is Cartesian closed, then so too is  $\mathbf{S}(\mathcal{C})$  and the projection preserves all this structure. In fact, an explicit construction of e.g., the exponential in  $\mathbf{S}(\mathcal{C})$  yields almost precisely the standard construction of the logical relation at function type. The benefit of relying on such technology, however, is that one can typically avoid these explicit calculations; once it is known that  $\mathbf{S}(\mathcal{C})$  is Cartesian closed, canonicity at ground type can be proven without ever needing to explicitly calculate the construction of exponentials.

The correspondence between logical relations and gluing was pushed further in multiple directions [Alt+01; AHS95; Fio02; Str98]. In particular, the insistence of logical relations of *closed terms* of a type can be relaxed by switching to a different comma category. We are interested in proving normalization, emphatically not a result solely concerned with closed terms.

Accordingly, rather than considering a model in  $\mathbf{S}(\mathcal{C})$ , we will consider a glued category valued in presheaves over renamings. It is perhaps more natural to expect one to examine presheaves over contexts: after all, terms organize into a presheaf over arbitrary contexts and substitutions, with substitution giving rise to the reindexing action of the presheaf. One would then consider a model in  $\mathbf{Gl}(\mathbf{y})$ ; the category given by gluing along the Yoneda embedding.

The first attempts at categorical proofs of normalization constructed models in  $\mathbf{Gl}(\mathbf{y})$ , so that one considered predicates not on the closed terms of type  $A$ , but over the family  $([\Gamma, A])_\Gamma$ . The presheaf condition corresponds to the monotonicity conditions familiar to all Kripke logical relations, but in particular those for normalization. Unfortunately, these presheaf conditions are too onerous to view normal forms as a presheaf on contexts — normal forms are never stable under all substitutions — so that the resultant normalization function is given as a map from terms to terms and indistinguishable from the identity function on terms.

This can be rectified taking presheaves over the wide subcategory of renamings  $\mathcal{R} \subseteq \mathbb{T}$ . By cutting down the substitutions allowed between contexts, normal forms may be organized into an object of  $\mathbf{PSh}(\mathcal{R})$  [AHS95; Fio02]. Accordingly, normalization proceeds not by gluing along the Yoneda lemma as still repeated in literature, but by gluing along a restriction of the Yoneda lemma, the nerve  $\mathbf{n} : \mathbb{T} \rightarrow \mathbf{PSh}(\mathcal{R})$ . Unlike canonicity, it is insufficient to simply construct a model of  $\mathbb{T}$  in  $\mathbf{Gl}(\mathbf{n})$ , extra structure is required which ‘sandwiches’ the predicate for a type  $A$  between the neutral and normal forms of  $A$ , so that the following commutative diagram exists in  $\mathbf{PSh}(\mathcal{R})$ :

$$(1) \quad \begin{array}{ccccc} & & \uparrow_A & & \downarrow_A \\ & & & & \\ \text{Ne}_A & \xrightarrow{\quad} & P_A & \xrightarrow{\quad} & \text{Nf}_A \\ & \searrow & \downarrow & \swarrow & \\ & & \mathbf{n}(A) & & \end{array}$$

The top two maps of this diagram are categorical realizations of the *reflect* and *reify* maps found in normalization-by-evaluation [Abe13] and are used to extract the normalization function. In the simply-typed case, these maps can be constructed by induction after the fact, but in the dependently-typed setting this approach does not scale and the two must be constructed simultaneously. Accordingly, we require

a type in our glued model to be not just a proof-relevant predicate  $P_A$  over  $\mathbf{n}(A)$ , but a triple of  $(P_A, \uparrow_A, \downarrow_A)$  [AHS95; Fio02].

One final remark is needed prior to the introduction of STC. Thus far we have limited consideration to simply-typed languages. The full utility of proof-relevance only becomes apparent when considering dependent type theory with universes. In these cases, it becomes necessary to associate with an element of a universe not merely a proposition (computable or not computable) but to instead pair an element with the *computability data of a type* again. This is beautifully handled by the proof-relevance of predicates: the predicate on universes at  $A : \mathbf{U}$  contains choices of computability data over  $\mathbf{El}(A)$  as witnesses [Coq19; Shu15].

**2.2. Synthetic Tait Computability.** Finally, we now turn to Synthetic Tait Computability. STC builds on the idea of working with glued categories, but instead of gluing together locally Cartesian closed categories presenting dependent type theories [GS20], STC is concerned with topoi. Suppose  $\mathbb{T}$  presents a type theory, and  $\mathcal{R} \subseteq \mathbb{T}$  is the subcategory of renamings. Rather than gluing together the locally Cartesian closed categories  $\mathbb{T}$  and  $\mathbf{PSh}(\mathcal{R})$  along  $\mathbf{n}$ , we instead glue together  $\mathbf{PSh}(\mathbb{T})$  and  $\mathbf{PSh}(\mathcal{R})$  along the inverse image of the essentially geometric morphism induced by the inclusion  $i : \mathcal{R} \rightarrow \mathbb{T}$ . As topoi are closed under gluing, the resultant glued category  $\mathbf{Gl}(i)$  is a topos. In fact, by a standard result [AGV72; CJ95] it is a presheaf topos.

As a result,  $\mathbf{Gl}(i)$  supports a model of extensional Martin-Löf type theory with a hierarchy of cumulative universes. Inside this type theory, two fibered modalities,  $\circ$  and  $\bullet$ , allow us to recover  $\mathbf{PSh}(\mathbb{T})$  and  $\mathbf{PSh}(\mathcal{R})$  inside of  $\mathbf{Gl}(i)$ . By relaxing the definition of a model of Martin-Löf type theory to non-representability universes moreover, we may specify and construct the glued model purely internally to  $\mathbf{Gl}(i)$ .

This internal approach offers a significant improvement over previous proofs of gluing for dependent type theory [Coq19; KHS19] which work externally and incur what Sterling [Ste21] has referred to as an ‘avalanche’ of naturality obligations. As we shall see in Section 7, these tedious calculations are replaced with programming exercises in extensional type theory (or MTT in our case). In particular, a model lying over the model in  $\mathbf{PSh}(\mathbb{T})$  corresponds to a series of constants subject to the requirement that they are sent by  $\circ$  to the constants defining the model in  $\mathbf{PSh}(\mathbb{T})$ .

We refer to the internal type theory of  $\mathbf{Gl}(i)$ —extensional type theory with cumulative universes, two fibered modalities, and a few axioms governing their behavior—as the language of Synthetic Tait Computability.

### 3. MTT COSMOI

While the basic theory of models of MTT is detailed by Gratzer, Kavvos, Nuyts, and Birkedal [Gra+20a], we generalize to locally Cartesian closed categories. This change, inspired by Gratzer and Sterling [GS20], is crucial to actually carrying out the normalization theorem as it removes *representability requirements* from the universes of types and terms. These prove superfluous for the theorem, and forcing the glued model to satisfy them is both unnatural and surprisingly challenging.

*Remark 2.* Henceforth, we shall consider MTT over a fixed mode theory  $\mathcal{M}$ .

**Definition 3.1** (Definition 18 [Awo18]). An internal lifting structure  $s : i \dashv \tau$  between a pair of morphisms  $i : A \rightarrow B$  and  $\tau : X \rightarrow Y$  is a section of canonical map  $X^B \rightarrow Y^B \times_{Y^A} X^A$ .



**Definition 3.2.** The 2-category  $\mathbf{Cat}_g$  consists of small categories, functors between them, and invertible natural transformations.

**Definition 3.3.** The locally full 2-subcategory  $\mathcal{V}$  of  $[\mathcal{M}, \mathbf{Cat}_g]$  is spanned by pseudofunctors  $F$  enjoying the following two properties:

- For each  $m : \mathcal{M}$ ,  $F(m)$  is a locally Cartesian closed category.
- For each  $\mu : n \rightarrow m$ ,  $F(\mu)$  is a right adjoint.

Note that we do not require that  $F(\mu)$  preserves the locally Cartesian closed structure of  $F(m)$ . We write  $F(\mu)_!$  for the left adjoint of  $F(\mu)$ .

**Definition 3.4.** An object  $F : \mathcal{V}$  is an MTT cosmos when equipped with the following structure:

- (1) In  $F(m)$ , there is a universe  $\tau_m : \dot{\mathcal{T}}_m \rightarrow \mathcal{T}_m$  with a choice of codes witnessing its closure under dependent sums, and booleans. Additionally, there is a choice of code making  $\tau_m$  closed under modal dependent products: a code for  $\mathbf{P}_{f_\mu(\tau_n)}(\tau_m)$  for each  $\mu : n \rightarrow m$ .
- (2) For each  $\mu$ , there exists a chosen commuting square

$$(2) \quad \begin{array}{ccc} F(\mu)(\dot{\mathcal{T}}_n) & \xrightarrow{\quad} & \dot{\mathcal{T}}_m \\ \downarrow & & \downarrow \\ F(\mu)(\mathcal{T}_n) & \xrightarrow{\quad \mathbf{Mod} \quad} & \mathcal{T}_m \end{array}$$

- (3) For each  $\mu : n \rightarrow m$  and  $\nu : o \rightarrow n$ , there is a chosen lifting structure  $F(\mu)(m) \dashv F(\mu \circ \nu)(\mathcal{T}_o) \times \tau_m$ , where  $m : F(\nu)(\dot{\mathcal{T}}_o) \rightarrow F(\nu)(\mathcal{T}_o) \times_{\tau_n} \dot{\mathcal{T}}_n$  is the comparison map induced by [Diagram 2](#).
- (4)  $\tau_\mu$  contains a subuniverse also closed under all these connectives.

To a first approximation, an MTT cosmos is a standard model of MTT formulated in the language in natural models without the requirements that the universes be legitimate representable natural transformations. By expanding the class of models in this way, we make it far easier to construct the glued model in [Section 7](#).

*Remark 3.* In order to simplify matters, we assume MTT has only a weak universe à la Tarski. Concretely, there exists a universe  $\mathbf{U}$  and a decoding operation  $\mathbf{El}(-)$  just as in Gratzer, Kavvos, Nuyts, and Birkedal [[Gra+20a](#)]. Unlike the original presentation of MTT however, we only obtain a chosen isomorphism between each connective and its decoded code e.g.  $\mathbf{dec}^\mathcal{P}(-) : \mathbf{El}((\mu \mid A) \rhd B) \cong (\mu \mid \mathbf{El}(A)) \rightarrow \mathbf{El}(B)$ .

Recent unpublished work by Gratzer and Sterling has shown that Synthetic Tait Computability [[Ste21](#)] can be strengthened to accommodate stricter universes, and can be applied to the following without issue. We remark, however, that there is experimental evidence that elaboration can be used to alleviate the tedium of these weaker universes in practice [[Red20](#)].

**Definition 3.5.** A morphism between MTT-structured cosmoi  $\alpha : F \rightarrow G$  is a 2-natural transformation  $\alpha$  such that  $\alpha_m$  is an LCCC functor and preserves all connectives strictly. Furthermore, we require that there is a natural isomorphism



$\beta_\mu : \alpha_n \circ F(\mu)_! \cong G(\mu)_! \circ \alpha_m$  such that the transposition of  $\alpha_\mu \circ \alpha_m(a)$  is  $\alpha_m(\hat{a}) \circ \beta_\mu$  when  $a : X \rightarrow F(\mu)(Y) : F(m)$ .

**Theorem 3.6.** *Any strict model of MTT induces a cosmos. In particular, the initial model of MTT  $\mathcal{S}[-]$  induces a cosmos.*

**Theorem 3.7** (Quasi-projectivity). *Given an arbitrary MTT cosmos  $G$  and a map  $\pi : G \rightarrow \mathcal{S}[-]$ , the following holds:*

- (1) *For every context  $\Gamma \mathbf{cx} @ m$ , there exists an object  $[\Gamma] : G(m)$ , together with a canonical isomorphism  $\alpha_\Gamma : \pi([\Gamma]) \cong \mathbf{y}(\Gamma)$ .*
- (2) *For every type  $\Gamma \vdash A @ m$ , there is a morphism  $[A] : [\Gamma] \rightarrow \mathcal{T}_m$  such that  $\pi([A]) \circ \alpha_\Gamma = [A]$ .*
- (3) *For every term  $\Gamma \vdash M : A @ m$ , there is a morphism  $[M] : [\Gamma] \rightarrow \dot{\mathcal{T}}_m$  lying over  $[A]$  such that  $\pi([M]) \circ \alpha_\Gamma = [M]$ .*

*Remark 4.* We have used ‘syntactic’ notation ( $\Gamma \mathbf{cx} @ m$ ,  $\Gamma \vdash M : A @ m$ , etc.) to denote the components of the initial model  $\mathcal{S}[-]$ . This is a mere notational convenience: we do not rely on the fact that  $\mathcal{S}[-]$  may be constructed out of ‘traditional’ syntax.

*Proof.* For clarity, we write  $\mathbf{El}_m$ ,  $\mathbf{Ty}_m$  and  $\mathbf{Tm}_m$  instead of  $\tau_m$ ,  $\mathcal{T}_m$ , and  $\tau_m$  in the syntactic model, reserving the latter exclusively for  $G$ . We write  $[\mu]$  for the functor sending  $\Gamma$  to  $\Gamma.\{\mu\}$ . We may replace  $G$  by an equivalent strict 2-functor such that  $\pi$  becomes strictly 2-natural, so we work under this assumption.

We construct a displayed model of MTT [KKA19] which lies over the syntactic model. Using the existing coherence result for MTT, we only ensure that  $\Gamma.\{\mu\}.\{\nu\}$  and  $\Gamma.\{\mu \circ \nu\}$  agree up to pseudonatural isomorphism.

- Contexts in mode  $m$  are interpreted by triples  $(X : G(m), \Gamma, \alpha : \pi(X) \cong \mathbf{y}(\Gamma))$ .
- A type in a context  $(X, \Gamma, \alpha)$  is a pair of  $(\bar{A} : X \rightarrow \mathcal{T}_m, \Gamma \vdash A @ m)$  such that  $\pi(\bar{A}) = [A] \circ \alpha$ .
- A term in a context  $(X, \Gamma, \alpha)$  of type  $(A^*, A)$  is a pair of  $M^* : X \rightarrow \dot{\mathcal{T}}_m$  and  $\Gamma \vdash M : A @ m$  such that  $\tau_m \circ M^* = A^*$  and  $\pi(M^*) = [M] \circ \alpha$ .
- A substitution between  $(X, \Gamma, \alpha)$  and  $(Y, \Delta, \beta)$  is a pair of  $f : X \rightarrow Y$  and  $\Gamma \vdash \delta : \Delta @ m$  such that  $\beta \circ \pi(f) = \mathbf{y}(\delta) \circ \alpha$

Once this model is constructed, the result is a direct consequence of the initiality of (strict) syntax. In fact, the construction of contexts, substitutions, terms, and types is relatively routine owing to the fact that  $\pi$  is a 2-natural transformation, preserves finite limits, and strictly commutes with all connectives. We show a few cases to give a flavor for the procedure.

*The action of a modality on a context.* Given a triple  $(X, \Gamma, \alpha)$  at mode  $n$  and a modality  $\mu : n \rightarrow m$ , we define the ‘locked’ context to be the following:

$$(G(\mu)_!(X), \Gamma.\{\mu\}, \gamma \circ [\mu]_! \alpha \circ \beta)$$

Here,  $\beta : \pi(G(\mu)_!(X)) \cong [\mu]_!(\pi(X))$  while  $\gamma : [\mu]_!(\mathbf{y}(\Gamma)) \cong \mathbf{y}(\Gamma.\{\mu\})$ .

*Context extension.* Given a context  $(X, \Gamma, \alpha)$  at mode  $m$ , a modality  $\mu : n \rightarrow m$  and a type  $(A^*, A)$  in the context  $(G(\mu)_!(X), \Gamma.\{\mu\}, \beta)$ , we form the context extension as the triple  $(X \times_{G(\mu)(\mathcal{T}_m)} G(\dot{\mathcal{T}}_m), \Gamma.(\mu \mid A), \beta)$ , where  $\beta$  is the composite:

$$\pi(X \times_{G(\mu)(\mathcal{T}_m)} G(\dot{\mathcal{T}}_m)) \cong \pi(X) \times_{[\mu]^*(\mathbf{Ty}_m)} [\mu]^*(\mathbf{Tm}_m) \cong \mathbf{y}(\Gamma.(\mu \mid A))$$

The first isomorphism follows from the fact that  $\pi$  preserves finite limits, is strictly 2-natural, and strictly preserves  $\tau_m$  while the second isomorphism is the universal property of  $\mathbf{y}(\Gamma.(\mu \mid A))$ .

*Modal types.* Suppose we are given a context  $(X, \Gamma, \alpha)$  and a type  $(A^*, A)$  in the context  $(G(\mu)!(\mu)(X), \Gamma.\{\mu\}, \gamma \circ \llbracket \mu \rrbracket^*(\alpha) \circ \beta)$ . We form the modal type as

$$(\mathbf{Mod}_\mu(\widehat{A^*}), \langle \mu \mid A \rangle)$$

It remains to check that these types are coherent. That is, that

$$\pi(\mathbf{Mod}_\mu(\widehat{A^*})) = \llbracket \mu \mid A \rrbracket \circ \alpha$$

By assumption,  $\pi(A^*) = [A] \circ \gamma \circ \llbracket \mu \rrbracket^*(\alpha) \circ \beta$ . By our assumption that  $\pi$  commutes with transposition,  $\pi(\widehat{A^*}) = \widehat{[A] \circ \gamma \circ \alpha}$ . The result follows from the fact that  $\pi$  preserves **Mod**.  $\square$

**3.1. An initial cosmos.** While the following material is not strictly necessary for our proof of normalization, we record it for general interest.

**Theorem 3.8.** *The category of MTT cosmoi can be sketched over  $\mathcal{V}$ .*

*Proof.* First we show that  $\mathcal{V}$  is 2-monadic over  $[[\mathcal{M}], \mathbf{Cat}_g]$ . This follows directly from the fact that locally Cartesian closure can be realized as a finitary 2-monad on  $\mathbf{Cat}_g$  [Lac09, Section 5.9] and pseudofunctors whose 1-cells are right adjoints are finitarily 2-monadic over  $[[\mathcal{M}], \mathbf{Cat}_g]$  [Lac09, Section 5.9]. Combining the operations and equations shows  $\mathcal{V}$  is finitarily 2-monadic over  $[[\mathcal{M}], \mathbf{Cat}_g]$ .

The constants and equations of MTT-cosmoi are sketchable over this 2-monad following the work of Kinoshita, Power, and Takeyama [KPT99].  $\square$

**Corollary 3.9.** *The category of MTT-cosmoi has a bi-initial object  $\mathcal{I}[-]$ .*

#### 4. FOUNDATIONS OF MULTIMODAL SYNTHETIC TAIT COMPUTABILITY

Eventually, we will wish to work with a collection of models of Synthetic Tait Computability (STC), connected by a variety of adjunctions. In order for this to work, however, we need to show that adjunctions between various categories lift to adjunctions between glued categories. The main result of this section is **Theorem 4.13**, which states that a collection of glued categories interconnected by right adjoints supports a model of “multimodal” STC.

**Theorem 4.1.** *Suppose we are given lex functors  $\rho_i : \mathbf{Sh}(\mathcal{X}_i) \rightarrow \mathbf{Sh}(\mathcal{Y}_i)$  and a 2-cell  $\alpha$  witnessing the weak commutativity of the following diagram:*

$$\begin{array}{ccc} \mathbf{Sh}(\mathcal{X}_0) & \xrightarrow{\rho_0} & \mathbf{Sh}(\mathcal{Y}_0) \\ f^* \downarrow & & \downarrow g^* \\ \mathbf{Sh}(\mathcal{X}_1) & \xrightarrow{\rho_1} & \mathbf{Sh}(\mathcal{Y}_1) \end{array}$$

Then there is an induced morphism  $[f, g] : \mathbf{Gl}(\rho_1) \rightarrow \mathbf{Gl}(\rho_0)$  such that  $[f, g]$  fits into the following diagram:

$$\begin{array}{ccccc} \mathcal{X}_1 & \hookrightarrow & \mathbf{Gl}(\rho_1) & \longleftarrow & \mathcal{Y}_1 \\ \downarrow & & \downarrow & & \downarrow \\ \mathcal{X}_0 & \hookrightarrow & \mathbf{Gl}(\rho_0) & \longleftarrow & \mathcal{Y}_0 \end{array}$$

Moreover, if both fringe functors  $\rho_i$  are continuous and  $f$  and  $g$  are essential, then  $[f, g]$  is also essential.

*Proof.* Explicitly,  $[f, g]^*$  sends  $F \rightarrow \rho_0(E)$  to  $g^*(F) \rightarrow \rho_1(f^*(E))$ . The 2-cell  $\alpha$  is used to obtain this morphism by correcting  $g(F \rightarrow \rho_0(E))$ .

Colimits and finite limits are determined pointwise in  $\mathbf{Sh}(\mathbf{Gl}(\rho_i))$ , so  $[f, g]^*$  preserves them because  $f^*$  and  $g^*$  both do. If both  $\rho_i$  are continuous, then all limits are determined pointwise, so  $[f, g]$  is essential if both  $f$  and  $g$  are.  $\square$

**Theorem 4.2.** *Suppose that  $f$  and  $g$  are essential, then  $[f, g]_!$  is a pointwise application of  $f_!$  and  $g_!$ .*

*Proof.* Explicitly, fix a morphism  $x : Y \rightarrow \rho_0(X)$ . We then define  $[f, g]_! x$  to be  $\tilde{x} : g_!(Y) \rightarrow \rho_1(g_!(X))$ , where  $\tilde{x} = \beta \circ g_!(x)$  and  $\beta$  is the canonical natural transformation determined by the following series of transposes:

$$\begin{aligned} \beta \in [g_! \circ \rho_1, \rho_0 \circ f_!] &\cong [\rho_1, g^* \circ \rho_1 \circ f_!] \\ &\cong [\rho_1, \rho_1 \circ f^* \circ f_!] \ni \rho_1(\eta) \end{aligned}$$

It remains to show that this defines an adjoint. We show that there is a bijection  $[[f, g]_!(x), x'] \cong [x, [f, g]^*(x')]$ . Given that  $[f, g]_!$  and  $[f, g]^*$  are defined pointwise, it suffices to show that the transpose operators for  $f_! \dashv f^*$  and  $g_! \dashv g^*$  lift.

Fix  $b : X \rightarrow f^* X'$  and  $a : Y \rightarrow g^*(Y')$  such that  $\alpha \circ g^*(x') \circ a = \rho_1(b) \circ x$ . We wish to show that  $x' \circ \hat{a} = \rho_0(\hat{b}) \circ \tilde{x}$ .

$$\begin{aligned} \rho_0(\hat{b}) \circ \tilde{x} &= \rho_0(\hat{b}) \circ \beta \circ g_!(x) \\ &= \rho_0(\hat{b}) \circ \epsilon \circ g_!(\alpha^{-1} \circ \rho_1(\eta)) \circ g_!(x) \\ &= \rho_0(\hat{b}) \circ \epsilon \circ g_!(\alpha^{-1} \circ \rho_1(\eta) \circ x) \\ &= \epsilon \circ g_!(g^*(\rho_0(\hat{b}))) \circ \alpha^{-1} \circ \rho_1(\eta) \circ x \\ &= \epsilon \circ g_!(\alpha^{-1} \circ \rho_1(f^*(\hat{b}))) \circ \rho_1(\eta) \circ x \\ &= \epsilon \circ g_!(\alpha^{-1} \circ \rho_1(b) \circ x) \\ &= \epsilon \circ g_!(g^*(x') \circ a) \\ &= x' \circ \hat{a} \end{aligned}$$

Next,  $b : f_!(X) \rightarrow X'$  and  $a : g_!(Y) \rightarrow Y'$  such that  $x' \circ a = \rho_0(b) \circ \tilde{x}$ . We wish to show that  $\alpha \circ g^*(x') \circ \hat{a} = \rho_1(\hat{b}) \circ x$ .

$$\begin{aligned} \alpha \circ g^*(x') \circ \hat{a} &= \alpha \circ g^*(x') \circ g^*(a) \circ \eta \\ &= \alpha \circ g^*(x' \circ a) \circ \eta \\ &= \alpha \circ g^*(\rho_0(b) \circ \tilde{x}) \circ \eta \end{aligned}$$

$$\begin{aligned}
&= \rho_1(f^*(b)) \circ \alpha \circ g^*(\epsilon \circ g!(\alpha^{-1} \circ \rho_1(\eta) \circ x)) \circ \eta \\
&= \rho_1(f^*(b)) \circ \alpha \circ g^*(\epsilon) \circ \eta \circ \alpha^{-1} \circ \rho_1(\eta) \circ x \\
&= \rho_1(f^*(b)) \circ \rho_1(\eta) \circ x \\
&= \rho_1(\widehat{b}) \circ x
\end{aligned}
\quad \square$$

It is natural to view  $\mathbf{Gl}(\rho_i)$  as a topos with a subterminal object  $\mathbf{syn}_i$  such that the open subtopos corresponding to  $\mathbf{syn}_i$  is equivalent to  $\mathcal{X}_i$ , while the closed subtopos is equivalent to  $\mathcal{Y}_i$ . For readability, we will identify these subtopoi with  $\mathcal{X}_i$  and  $\mathcal{Y}_i$  in what follows. Recall that the inverse image of the inclusion  $\mathcal{X}_i \hookrightarrow \mathbf{Gl}(\rho_i)$  is given by  $-\mathbf{syn}_i$ . The inverse image of  $\mathcal{Y}_i \hookrightarrow \mathbf{Gl}(\rho_i)$  is given by  $-\star \mathbf{syn}_i$ , where  $A \star \mathbf{syn}_i$  is the pushout  $A \amalg_{A \times \mathbf{syn}_i} \mathbf{syn}_i$ .

**4.1. Open and closed subtopoi.** We now devote substantial effort to showing that the decomposition of  $\mathbf{Gl}(\rho_i)$  is preserved by  $[f, g]$ .

**Lemma 4.3.** *If  $A : \mathbf{Gl}(\rho_0)$  is contained in  $\mathbf{Sh}(\mathcal{X}_0)$ , then  $[f, g](A) : \mathbf{Sh}(\mathcal{X}_1)$ .*

*Proof.* Observe that  $A$  is in the open subtopos  $\mathbf{Sh}(\mathcal{X}_0)$  if  $A \cong \rho_0(E) \rightarrow \rho_0(E)$ , in which case  $[f, g](A) \cong \rho_1(f(E)) \rightarrow \rho_1(f(E))$  which then lies in the open subtopos.  $\square$

**Lemma 4.4.** *If  $A : \mathbf{Sh}(\mathbf{Gl}(\rho_0))$  is contained in  $\mathbf{Sh}(\mathcal{Y}_0)$ , then  $[f, g](A) : \mathbf{Sh}(\mathcal{Y}_1)$ .*

*Proof.* Again, observe that  $A : \mathbf{Sh}(\mathcal{Y}_0)$  if  $A \cong F \rightarrow \rho_0(\mathbf{1}_{\mathbf{Sh}(\mathcal{X}_0)})$ , in which case

$$[f, g]^*(A) \cong g(F) \rightarrow \rho_1(f(\mathbf{1}_{\mathbf{Sh}(\mathcal{X}_0)})) \cong g(F) \rightarrow \rho_1(\mathbf{1}_{\mathbf{Sh}(\mathcal{X}_1)})$$

This then lies in the closed subtopos of  $\mathbf{Gl}(\rho_1)$ .  $\square$

In fact, a more general theorem is true. Reflecting an object into either subtopos and applying  $[f, g]$  is (naturally) isomorphic to applying  $[f, g]$  and then reflecting.

**Lemma 4.5.** *There is an isomorphism  $\mathbf{syn}_1 \rightarrow [f, g](\mathbf{syn}_0)$ .*

*Proof.* Unfolding definitions:

$$\begin{aligned}
\mathbf{syn}_0 &= \mathbf{0}_{\mathbf{Sh}(\mathcal{Y}_0)} \rightarrow \rho_0(\mathbf{1}_{\mathbf{Sh}(\mathcal{X}_0)}) \\
\mathbf{syn}_1 &= \mathbf{0}_{\mathbf{Sh}(\mathcal{Y}_1)} \rightarrow \rho_1(\mathbf{1}_{\mathbf{Sh}(\mathcal{X}_1)})
\end{aligned}$$

The result is immediate by computation.  $\square$

**Corollary 4.6.**

- (1)  $[f, g](\mathbf{syn}_0) \cong \mathbf{syn}_1$
- (2)  $[f, g](-\mathbf{syn}_0) \cong ([f, g](-))^{\mathbf{syn}_1}$
- (3)  $[f, g](\mathbf{syn}_0 \star -) \cong \mathbf{syn}_1 \star [f, g](-)$

*Remark 5.* The last fact of [Corollary 4.6](#) is a consequence of the first;  $\star$  is preserved by  $[f, g]$  because this functor is lex and cocontinuous and  $\star$  is defined by finite products and pushouts. Exponentiation, however, is not generally preserved by  $[f, g]$ . It is a special fact of  $\mathbf{syn}_0$  that exponentiation by  $\mathbf{syn}_0$  is preserved.

As idempotent lex modalities, both of these actions internalize into the internal type theory of the topoi. In particular, they have actions on a family given by applying the operation to the entire family, then pulling back along the unit. This is automatically preserved by  $[f, g]^*$ , as  $[f, g]^*$  preserves each step in this construction.

We write  $\circ$  for the action of  $-\mathbf{syn}_0$  on families, and  $\bullet$  for  $-\star\mathbf{syn}_0$ . We will abusively use the same notation for  $-\mathbf{syn}_0$  and  $-\mathbf{syn}_1$  as well as  $-\star\mathbf{syn}_0$  and  $-\star\mathbf{syn}_1$ .

**Theorem 4.7.** *Both  $\circ$  and  $\bullet$  are preserved by  $[f, g]^*$*

We would like to show a slightly stronger result, namely that the dependent versions of these modalities are preserved [RSS20]. Prior to this, we must address the question of universes and models of type theory in this situation.

As a morphism of logoi, we know that  $[f, g]$  is a cocontinuous left exact functor. We would like to show that it induces a dependent right adjoint [Bir+20], but this is complicated by the murky definition of a model of dependent type theory with universes in an arbitrary logoi. Accordingly, prove a theorem which is sufficient for the case we have in mind.

**Definition 4.8.** A functor  $F : \mathbf{PSh}(\mathcal{C}) \rightarrow \mathbf{PSh}(\mathcal{D})$  between presheaf categories is said to preserve a Grothendieck universe  $\mathcal{U}$  when it sends  $\mathcal{U}$ -small families in  $\mathbf{PSh}(\mathcal{C})$  to  $\mathcal{U}$ -small families in  $\mathbf{PSh}(\mathcal{D})$ .

**Theorem 4.9.** *A right adjoint  $F : \mathbf{PSh}(\mathcal{C}) \rightarrow \mathbf{PSh}(\mathcal{D})$  which preserves Grothendieck universes larger than  $|\mathcal{C}|$  and  $|\mathcal{D}|$  induces a dependent right adjoint.*

*Remark 6.* To be precise, we must specify the particular models of dependent type theory we consider in  $\mathbf{PSh}(\mathcal{C})$  and  $\mathbf{PSh}(\mathcal{D})$ . We take the model of type theory in which a type is an element of the universe of small presheaves [HS97].

*Proof.* Fix a Grothendieck universe  $\mathcal{V}$  which is large enough to contain both  $\mathcal{C}$  and  $\mathcal{D}$ . We wish to show that  $F$  induces a weak CwF morphism. A type in  $\mathbf{PSh}(\mathcal{C})$  is given by a generalized element of  $\mathcal{T}_{\mathcal{C}}$ , the Hofmann-Streicher universe of  $\mathcal{V}$ -small presheaves, so it suffices to show that there is a pullback square of the following shape:

$$\begin{array}{ccc} F(\dot{\mathcal{T}}_{\mathcal{C}}) & \longrightarrow & \dot{\mathcal{T}}_{\mathcal{D}} \\ F(\tau_{\mathcal{C}}) \downarrow \lrcorner & & \downarrow \tau_{\mathcal{D}} \\ F(\mathcal{T}_{\mathcal{C}}) & \longrightarrow & \mathcal{T}_{\mathcal{D}} \end{array}$$

Proving this is equivalent to showing that  $F(\tau_{\mathcal{C}})$  is  $\mathcal{U}$ -small because  $\tau_{\mathcal{D}}$  is generic for such maps. This immediately defines a weak CwF morphism using the observation that context extension in these models is defined by pullback.  $\square$

**Theorem 4.10.** *A cocontinuous lex functor  $F : \mathbf{PSh}(\mathcal{C}) \rightarrow \mathbf{PSh}(\mathcal{D})$  preserves a Grothendieck universe  $\mathcal{U}$  if and only if it restricts to a functor  $\mathbf{PSh}_{\mathcal{U}}(\mathcal{C}) \rightarrow \mathbf{PSh}_{\mathcal{U}}(\mathcal{D})$ .*

*Proof.* The only if direction is clear, so it remains to show that preserving  $\mathcal{U}$ -objects ensures that  $F$  preserves  $\mathcal{U}$ -families. Fix a  $\mathcal{U}$ -small family  $f : X \rightarrow Y : \mathbf{PSh}(\mathcal{C})$ .

First, we observe that  $F(Y) = \text{colim}_i F(\mathbf{y}(C_i))$  using the canonical decomposition of  $Y$  into colimits. By Yoneda, a morphism  $\mathbf{y}(D) \rightarrow F(Y)$  then factors through some  $F(\mathbf{y}(C_i))$ , followed by  $F([y_i]) : F(\mathbf{y}(C_i)) \rightarrow F(Y)$ , so it suffices to show that the pullback of  $F(f)$  to  $F(\mathbf{y}(C_i))$  is a small family.

As  $F(f)$  and  $F([y_i])$  are both in the image of  $F$ , this pullback can be computed in  $\mathbf{PSh}(\mathcal{C})$ . Therefore, we must show that  $F(X \times_Y \mathbf{y}(C_i)) \rightarrow F(\mathbf{y}(C_i))$  is a small family. Using the assumption that  $f$  was a small family, we observe that  $X \times_Y \mathbf{y}(C_i)$

is a small object, so  $F(X \times_Y \mathbf{y}(C_i))$  is  $\mathcal{U}$ -small. Therefore, this family is a small family, completing the proof.  $\square$

While these results are limited to presheaf topoi, they are sufficient for our purposes; we are interested in gluing together presheaf topoi along a continuous functor. The following result shows that the resultant topos is of presheaf type:

**Theorem 4.11** (Artin, Grothendieck, and Verdier [AGV72] and Carboni and Johnstone [CJ95]). *Gluing together presheaf topoi along a continuous functor results in a presheaf topos.*

Motivated by this, we now assume that  $\mathcal{E}_i$  and  $\mathcal{F}_i$  are all presheaf logoi, and all morphisms involved preserve all Grothendieck universes larger than  $\mathcal{U}$ . In this case, we may then pick a Grothendieck universe  $\mathcal{U}$  large enough to be preserved by both  $f$  and  $g$  and so the bases of  $\mathbf{Gl}(\rho_1)$  and  $\mathbf{Gl}(\rho_1)$  are small for this universe. In this case,  $[f, g]$  is a DRA which preserves all universes larger than  $\mathcal{U}$ .

We now show that  $[f, g]$  preserves the dependent version of  $\circ$  and  $\bullet$ .

**Theorem 4.12.**  $[f, g]^*$  preserves  $\hat{\circ}$  and  $\hat{\bullet}$  on any universe preserved by  $[f, g]^*$ .

*Proof.* We will show this for  $\hat{\circ}$ , though the proof is identical for  $\hat{\bullet}$  and indeed for any lex modality preserved by  $[f, g]$ .

Let us pick some universe  $\mathcal{U}$  large enough to be preserved by  $[f, g]$  and write  $\tau_0 : \dot{\mathcal{T}}_0 \rightarrow \mathcal{T}_0$  for the Hofmann-Streicher universe it induces in  $\mathbf{Sh}(\mathbf{Gl}(\rho_0))$  and  $\tau_1 : \dot{\mathcal{T}}_1 \rightarrow \mathcal{T}_1$  for the Hofmann-Streicher universe in  $\mathbf{Sh}(\mathbf{Gl}(\rho_1))$ . By assumption, we have a pullback squares

$$\begin{array}{ccc} [f, g]^* \dot{\mathcal{T}}_0 & \longrightarrow & \dot{\mathcal{T}}_1 \\ [f, g]^* \tau_0 \downarrow & \lrcorner & \downarrow \tau_1 \\ [f, g]^* \mathcal{T}_0 & \xrightarrow{i} & \mathcal{T}_1 \end{array} \quad \begin{array}{ccc} \circ \dot{\mathcal{T}}_i & \longrightarrow & \dot{\mathcal{T}}_i \\ \circ \tau_i \downarrow & \lrcorner & \downarrow \tau_i \\ \circ \mathcal{T}_i & \xrightarrow{\hat{\circ}} & \mathcal{T}_i \end{array}$$

Finally, we note that  $[f, g]^*$  preserves  $\circ$ , meaning that there is a natural isomorphism  $\alpha : [f, g]^* \circ \circ \rightarrow \circ \circ [f, g]^*$ .

We may now rephrase: our goal is to show that  $i \circ [f, g]^*(\hat{\circ})$  and  $\hat{\circ} \circ \circ i \circ \alpha$  represent the same type. First, compute the type represented by  $\hat{\circ} \circ \circ i \circ \alpha$ :

$$\begin{array}{ccccc} [f, g]^* \circ \dot{\mathcal{T}}_0 & \longrightarrow & \circ \dot{\mathcal{T}}_1 & \longrightarrow & \dot{\mathcal{T}}_1 \\ \downarrow & \lrcorner & \downarrow & \lrcorner & \downarrow \\ [f, g]^* \circ \mathcal{T}_0 & \xrightarrow{\circ i \circ \alpha} & \circ \mathcal{T}_1 & \xrightarrow{\hat{\circ}} & \mathcal{T}_1 \end{array}$$

The left-hand square is a pullback because  $\circ$  is lex. Next, we compute  $i \circ [f, g]^*(\hat{\circ})$ :

$$\begin{array}{ccccc}
 [f, g]^* \circ \dot{\mathcal{T}}_0 & \longrightarrow & [f, g]^* \dot{\mathcal{T}}_1 & \longrightarrow & \dot{\mathcal{T}}_1 \\
 \downarrow \lrcorner & & \downarrow \lrcorner & & \downarrow \\
 [f, g]^* \circ \mathcal{T}_0 & \xrightarrow{[f, g]^* \circ} & [f, g]^* \mathcal{T}_1 & \xrightarrow{i} & \mathcal{T}_1
 \end{array}$$

The left-hand square is now a pullback because  $[f, g]^*$  is lex.  $\square$

**4.2. Interpreting MTT in glued topoi.** The results of the previous subsection show that  $[f, g]$  preserves the open and closed modalities of a glued topos. As a dependent right adjoint, moreover,  $[f, g]$  can be internalized into a model of MTT. It is this model of MTT that we use to substantiate the language of ‘multimodal’ synthetic Tait computability.

**Theorem 4.13.** *Fix a 2-category  $\mathcal{M}$  and a 2-natural transformation  $\alpha$  between a pair of 2-functors  $\llbracket - \rrbracket_0, \llbracket - \rrbracket_1 : \mathcal{M}^{\text{op}} \rightarrow \mathbf{Cat}$  such that both 2-functors assigns 0-cells to presheaf topoi and 1-cells to essential geometric morphisms between them. There exists a model of MTT with  $\mathcal{M}$  such that mode  $m$  is given by  $\mathbf{Gl}(\llbracket \alpha_m \rrbracket^*)$  and the modality  $\mu$  is interpreted by  $\llbracket \mu \rrbracket_0^*, \llbracket \mu \rrbracket_1^*$ .*

*Furthermore, each mode has an open and closed modality preserved by the MTT modalities such that the open (resp. closed) subtopos of mode  $m$  is equivalent to  $\llbracket m \rrbracket_0$ , (resp.  $\llbracket m \rrbracket_1$ ).*

In this model, each mode supports a model of synthetic Tait computability and the extra structure of this language (the open and closed modalities) is preserved by the modalities between modes.

## 5. THE CATEGORY OF RENAMINGS FOR MTT

We now isolate a class of renamings:  $\text{Ren}_m$ . We also define the neutral and normal forms of terms in MTT. Unlike the syntax of MTT, neutral and normal forms are not taken up to a complex equivalence relation, and so their conversion problem is immediately reducible to conversion in the mode theory.

The judgments for these are slightly atypical, in that the typing judgments for both neutral and normal forms are defined (1) with formal telescopes as contexts and (2) inductive-recursively with the inclusion of neutral and normals into terms, and renamings into contexts and substitutions. This dependence of the unquoted raw syntax on actual terms is necessary in order to ensure that the typing relation can be sensibly defined.

The full collection of neutral and normal forms is given in [Appendix A](#). We present a few representative cases below.

$$\boxed{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \vdash \psi : \Psi @ m \rightsquigarrow \gamma}$$

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma}{\Theta.\{\mu\} \text{ tele } @ n \rightsquigarrow \Gamma.\{\mu\}} \quad \frac{\Gamma, \Delta \text{ tele } @ m \rightsquigarrow \Theta, \Psi \quad \Theta \vdash r : \Psi @ m \rightsquigarrow \delta}{\Theta.\{\mu\} \vdash r.\{\mu\} : \Psi.\{\mu\} @ n \rightsquigarrow \delta.\{\mu\}}$$



$$\begin{array}{c}
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \mu, \nu : n \longrightarrow m \quad \alpha : \nu \longrightarrow \mu}{\Theta.\{\mu\} \vdash \{\alpha\}_\Theta : \Theta.\{\nu\} @ n \rightsquigarrow \{\alpha\}_\Gamma} \\
\\
\boxed{\Theta \models^{\text{nf}} \tau @ m \rightsquigarrow A} \\
\\
\frac{\Theta \models^{\text{nf}} \tau @ n \rightsquigarrow A \quad \Theta.(\text{id}_m \mid A) \models^{\text{nf}} \sigma @ m \rightsquigarrow B}{\Theta \models^{\text{nf}} \tau \times \sigma @ m \rightsquigarrow A \times B} \quad \frac{\Theta.\{\mu\} \models^{\text{nf}} \tau @ n \rightsquigarrow A}{\Theta \models^{\text{nf}} \langle \mu \mid \tau \rangle @ m \rightsquigarrow \langle \mu \mid A \rangle} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^{\text{nf}} u : \mathbf{U} @ m \rightsquigarrow A}{\Theta \models^{\text{nf}} \text{El}(u) @ m \rightsquigarrow A} \\
\\
\boxed{\Theta \models^{\text{ne}} e : A @ m \rightsquigarrow M \quad \Theta \models^{\text{nf}} u : A @ m \rightsquigarrow M} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta(k) = (\mu \mid A) \quad \text{mods}(\Theta, k) = \nu \quad \alpha : \mu \longrightarrow \nu}{\Theta \models^{\text{ne}} \mathbf{v}_k^\alpha : A[\{\alpha\} \circ (\uparrow.\{\nu_{k-1}\}) \cdots \circ (\uparrow.\{\nu_0\})] @ m \rightsquigarrow \mathbf{v}_0[\{\alpha\} \circ (\uparrow.\{\nu_{k-1}\}) \cdots \circ (\uparrow.\{\nu_0\})]} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta.(\mu \mid A) \models^{\text{nf}} u : B @ m \rightsquigarrow M}{\Theta \models^{\text{nf}} \lambda(u) : (\mu \mid A) \rightarrow B @ m \rightsquigarrow \lambda(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta \models^{\text{ne}} e : (\mu \mid A) \rightarrow B @ m \rightsquigarrow M \quad \Theta \models^{\text{nf}} u : A @ m \rightsquigarrow N}{\Theta \models^{\text{ne}} e(u) : B[\text{id}.N] @ m \rightsquigarrow M(N)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Theta.\{\mu\} \models^{\text{nf}} u : A @ n \rightsquigarrow M}{\Theta \models^{\text{nf}} \text{mod}_\mu(u) : \langle \mu \mid A \rangle @ m \rightsquigarrow \text{mod}_\mu(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Theta \models^{\text{ne}} e : \langle \mu \mid A \rangle @ m \rightsquigarrow M}{\Theta \models^{\text{nf}} \text{up}(e) : \langle \mu \mid A \rangle @ m \rightsquigarrow M} \\
\\
\frac{\nu : o \longrightarrow n \quad \mu : n \longrightarrow m \quad \Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\}.\{\nu\} \vdash A @ o \quad \Theta.\{\mu\} \models^{\text{ne}} u : \langle \nu \mid A \rangle @ n \rightsquigarrow M \quad \Theta.(\mu \mid \langle \nu \mid A \rangle) \models^{\text{nf}} \tau @ m \rightsquigarrow B \quad \Theta.(\mu \circ \nu \mid A) \models^{\text{nf}} u : B[\uparrow.\text{mod}_\nu(\mathbf{v}_0)] @ m \rightsquigarrow}{\Theta \models^{\text{ne}} \text{letmod}(\mu; \nu; \tau; e; u) : B[\text{id}.M] @ m \rightsquigarrow \text{let}_\nu \text{mod}_\mu(\_) \leftarrow M \text{ in } N} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^{\text{ne}} e : \text{El}(\widehat{\text{bool}}) @ m \rightsquigarrow M}{\Theta \models^{\text{ne}} \text{dec}^\triangleright(e) : \text{bool} @ m \rightsquigarrow \text{dec}^\triangleright(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^{\text{nf}} u : \text{bool} @ m \rightsquigarrow M}{\Theta \models^{\text{nf}} \text{dec}^\triangleleft(u) : \text{El}(\widehat{\text{bool}}) @ m \rightsquigarrow \text{dec}^\triangleleft(M)}
\end{array}$$

There are no equations imposed on any of these generators, except for renamings which have the necessary equations to organize them into a 2-functor into **Cat**. We further define a substitution action applying a renaming to a normal or neutral form, which we must show respects these equations. This operation and its pertinent

proofs are largely standard, with the exception being the case for variables. We reproduce the definition of the substitution action on variables here:

$$\begin{aligned}
\mathbf{v}_k^\alpha[\text{id}] &= \mathbf{v}_k^\alpha \\
\mathbf{v}_k^\alpha[\uparrow.\{\mu\}] &= \mathbf{v}_{k+1}^\alpha \\
\mathbf{v}_k^\alpha[r.\mathbf{v}_j^\beta.\{\mu\}] &= \begin{cases} \mathbf{v}_j^{(\beta \star \text{id}_\mu) \circ \alpha} & k = 0 \\ \mathbf{v}_{k-1}^\alpha[r.\{\mu\}] & \text{otherwise} \end{cases} \\
\mathbf{v}_k^\alpha[(r \circ s).\{\mu\}] &= \mathbf{v}_k^\alpha[r.\{\mu\}][s.\{\mu\}] \\
\mathbf{v}_k^\alpha[\{\beta\}\Theta] &= \mathbf{v}_k^{\alpha \circ \beta}
\end{aligned}$$

*Remark 7.* Note that these normal forms do not necessarily enjoy decidable equality. Rather, the problem of deciding when two normal forms are convertible is precisely the problem of deciding whether certain 1- and 2-cells of the mode theory  $\mathcal{M}$  are equal. Accordingly, it is possible that MTT may enjoy normalization, but not decidable type-checking. The mode theories used in instantiations of MTT thus far are evidently decidable, so we do not expect this to be an issue for practical usage.

## 6. PREREQUISITES FOR THE NORMALIZATION MODEL

**6.1. Key syntactic categories.** Recall from [Section 5](#) that for each mode there is a category  $\mathbf{Ren}_m$  of telescopes and renamings, together with a functor  $\mathbf{i}[m] : \mathbf{Ren}_m \rightarrow \mathbf{Cx}_m$  which sends a telescope the corresponding context and formal renaming to substitution. The collection of these functors induces a natural transformation between the 2-functors  $\mathbf{Cx}_-, \mathbf{Ren}_- : \mathcal{M}^{\text{coop}} \rightarrow \mathbf{Cat}$ .

The functor  $\mathbf{i}[m]$  induces a functor  $\mathbf{i}[m]^* : \mathbf{PSh}(\mathbf{Cx}_m) \rightarrow \mathbf{PSh}(\mathbf{Ren}_m)$ . Moreover, because  $\mathbf{PSh}(-)$  preserves strict equalities, this gives rise to a 2-functor  $\mathcal{M} \times \Delta^1 \rightarrow \mathbf{Cat}$ . As a morphism defined by precomposition each  $\mathbf{i}[m]^*$  has a left adjoint and a right adjoint:  $\mathbf{i}[m]_! \dashv \mathbf{i}[m]^* \dashv \mathbf{i}[m]_*$ . The left adjoint extends  $\mathbf{i}[m]$  in the sense that  $\mathbf{i}[m]_!(\mathbf{y}(\Theta)) \cong \mathbf{y}(\mathbf{i}[m](\Theta))$ . We may now apply [Theorem 4.13](#).

**Theorem 6.1.** *There exist a model  $\mathcal{G}[-]$  of MTT with a hierarchy of cumulative universes in  $\mathbf{Gl}(\mathbf{i}[m]^*)$ .*

*In this model,  $\mathcal{G}[m]$  is  $\mathbf{Gl}(\mathbf{i}[m]^*)$ , and  $\mathcal{G}[\mu]$  is interpreted by precomposition with  $\mu$  in both components. Furthermore, there is a (dependent) open and closed modality in  $\mathcal{G}[m]$ , preserved by all modalities.*

In what follows, we will freely use MTT to work with this model. As the interpretation of identity types in this model supports equality reflection, we will work with extensional equality when reasoning with MTT.

**6.2. Basic properties of  $\mathcal{G}[m]$ .** The MTT modalities in this model are particularly well-behaved because they arise from the inverse image of essential geometric morphisms. In particular, they preserve colimits internally to the theory.

**Theorem 6.2.** *For each  $\mu$  there is an equivalence  $\langle \mu \mid A + B \rangle \simeq \langle \mu \mid A \rangle + \langle \mu \mid B \rangle$ .*

*Proof.* Recall that each modality in this model is interpreted by a morphism which is simultaneously a left and right adjoint. Therefore, this model can be extended to the mode theory  $\mathcal{M}_{\text{adj}}$ , which extends  $\mathcal{M}$  by adding a right adjoint  $\bar{\mu}$  for each  $\mu$  from  $\mathcal{M}$  [[Gra+20a](#), Section 10.2].

The *crisp induction* principles derived in Gratzer, Kavvos, Nuyts, and Birkedal [Gra+20a, Section 10.5] now show that  $\langle \mu \mid - \rangle$  preserves coproducts internally with this extended mode theory, so the equivalence is validated by the model from [Theorem 6.1](#) (though is no longer internally derivable).  $\square$

*Remark 8.* We note that this model of MTT enjoys  $\Gamma.(\mu \circ \nu \mid A) \cong \Gamma.(\mu \mid \langle \nu \mid A \rangle)$ : the modalities are interpreted by proper dependent right adjoints. We therefore blur the distinction between a variable  $(\mu \mid x : A)$  and a variable  $(\text{id} \mid x : \langle \mu \mid A \rangle)$ .

**Notation 6.3.** It's convenient to recall that  $\circ$  is defined by exponentiation with a subterminal (a proposition) **syn**. We accordingly write  $\circ_{z:\mathbf{syn}} A(z)$ .

*Remark 9.* Strictly speaking, **syn** should contain a mode annotation. However,  $\langle \mu \mid \mathbf{syn}_n \rangle \cong \mathbf{syn}_m$ , and as both are subterminal we identify  $\mathbf{syn}_m$  and  $\langle \mu \mid \mathbf{syn}_n \rangle$ .

Finally, we note also that  $\circ$  and  $\bullet$  enjoy two important properties:

**Theorem 6.4** (Fracture [AGV72]). *For any type  $A$  we have  $A \cong \circ A \times_{\bullet \circ A} \bullet A$ .*

**Theorem 6.5** (Internal realignment [OP18; SH20]). *Let us denote the type of types isomorphic to  $A$  as  $\text{Iso}(A)$ :*

$$\text{Iso}(A) = \sum_{B:\mathbf{U}} (A \cong B)$$

*There is a section  $\text{re}$  to  $\eta : \text{Iso}(A) \rightarrow \circ \text{Iso}(A)$ .*

*Remark 10.* The following constructive proof is due to Christian Sattler.

*Proof.* This follows directly from the argument given by Orton and Pitts [OP18, Theorem 8.4]. It suffices to show that **syn** is levelwise decidable externally. This is immediate when we unfold the presentation of  $\mathcal{G}[[m]]$  as a presheaf topos.

Explicitly, Carboni and Johnstone [CJ95] show that  $\mathcal{G}[[m]]$  can be presented as  $\mathbf{PSh}(\mathbf{Coll}([- , \mathbf{i}[m](-)]))$ . While the precise definition of  $\mathbf{Coll}([- , \mathbf{i}[m](-)])$  is given by Carboni and Johnstone [CJ95], for our purposes it suffices to recall that the collection of objects of this *collage* is the disjoint union of the objects of  $\mathbf{Cx}_m$  and  $\mathbf{Ren}_m$ . Under this presentation moreover, **syn** can be defined as follows:

$$\mathbf{syn}(\Theta) = \emptyset \quad \mathbf{syn}(\Gamma) = \{\star\}$$

The levelwise decidability of **syn** is then immediate.  $\square$

In anticipation of using this model of MTT to construct the normalization model in the next section, we also introduce several constants interpreted by specific presheaves in  $\mathcal{G}[[m]]$ .

**Definition 6.6.** There is a constant  $\vdash \text{Ty}_m : \circ \mathbf{U} @ m$  in  $\mathcal{G}[[m]]$  given by  $\mathcal{T}_m$ .

**Definition 6.7.** There is a type  $z : \mathbf{syn}, \text{Ty}_m(z) \vdash \text{Tm}_m : \mathbf{U}$  in  $\mathcal{G}[[m]]$  whose total space is given by  $\tau_m : \tilde{\mathcal{T}}_m \rightarrow \mathcal{T}_m$ .

**Definition 6.8.** There is pair of types  $\circ_z \text{Ty}_m(z) \vdash \text{Nf}_m$  and  $\circ_z \text{Ty}_m(z) \vdash \text{Ne}_m$  in  $\mathcal{G}[[m]]$  given by normal and neutral forms viewed as a presheaf over  $\mathbf{i}[m]^*(\tilde{\mathcal{T}}_m)$ . Similarly, there is a closed type  $\text{NfTy}_m$  of normal types lying over  $\mathbf{i}[m]^*(\mathcal{T}_m)$ .

**Definition 6.9.** There is a type  $\circ_z \text{Ty}_m(z) \vdash \mathbf{V}_m$  given by viewing the presheaf of variables as lying over  $\mathbf{i}[m]^*(\tilde{\mathcal{T}}_m)$ . There is an induced inclusion  $\mathbf{V}_m(A) \hookrightarrow \text{Ne}_m(A)$  which we shall treat as silent.

Using realignment if necessary, we choose  $\text{Nf}_m$ ,  $\text{Ne}_m$ ,  $\text{V}_m$  such that under the assumptions  $z : \mathbf{syn}$  and  $A : \circ_z \text{Ty}_m(z)$ , we have the following equations:

$$\text{V}_m(A) = \text{Ne}_m(A) = \text{Nf}_m(A) = \text{Tm}_m(z, A(z))$$

Similarly, under just the assumption of  $z$ , we have  $\text{NfTy}_m = \text{Ty}_m(z)$ .

**6.3. An MTT cosmos, internally.** In [Section 7](#) we will construct a model of MTT in  $\mathcal{G}[-]$  lying over the syntactic model. To that end, we recast the existence of a model of MTT into a structure we can express in the internal language of  $\mathcal{G}[-]$ . Explicitly, we rephrase [Definition 3.3](#) as a sequence of constants in the internal type theory e.g., a pair  $\text{Ty}_m$  and  $\text{Ty}_m \vdash \text{Tm}_m$ , combinators for all the type codes, etc. The full list of constants is described in [Appendix B](#).

*Remark 11.* There is potential for confusion here: we are using the internal language of  $\mathcal{G}[-]$  (extensional MTT) as a framework to express the structure of an MTT cosmos. This is not a circularity; the interpretation of extensional MTT into  $\mathcal{G}[-]$  is formulated in terms of the models and metatheory developed already in Gratzner, Kavvos, Nuyts, and Birkedal [[Gra+20a](#)].

Already, we can immediately obtain an MTT cosmos internal to  $\mathcal{G}[-]$  by restricting to the open subtopos:  $\mathcal{S}[-]$ . In this cosmos, types are interpreted as elements of the constant  $\text{Ty}_m(z)$  while terms are elements of  $\text{Tm}_m(z, -)$ . This the *syntactic* MTT cosmos. Externally, this is the MTT cosmos  $\mathcal{S}[-]$  where universes are interpreted by the traditional representable natural transformations [[Awo18](#)]. Eventually, we will construct the normalization model as a second cosmos in  $\mathcal{G}[-]$  which lies strictly over the syntactic cosmos.

We will use the “unqualified” names such as  $\text{Prod}(z)$  to refer to elements of this syntactic cosmos. So, for instance, under the assumption  $z : \mathbf{syn}$  we the following:

$$\begin{aligned} \text{Mod}[\mu] &: (\mu \mid A : \text{Ty}_n(z)) \rightarrow \text{Ty}_m(z) \\ \text{m}[\mu] &: (\mu \mid A : \text{Ty}_n(z))(\mu \mid a : \text{Tm}_n(z, A)) \rightarrow \text{Tm}_m(\text{Mod}[\mu](A)) \\ \text{letmod}[\mu; \nu] &: (\nu \circ \mu \mid A : \text{Ty}_n(z)) [B : \langle \mu \mid \text{Tm}_n(z, \text{Mod}[\mu](A)) \rangle \rightarrow \text{Ty}_o(z)] \rightarrow \\ &\quad [(\nu \circ \mu \mid x : \text{Tm}_n(z, A)) \rightarrow \text{Tm}_o(z, B(\text{m}[\mu](A, x)))] \rightarrow \\ &\quad (\nu \mid a : \text{Tm}_m(z, \text{Mod}[\mu](A))) \rightarrow \\ &\quad \text{Tm}_o(z, B(a)) \\ - : (\nu \circ \mu \mid A : \text{Ty}_n(z)) [B : \langle \mu \mid \text{Tm}_n(z, \text{Mod}[\mu](A)) \rangle \rightarrow \text{Ty}_o(z)] \rightarrow \\ &\quad [b : (\nu \circ \mu \mid x : \text{Tm}_n(z, A)) \rightarrow B(\text{m}[\mu](A, x))] \rightarrow \\ &\quad (\nu \circ \mu \mid a : \text{Tm}_n(z, A)) \rightarrow \text{letmod}[\mu; \nu](A, B, b, \text{m}[\mu](A, a)) = b(a) \end{aligned}$$

Now, because  $\circ$  commutes with dependent products, sums, and modalities, we may equivalently view, e.g.  $\text{Prod}$  as a constant with the following type:

$$\left[ \sum_{A : \langle \mu \mid \circ_z \text{Ty}_n(z) \rangle} \langle \mu \mid \circ_z \text{Tm}_n(z, A) \rangle \rightarrow \circ_z \text{Ty}_m(z) \right] \rightarrow \circ_z \text{Ty}_m(z)$$

Similarly, we will abusively write  $\text{Prod}(A, B)$  for  $\lambda z. \text{Prod}(A(z), B(-, z))$ .

**6.4. Higher-order abstract syntax for neutral and normal forms.** The types  $\text{Ne}_m$ ,  $\text{Nf}_m$  can be used to encode a form of HOAS inside of  $\mathcal{G}[m]$  [[Hof99](#)]. We record the constants that result from this in [Appendix C](#).

In the proof of normalization we will take advantage of the fact that under an open modality, a normal and neutral form decode to the appropriate terms. Therefore, under the open modality  $\mathbf{proj}_0(\mathbf{pair}(M, N))$  is well-typed and equal to  $M$  along with other expected equations.

## 7. THE NORMALIZATION MODEL

We now construct a model of MTT in  $\mathcal{G}[-]$  which lies over the over the syntactic model of MTT built around  $\mathbf{Ty}_m$  and  $\mathbf{Tm}_m$  introduced in [Section 6.3](#). Concretely, this means that we must construct a series of constants ( $\mathbf{Ty}_m^*$ ,  $\mathbf{Tm}_m^*$ ,  $\mathbf{Prod}^*$ , etc.) in  $\mathcal{G}[-]$  such that under  $z : \mathbf{syn}$  these constants are equal to their corresponding syntactic components ( $\mathbf{Ty}_m(z)$ ,  $\mathbf{Tm}_m(z, -)$ ,  $\mathbf{Prod}(z)$ , etc.). This family of constants will define a model in  $\mathcal{G}[-]$  and the ‘alignment’ condition ensures that there is a morphism of models ([Theorem 7.9](#)). See [Appendix B](#) for the full list of constants. Throughout this section we will use extensional MTT as an internal language for  $\mathcal{G}[-]$  ([Theorem 6.1](#)) to define these constants. The strict equations will follow from repeated applications of the realignment theorem ([Theorem 6.5](#)). Many of these computations will be familiar to readers experienced with STC. The main novelties are modal connectives [Lemmas 7.5](#) and [7.6](#).

**Definition 7.1.** We define  $\mathbf{Ty}_m^*$  as the realignment of the following over  $\mathbf{Ty}_m$ :

$$\begin{aligned} \mathbf{record } T : \{ \mathbf{U}_2 \mid z : \mathbf{syn} \mapsto \mathbf{Ty}_m \} \text{ where} \\ \mathfrak{A} : \mathbf{NfTy}_m \\ \Phi : \{ \mathbf{U}_1 \mid z : \mathbf{syn} \mapsto \mathbf{Tm}_m(z, \mathfrak{A}) \} \\ \uparrow : \{ \mathbf{Ne}_m(\odot \mathfrak{A}) \rightarrow \Phi \mid z : \mathbf{syn} \mapsto \mathbf{id} \} \\ \downarrow : \{ \Phi \rightarrow \mathbf{Nf}_m(\odot \mathfrak{A}) \mid z : \mathbf{syn} \mapsto \mathbf{id} \} \end{aligned}$$

Taking advantage of the isomorphism  $\mathbf{Ty}_m^* \cong T$ , we construct elements of  $\mathbf{Ty}_m^*$  by specifying  $\mathfrak{A}$ ,  $\Phi$ ,  $\uparrow$ , and  $\downarrow$ . By realignment,  $\eta(\{\mathfrak{A}; \Phi; \downarrow; \uparrow\}) = \eta(\mathfrak{A})$ .

**Definition 7.2.** We define  $A : \mathbf{Ty}_m^* \vdash \mathbf{Tm}_m^*(A) @ m$  by  $\mathbf{Tm}_m^*(\{-; \Phi; -, -\}) = \Phi$ .

**Lemma 7.3.**  $\mathbf{Tm}_m^*$  lies strictly over  $\mathbf{Tm}_m$ . Explicitly for each  $A : \mathbf{Ty}_m^* = \mathbf{Ty}_m(z)$ ,  $\mathbf{Tm}_m^*(A) = \mathbf{Tm}_m(z, A)$ .

*Proof.* Using the boundary condition on  $\Phi$ , we compute  $\mathbf{Tm}_m^*(A) = \Phi_A = \mathbf{Tm}_m(z, \mathfrak{A}_A)$ . Under the hypothesis  $z : \mathbf{syn}$ , we obtain an equality  $A = \mathfrak{A}_A$  because  $\eta(A) = \eta(\mathfrak{A}_A)$  by definition. The conclusion now follows.  $\square$

**Lemma 7.4.** Fixing  $T_0 : \mathbf{Ty}_m^*$  and  $T_1 : \mathbf{Tm}_m^*(T_0) \rightarrow \mathbf{Ty}_m^*$ , there exists a pair of constants:

$$\begin{aligned} \mathbf{Sig}^* : \{ \mathbf{Ty}_m^* \mid z : \mathbf{syn} \mapsto \mathbf{Sig}(z, T_0, T_1) \} \\ \alpha_{\mathbf{Sig}^*} : \left\{ \mathbf{Tm}_m(\mathbf{Sig}^*(T_0, T_1)) \cong \sum_{t : \mathbf{Tm}_m^*(T_0)} \mathbf{Tm}_m^*(T_1(t)) \mid z : \mathbf{syn} \mapsto \alpha_{\mathbf{Sig}}(z, T_0, T_1) \right\} \end{aligned}$$

*Proof.* Let us start by apply [Theorem 6.5](#) to  $\alpha_{\mathbf{Sig}}(z)$  and  $\sum_{t_0 : \Phi_{T_0}} \Phi_{T_1(t_0)}$  to produce  $\Psi : \mathbf{U}_1$  such that  $z : \mathbf{syn} \vdash \Psi = \mathbf{Sig}(z, T_0, T_1)$  and  $\alpha_{\mathbf{Sig}^*} : \Psi \cong \sum_{t_0 : \Phi_{T_0}} \Phi_{T_1(t_0)}$  which restricts to  $\alpha_{\mathbf{Sig}}$  under  $z : \mathbf{syn}$ .

We now define  $\mathbf{Sig}^*(T_0, T_1)$  as follows:

$$\begin{aligned} \mathfrak{A}_{\mathbf{Sig}^*(T_0, T_1)} &= \mathbf{Sum}(\mathfrak{A}_{T_0}, \lambda v. \mathfrak{A}_{T_1(\uparrow_{T_0} v)}) \\ \Phi_{\mathbf{Sig}^*(T_0, T_1)} &= \Psi \end{aligned}$$

$$\begin{aligned}\uparrow_{\text{Sig}^*(T_0, T_1)} &= \lambda e. \alpha_{\text{Sig}^*}^{-1} \langle \uparrow_{T_0}(\mathbf{proj}_0(e)), \uparrow_{T_1(\uparrow_{T_0}(\mathbf{proj}_0(e)))}(\mathbf{proj}_1(e)) \rangle \\ \downarrow_{\text{Sig}^*(T_0, T_1)} &= \lambda t. \mathbf{pair}(\downarrow_{T_0}(\alpha_{\text{Sig}^*}(t)_0), \downarrow_{T_1(\alpha_{\text{Sig}^*}(t)_0)}(\alpha_{\text{Sig}^*}(t)_1))\end{aligned}$$

The fact that  $\downarrow$  and  $\uparrow$  lie over the identity follows directly from the  $\beta$  and  $\eta$  laws of dependent sums in MTT. We show the calculations for  $\uparrow$ . Fix  $z : \mathbf{syn}$ :

$$\begin{aligned}\uparrow_{\text{Sig}^*(T_0, T_1)}(e) &= \alpha_{\text{Sig}^*}^{-1} \langle \uparrow_{T_0}(\mathbf{proj}_0(e)), \uparrow_{T_1(\uparrow_{T_0}(\mathbf{proj}_0(e)))}(\mathbf{proj}_1(e)) \rangle \\ &= \alpha_{\text{Sig}^*}^{-1} \langle \mathbf{proj}_0(e), \mathbf{proj}_1(e) \rangle \\ &= \alpha_{\text{Sig}^*}^{-1} \langle \alpha_{\text{Sig}(T_0, T_1)}(e)_0, \alpha_{\text{Sig}(T_0, T_1)}(e)_1 \rangle \\ &= e\end{aligned}$$

The fact that  $\mathfrak{A}_{\text{Sig}(T_0, T_1)}$  and  $\Phi_{\text{Sig}^*(T_0, T_1)}$  lie over  $\text{Sig}(T_0, T_1)$  and  $\mathsf{Tm}_m(z, \text{Sig}(z, T_0, T_1))$  respectively follows from definition and realignment.  $\square$

**Lemma 7.5.** *Fixing  $T_0 : (\mu \mid \mathsf{Ty}_n^*)$  and  $T_1 : (\mu \mid \mathsf{Tm}_n^*(T_0)) \rightarrow \mathsf{Ty}_m^*$ , there exists a pair of constants:*

$$\begin{aligned}\text{Prod}^* &: \{\mathsf{Ty}_m \mid z : \mathbf{syn} \mapsto \text{Prod}^*(z, T_0, T_1)\} \\ \alpha_{\text{Prod}^*} &: \left\{ \mathsf{Tm}_m(\text{Prod}(T_0, T_1)) \cong \prod_{t: \langle \mu \mid \mathsf{Tm}_n^*(T_0) \rangle} \mathsf{Tm}_m^*(T_1(t)) \mid z : \mathbf{syn} \mapsto \alpha_{\text{Prod}}(z, T_0, T_1) \right\}\end{aligned}$$

*Proof.* Apply [Theorem 6.5](#) to  $\alpha_{\text{Prod}(z)}$  and  $(\mu \mid t_0 : \Phi_{T_0}) \rightarrow \Phi_{T_1(t_0)}$  to produce  $\Psi : \mathsf{U}_1$  such that  $z : \mathbf{syn} \vdash \Psi = \text{Prod}(z, T_0, T_1)$  and  $\alpha_{\text{Prod}^*} : \Psi \cong (\mu \mid t_0 : \Phi_{T_0}) \rightarrow \Phi_{T_1(t_0)}$  which restricts to  $\alpha_{\text{Prod}}$  under  $z : \mathbf{syn}$ . We now define  $\text{Prod}(T_0, T_1)$  as follows:

$$\begin{aligned}\mathfrak{A}_{\text{Prod}^*(T_0, T_1)} &= \mathbf{Prod}(\mathfrak{A}_{T_0}, \lambda v. \mathfrak{A}_{T_1(\uparrow_{T_0} v)}) \\ \Phi_{\text{Prod}^*(T_0, T_1)} &= \Psi \\ \uparrow_{\text{Prod}^*(T_0, T_1)} &= \lambda e. \alpha_{\text{Prod}}^{-1}(\lambda t. \uparrow_{T_1(t)}(\mathbf{app}(e)(\downarrow_{T_0}(t)))) \\ \downarrow_{\text{Prod}^*(T_0, T_1)} &= \lambda t. \mathbf{lam}(\lambda v. \uparrow_{T_1(\uparrow_{T_0}(v))}(\alpha_{\text{Prod}}(t)(\uparrow_{T_0}(v))))\end{aligned}$$

Once again, the fact that everything lies over the correct terms follows from the  $\beta$  and  $\eta$  laws for dependent products, as well as the conclusions of realignment.  $\square$

**Lemma 7.6.** *Fixing  $T : (\mu \mid \mathsf{Ty}_n^*)$ , the following constants exist:*

$$\begin{aligned}\text{Mod}^*[\mu] &: \{\mathsf{Ty}_m \mid z : \mathbf{syn} \mapsto \text{Mod}[\mu](z, T)\} \\ \mathbf{m}^*[\mu] &: (\mu \mid a : \mathsf{Tm}_n^*(T)) \rightarrow \{\mathsf{Tm}_m^*(\text{Mod}^*[\mu](T)) \mid z : \mathbf{syn} \mapsto \mathbf{m}[\mu](z, T, a)\} \\ \text{letmod}^*[\mu; \nu] &: [B : \langle \nu \mid \mathsf{Tm}_n^*(\text{Mod}^*[\mu](T)) \rangle \rightarrow \mathsf{Ty}_o^*] \\ &\rightarrow [b : (\nu \circ \mu \mid x : \mathsf{Tm}_n^*(T)) \rightarrow B(\mathbf{m}^*[\mu](T, x))] \\ &\rightarrow (\nu \mid a : \mathsf{Tm}_m^*(\text{Mod}^*[\mu](T))) \\ &\rightarrow \{\mathsf{Tm}_o^*(B(a)) \mid z : \mathbf{syn} \mapsto \text{letmod}[\mu; \nu](z, T, B, b, a)\} \\ - : [B : \langle \nu \mid \mathsf{Tm}_n^*(\text{Mod}^*[\mu](T)) \rangle \rightarrow \mathsf{Ty}_o^*] \\ &\rightarrow [b : (\nu \circ \mu \mid x : \mathsf{Tm}_n^*(T)) \rightarrow \mathsf{Tm}_o^*(B(\mathbf{m}^*[\mu](T, x)))] \\ &\rightarrow (\nu \circ \mu \mid a : \mathsf{Tm}_n^*(T)) \\ &\rightarrow \text{letmod}^*[\mu; \nu](B, b, \mathbf{m}^*[\mu](a)) = b(a)\end{aligned}$$

*Proof.* As is typical now, we will proceed by realignment. We define  $\Phi$  to be the realignment of the following type along the subsequent isomorphism:

$$\Psi = \sum_{m:\text{Nf}_m(\text{Mod}[\mu](T))} \bullet \left[ \left( \sum_{e:\text{Ne}_m(\text{Mod}[\mu](T))} \mathbf{up}(e) = m \right) + \left( \sum_{a:\langle \mu | \Phi_T \rangle} \mathbf{mod}_\mu(\downarrow_T a) = m \right) \right]$$

$$\alpha_\circ : \circ \Psi \cong \sum_{m:\circ \text{Nf}_m(\text{Mod}[\mu](T))} \mathbf{1} \cong \circ_z \mathbf{Tm}_m(z, \text{Mod}[\mu](z, T))$$

This gives a type  $\Phi$  and an isomorphism  $\alpha : \Phi \cong \Psi$  such that  $\alpha^{-1}(\langle m, \dots \rangle) = m$  in a context with  $z : \mathbf{syn}$ .

We now use this to define  $\text{Mod}^*[\mu](T)$ :

$$\begin{aligned} \mathfrak{A}_{\text{Mod}^*[\mu](T)} &= \mathbf{Mod}_\mu(\mathfrak{A}_T) \\ \Phi_{\text{Mod}^*[\mu](T)} &= \Phi \\ \uparrow_{\text{Mod}^*[\mu](T)} &= \lambda e. \alpha^{-1}(\langle \mathbf{up}(e), \eta(\text{in}_0(\langle e, \star \rangle)) \rangle) \\ \downarrow_{\text{Mod}^*[\mu](T)} &= \pi_0 \circ \alpha \end{aligned}$$

Unlike with dependent sums and products, however, this is not the end of the story. We must also define the introduction and elimination forms for this type. First, the intro form:

$$\mathbf{m}^*[\mu](T, m) = \alpha^{-1}(\langle \mathbf{mod}_\mu(\downarrow_T(m)), \eta(\text{in}_1(m, \star)) \rangle)$$

In order to define the elimination form, we need the elimination principle for  $\bullet A$ . Recall  $\bullet A = \mathbf{syn} \star A$  so we may use the induction principle of a pushout on  $\bullet A$ .

With this in mind, we define the elimination rule  $\text{letmod}^*[\mu; \nu](T, T_m, b, s)$  as follows. First, recall that  $s : \langle \nu | \mathbf{Tm}_m^*(\text{Mod}^*[\mu](T)) \rangle$ . We use the fact that  $\langle \nu | - \rangle$  preserves dependent sums, the closed modality (Theorem 6.1), and coproducts (Theorem 6.2) to decompose  $s$  into two terms:

$$\begin{aligned} m &: \langle \nu | \text{Nf}_n(\text{Mod}[\mu](T)) \rangle \\ q &: \bullet \left[ \sum_{e:\langle \nu | \text{Ne}_m(\text{Mod}[\mu](T)) \rangle} \langle \nu | \mathbf{up}(e) = m \rangle + \sum_{a:\langle \nu \circ \mu | \Phi_T \rangle} \langle \nu | \mathbf{mod}_\mu(\downarrow_T a) = m \rangle \right] \end{aligned}$$

We may now perform induction on  $q$ :

$$\begin{cases} \text{letmod}[\mu; \nu](z, T, T_m, b, s) & q = \text{in}_0(z) \\ \downarrow_{T_m(s)} \text{letmod}_{\mu; \nu}(T, T_m, \lambda x. \downarrow_{T_m(\mathfrak{m}[\mu](\uparrow_T x))} b(\uparrow_T x), e) & q = \text{in}_1(\text{in}_0(\text{mod}_\nu(e), -)) \\ b(a) & q = \text{in}_1(\text{in}_1(\text{mod}_\nu(a), -)) \end{cases}$$

The fact that these agree on overlaps follows from the  $\beta$  rule for modal types from MTT. By construction, the elimination and introduction forms lie over their syntactic counterparts, and calculation shows that the  $\beta$  equation holds.  $\square$

**Lemma 7.7.** *The following constants exist:*

$$\begin{aligned} \text{Bool}^* &: \{\text{Ty}_m^* \mid z : \mathbf{syn} \mapsto \text{Bool}(z)\} \\ \text{true}^* &: \{\text{Tm}_m^*(\text{Bool}) \mid z : \mathbf{syn} \mapsto \text{true}\} \\ \text{false}^* &: \{\text{Tm}_m^*(\text{Bool}) \mid z : \mathbf{syn} \mapsto \text{false}\} \\ \text{if}^* &: (T : \text{Tm}_m^*(\text{Bool}(z)) \rightarrow \text{Ty}_m^*) \rightarrow \\ &\quad \text{Tm}_m^*(T(\text{true}^*)) \rightarrow \text{Tm}_m^*(T(\text{false}^*)) \rightarrow (b : \text{Tm}_m^*(\text{Bool}^*)) \rightarrow \\ &\quad \{\text{Tm}_m^*(T(b)) \mid z : \mathbf{syn} \mapsto \text{if}(T, t, f, b)\} \\ - &: (T : \text{Tm}_m^*(\text{Bool}^*) \rightarrow \text{Ty}_m^*)(t : \text{Tm}_m^*(T(\text{true}^*))(f : \text{Tm}_m^*(T(\text{false}^*))) \rightarrow \\ &\quad (\text{if}^*(T, t, f, \text{true}^*) = t) \times (\text{if}^*(T, t, f, \text{false}^*) = f) \end{aligned}$$



*Proof.* The proof proceeds much as [Lemma 7.6](#), so we present the relevant definitions without much commentary. First, we realign  $\Psi$  along  $\alpha_\circ$ :

$$\begin{aligned}\Psi &= \sum_{m:\text{Nf}_m(\text{Bool}^*)} \bullet \left[ \left( \sum_{e:\text{Ne}_m(\text{Bool})} \mathbf{up}(e) = m \right) + \left( \sum_{b:\mathbf{2}} m = \mathbf{rec}_2(b; \mathbf{tt}; \mathbf{ff}) \right) \right] \\ \alpha_\circ : \circ_z \Psi &\cong \sum_{m:\text{Nf}_m(\text{Bool}^*)} \mathbf{1} \cong \mathbf{Tm}_m(z, \text{Bool}(z))\end{aligned}$$

As a result, we obtain  $\Phi$  along with  $\alpha : \Phi \cong \Psi$  such that  $z : \mathbf{syn} \vdash \alpha^{-1}\langle m, \dots \rangle = m$ . We may now define  $\text{Bool}^*$ :

$$\begin{aligned}\mathfrak{A}_{\text{Bool}^*} &= \mathbf{Bool} \\ \Phi_{\text{Bool}^*} &= \Phi \\ \uparrow_{\text{Bool}^*} &= \lambda e. \alpha^{-1}\langle \mathbf{up}(e), \eta(\mathbf{in}_0(e, \star)) \rangle \\ \downarrow_{\text{Bool}^*} &= \pi_0 \circ \alpha\end{aligned}$$

The true and false constants are defined as follows:

$$\begin{aligned}\mathbf{true}^* &= \langle \mathbf{tt}, \eta(\mathbf{in}_1(0, \star)) \rangle \\ \mathbf{false}^* &= \langle \mathbf{ff}, \eta(\mathbf{in}_1(1, \star)) \rangle\end{aligned}$$

Again, the elimination principle is defined using the induction principle for  $\bullet A$ .

$$\begin{aligned}\mathbf{if}^*(T_m, t_0, t_1, s = \alpha^{-1}\langle b, q \rangle) &= \\ \begin{cases} \mathbf{if}(z, T_m, t_0, t_1, m) & q = \mathbf{in}_0(z) \\ \downarrow_{T_m(s)} \mathbf{if}(T_m, \downarrow_{T_m(\mathbf{true}^*)} t_0, \downarrow_{T_m(\mathbf{false}^*)} t_1, e) & q = \mathbf{in}_1(\mathbf{in}_0(e, \star)), \mathbf{up}(e) = b \\ t_i & q = \mathbf{in}_1(\mathbf{in}_1(i, \star)), \mathbf{rec}(i; \mathbf{tt}; \mathbf{ff}) = b \end{cases}\end{aligned}$$

The boundary conditions and the computation rule follow from computation.  $\square$

**Lemma 7.8.** *The glued model supports a universe. Specifically, this means that the following constants exist:*

$$\begin{aligned}\mathbf{Uni}^* &: \{\mathbf{Ty}_m^* \mid z : \mathbf{syn} \mapsto \mathbf{Uni}\} \\ \mathbf{El}^* &: (T : \mathbf{Tm}_m^*(\mathbf{Uni}^*)) \rightarrow \{\mathbf{Ty}_m^* \mid z : \mathbf{syn} \mapsto \mathbf{El}(T)\}\end{aligned}$$

Moreover,  $\mathbf{Uni}^*$  is closed under various type-formers:

$$\begin{aligned}\widehat{\mathbf{Sig}}^* &: \left\{ \left[ \sum_{A:\mathbf{Tm}_m^*(\mathbf{Uni}^*)} \mathbf{Tm}_m^*(\mathbf{El}^*(A)) \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}) \right] \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*) \mid z : \mathbf{syn} \mapsto \widehat{\mathbf{Sig}} \right\} \\ \widehat{\mathbf{Prod}}^* &: \left\{ \left[ \sum_{A:\langle \mu \mid \mathbf{Tm}_n^*(\mathbf{Uni}^*) \rangle} [\langle \mu \mid \mathbf{Tm}_n^*(\mathbf{El}^*(A)) \rangle] \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*) \right] \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*) \mid z : \mathbf{syn} \mapsto \widehat{\mathbf{Prod}} \right\} \\ \widehat{\mathbf{Bool}}^* &: \{\mathbf{Tm}_m^*(\mathbf{Uni}^*) \mid z : \mathbf{syn} \mapsto \widehat{\mathbf{Bool}}\} \\ \widehat{\mathbf{Mod}}^* &: \{\langle \mu \mid \mathbf{Tm}_n^*(\mathbf{Uni}^*) \rangle \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*) \mid z : \mathbf{syn} \mapsto \widehat{\mathbf{Mod}}\} \\ \mathbf{dec}_{\widehat{\mathbf{Sig}}}^* &: (A : \mathbf{Tm}_m^*(\mathbf{Uni}^*)) (B : \mathbf{Tm}_m^*(\mathbf{El}^*(A)) \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*)) \rightarrow \\ &\quad \mathbf{Tm}_m^*(\mathbf{El}^*(\widehat{\mathbf{Sig}}^*(A, B))) \cong \mathbf{Tm}_m^*(\mathbf{Sig}^*(\mathbf{El}^*(A), \mathbf{El}^* \circ B)) \\ \mathbf{dec}_{\widehat{\mathbf{Prod}}}^* &: (\mu \mid A : \mathbf{Tm}_m^*(\mathbf{Uni}^*)) (B : (\mu \mid \mathbf{Tm}_n^*(\mathbf{El}^*(A))) \rightarrow \mathbf{Tm}_m^*(\mathbf{Uni}^*)) \rightarrow \\ &\quad \mathbf{Tm}_m^*(\mathbf{El}^*(\widehat{\mathbf{Prod}}^*(A, B))) \cong \mathbf{Tm}_m^*(\mathbf{Prod}^*(\mathbf{El}^*(A), \mathbf{El}^* \circ B)) \\ \mathbf{dec}_{\widehat{\mathbf{Bool}}}^* &: \mathbf{Tm}_m^*(\mathbf{El}^*(\widehat{\mathbf{Bool}}^*)) \cong \mathbf{Tm}_m^*(\mathbf{Bool}^*) \\ \mathbf{dec}_{\widehat{\mathbf{Mod}}}^* &: (\mu \mid A : \mathbf{Tm}_m^*(\mathbf{Uni}^*)) \rightarrow \mathbf{Tm}_m^*(\mathbf{El}^*(\widehat{\mathbf{Mod}}^*(A))) \cong \mathbf{Tm}_m^*(\mathbf{Mod}^*[\mu](\mathbf{El}^*(A)))\end{aligned}$$

Furthermore, when given  $z : \mathbf{syn}$  we require the following equations:

$$\widehat{\text{dec}}_{\widehat{\text{Sig}}}^* = \text{dec}_{\widehat{\text{Sig}}}(z) \quad \widehat{\text{dec}}_{\widehat{\text{Prod}}}^* = \text{dec}_{\widehat{\text{Prod}}}(z) \quad \widehat{\text{dec}}_{\widehat{\text{Bool}}}^* = \text{dec}_{\widehat{\text{Bool}}}(z) \quad \widehat{\text{dec}}_{\widehat{\text{Mod}}}^* = \text{dec}_{\widehat{\text{Mod}}}(z)$$

*Proof.* At this point we take advantage of the fact that  $\Phi_{\text{Ty}_m^*}$  is an element of  $\mathbf{U}_1$ ; in particular, we use the fact that is a universe  $\mathbf{U}_0$  small enough to fit inside  $\mathbf{U}_1$ .

*The universe and  $\text{El}(-)$  constant.* We may then define  $\Psi$  by realigning the following element of  $\mathbf{U}_1$  along the evident isomorphism to  $\text{Tm}_m^*(z, \text{Uni}(z))$ :

$$(3) \quad \begin{aligned} &\mathbf{record} \ T : \mathbf{U}_1 \ \mathbf{where} \\ &\quad \mathfrak{A} : \mathbf{Nf}_m(\mathbf{Uni}) \\ &\quad \Phi : \{\mathbf{U}_0 \mid z : \mathbf{syn} \mapsto \text{Tm}_m(z, \text{El}(z, \mathfrak{A}))\} \\ &\quad \uparrow : \{\text{Ne}_m(\circ_z \text{El}(z, \mathfrak{A})) \rightarrow \Phi \mid z : \mathbf{syn} \mapsto \text{id}\} \\ &\quad \downarrow : \{\Phi \rightarrow \mathbf{Nf}_m(\circ_z \text{El}(z, \mathfrak{A})) \mid z : \mathbf{syn} \mapsto \text{id}\} \end{aligned}$$

As a result, we obtain  $\Psi : \mathbf{U}_1$  and  $\alpha : \Psi \cong T$  such that  $\alpha^{-1}(\{\mathfrak{A}; \dots\}) = \mathfrak{A}$  given  $z : \mathbf{syn}$ . With  $\Psi$  in hand, we may define  $\text{Uni}^*$ :

$$\begin{aligned} \mathfrak{A}_{\text{Uni}^*} &= \mathbf{Uni} \\ \Phi_{\text{Uni}^*} &= \Psi \\ \uparrow_{\text{Uni}^*} &= \lambda e. \alpha^{-1}(\langle \mathbf{up}(e); \text{Ne}_m; \text{id}; \lambda e. \mathbf{up}(e) \rangle) \\ \downarrow_{\text{Uni}^*} &= \pi_0 \circ \alpha \end{aligned}$$

The definition of  $\text{El}^*$  is essentially “just” cumulativity:

$$\text{El}^*(\alpha^{-1}(\langle \mathfrak{A}; \Phi; \downarrow; \uparrow \rangle)) = \alpha(\langle \text{El}(\mathfrak{A}); \Phi; \downarrow; \uparrow \rangle)$$

Here we have used both the isomorphism realigning  $\text{Uni}^*$  and the isomorphism realigning  $\text{Ty}_m^*$ . When restricting with  $z : \mathbf{syn}$ ,  $\text{El}^*$  restricts to  $A \mapsto \text{El}(A)$ .

It remains to show that the universe is closed under the various codes for dependent sums, products, etc. For the sake of space, we show only two of these cases: dependent products and modal types.

Let us fix  $t_0 : (\mu \mid \text{Tm}_n^*(\text{Uni}^*))$  and  $t_1 : (\mu \mid \text{Tm}_n^*(\text{El}^*(t_0))) \rightarrow \text{Tm}_m^*(\text{Uni}^*)$ . First we realign  $\text{Tm}_m^*(\text{Prod}^*(\text{El}^*(A), \text{El}^*(B)))$  along  $\widehat{\text{dec}}_{\widehat{\text{Prod}}}$  to obtain  $\Psi$  which lies strictly over  $\text{Tm}_m^*(z, \text{Prod}(z, \text{El}(z, t_0), \text{El}(z) \circ t_1))$ .

We also get an isomorphism  $\widehat{\text{dec}}_{\widehat{\text{Prod}}}^* : \Psi \cong \text{Tm}_m^*(\text{Prod}^*(\text{El}^*(A), \text{El}^*(B)))$  lying over  $\widehat{\text{dec}}$ . Note that by the uniqueness of inverses,  $(\widehat{\text{dec}}^*)^{-1}$  lies over  $\widehat{\text{dec}}^{-1}$ .

We use this now to define  $\widehat{\text{Prod}}^*(t_0, t_1)$ :

$$\begin{aligned} \mathfrak{A}_{\widehat{\text{Prod}}^*(t_0, t_1)} &= (\mu \mid \mathfrak{A}_{t_0}) \widehat{\rhd} \lambda x. \mathfrak{A}_{t_1(\uparrow_{t_0}(x))} \\ \Phi_{\widehat{\text{Prod}}^*(t_0, t_1)} &= \Psi \\ \uparrow_{\widehat{\text{Prod}}^*(t_0, t_1)} &= \lambda e. (\widehat{\text{dec}}_{\widehat{\text{Prod}}}^*)^{-1}(\uparrow_{\text{Prod}^*(\text{El}^*(t_0), \text{El}^*(t_1))} \widehat{\text{dec}}^\triangleright(e)) \\ \downarrow_{\widehat{\text{Prod}}^*(t_0, t_1)} &= \lambda f. \widehat{\text{dec}}^\triangleleft(\downarrow_{\text{Prod}^*(\text{El}^*(t_0), \text{El}^*(t_1))} \widehat{\text{dec}}_{\widehat{\text{Prod}}}^*(f)) \end{aligned}$$

The procedure for modal types is similar. Let us fix  $t : (\mu \mid \text{Tm}_n^*(\text{Uni}^*))$ . Again, we realign  $\text{Tm}_m^*(\text{Mod}^*[\mu](\text{El}^*(t)))$  along the isomorphism  $\widehat{\text{dec}}_{\widehat{\text{Mod}}}$  to obtain  $\Psi$  and  $\widehat{\text{dec}}_{\widehat{\text{Mod}}[\mu]}^*$ . The actual construction of  $\widehat{\text{Mod}}^*$  is almost identical to  $\widehat{\text{Prod}}^*$ :

$$\begin{aligned} \mathfrak{A}_{\widehat{\text{Mod}}^*(t)} &= \langle \mu \mid \mathfrak{A}_t \rangle \\ \Phi_{\widehat{\text{Mod}}^*(t)} &= \Psi \end{aligned}$$

$$\begin{aligned}\uparrow_{\widehat{\text{Mod}}^*(t)} &= \lambda e. (\text{dec}_{\widehat{\text{Mod}}}^*)^{-1}(\uparrow_{\text{Mod}^*[\mu](\text{El}^*(t))} \text{dec}^\triangleright(e)) \\ \downarrow_{\widehat{\text{Mod}}^*(t)} &= \lambda m. \text{dec}^\triangleleft(\downarrow_{\text{Prod}^*(\text{El}^*(t))} \text{dec}_{\widehat{\text{Mod}}}^*(m))\end{aligned}$$

The checks that all constructions lie over their syntactic counterparts follow immediately from the conclusions of realignment.  $\square$

**Theorem 7.9.** *There exists an MTT cosmos built around  $\text{Ty}_m^*$  and  $\text{Tm}_m^*$  which lies strictly over the syntactic MTT cosmos. Externally, there is an MTT cosmos  $\mathcal{G}[-] : \mathcal{M} \rightarrow \mathbf{Gl}(\mathbf{i}[-]^*)$  together with a morphism of cosmoi  $\pi : \mathcal{G}[-] \rightarrow \mathcal{S}[-]$ .*

*Proof.* This is combination of [Definitions 7.1](#) and [7.2](#) and [Lemmas 7.3](#) to [7.8](#). That  $\pi$  is 2-natural is immediate from [Theorem 4.1](#) and that it commutes with transposition is precisely [Theorem 4.2](#).  $\square$

## 8. THE NORMALIZATION FUNCTION

**8.1. Initiality, revisited.** As formulated, the cosmos based around the generalized algebraic syntax of MTT  $\mathcal{S}[-]$  is not initial. By [Corollary 3.9](#), there is an initial cosmos  $\mathcal{I}[-]$  but it is a distinct object and  $\mathcal{I}[m]$  is not even cocomplete, let alone a presheaf topos.

Despite this,  $\mathcal{S}[-]$  enjoys a certain distinguished place among MTT cosmoi owing to [Theorem 3.7](#). In particular, we have the following

- (1) For every context  $\Gamma \text{ cx} @ m$ , there exists an object  $\mathcal{G}[\Gamma] : \mathcal{G}[m]$ , together with a canonical isomorphism  $\alpha_\Gamma : \pi(\mathcal{G}[\Gamma]) \cong \mathbf{y}(\Gamma)$ .
- (2) For every type  $\Gamma \vdash A @ m$ , there is a morphism  $\mathcal{G}[A] : \mathcal{G}[\Gamma] \rightarrow \text{Ty}_m^*$  such that  $\pi(\mathcal{G}[A]) \circ \alpha_\Gamma = [A]$ .
- (3) For every term  $\Gamma \vdash M : A @ m$ , there is a morphism  $\mathcal{G}[M] : \mathcal{G}[M] \rightarrow \text{Tm}_m^*$  lying over  $\mathcal{G}[A]$  such that  $\pi(\mathcal{G}[M]) \circ \alpha_\Gamma = [M]$ .

**8.2. The proof of the normalization theorem.** Prior to the proof of the normalization theorem ([Theorem 8.5](#)), we require several preliminary results.

Recall that there is a closed immersion of topoi  $\mathbf{i} : \mathbf{PSh}(\text{Ren}_m) \rightarrow \mathcal{G}[m]$ . In fact, this closed immersion is essential because the morphism  $\mathbf{PSh}(\text{Ren}_m) \rightarrow \mathbf{PSh}(\text{Cx}_m)$  is essential. Accordingly, for each  $X : \mathbf{PSh}(\text{Ren}_m)$  there is an object  $\mathbf{i}_!(X) : \mathcal{G}[m]$ . This functor has a familiar result when applied to a representable in  $\mathbf{PSh}(\text{Ren}_m)$ . Consider a telescope  $\Theta \text{ tele} @ m \rightsquigarrow \Gamma$ , we then have the following:

$$(\downarrow \Theta) \triangleq \mathbf{i}_!(\mathbf{y}(\Theta)) = \mathbf{y}(\Theta) \rightarrow \mathbf{i}[m]^*(\mathbf{y}(\Gamma))$$

**Lemma 8.1.** *For an arbitrary  $X : \mathcal{G}[m]$ , we have  $[(\downarrow \Theta), X] \cong \mathbf{i}^*(X)(\Theta)$ .*

By the previous discussion, for a telescope  $\Theta \text{ tele} @ m \rightsquigarrow \Gamma$ , there is also the object  $\mathcal{G}[\Gamma]$ , which lies over  $\Gamma$  in  $\mathcal{G}[m]$ . While this object is different than  $(\downarrow \Theta)$ , there is a canonical natural transformation between them.

**Lemma 8.2.** *Given  $\Theta \text{ tele} @ m \rightsquigarrow \Gamma$ , there is a morphism  $\text{atoms}_\Theta : (\downarrow \Theta) \rightarrow \mathcal{G}[\Gamma]$ . Moreover, up to the canonical isomorphism  $\mathbf{j}^* \mathcal{G}[\Gamma] \cong \mathbf{y}(\Gamma)$ ,  $\mathbf{j}^* \text{atoms}_\Theta = \text{id}_{\mathbf{y}(\Gamma)}$ .*

*Proof.* By applying [Lemma 8.1](#), this proof is equivalent to constructing an element of  $g \in \mathbf{i}^* \mathcal{G}[\Gamma](\Theta)$  such that the image of  $g$  under the canonical map to  $\mathbf{i}[m]^*(\mathbf{y}(\Gamma))(\Theta)$  is the identity. This morphism is constructed by induction on  $\Theta$ .

Case:

$$\frac{}{\mathbf{1} \text{ tele } @ m \rightsquigarrow \mathbf{1}}$$

In this case these two are isomorphic.

Case:

$$\frac{\mu : n \longrightarrow m \quad \Theta \text{ tele } @ m \rightsquigarrow \Gamma}{\Theta.\{\mu\} \text{ tele } @ n \rightsquigarrow \Gamma.\{\mu\}}$$

In this case we choose  $\text{atoms}_{\Theta.\{\mu\}} = \mathcal{G}[\mu]_!(\text{atoms}_{\Theta_0})$  and correcting by the canonical isomorphisms  $(\Theta_0.\{\mu\}) \cong \mathcal{G}[\mu]_!(\Theta)$  which exists by [Theorem 4.2](#) and  $\mathcal{G}[\Gamma.\{\mu\}] \cong \mathcal{G}[\mu]_!(\mathcal{G}[\Gamma])$  which exists by construction.

Case:

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n}{\Theta.(\mu \mid A) \text{ tele } @ m \rightsquigarrow \Gamma.(\mu \mid A)}$$

We construct this morphism using the characterization from [Lemma 8.1](#). Our induction hypothesis gives us an element  $g \in \mathcal{G}[\Gamma]_0(\Theta)$  lying over  $\alpha_\Gamma^{-1}(\text{id})$ , whence  $g' \in \mathcal{G}[\Gamma.(\mu \mid A)](\Theta.(\mu \mid A))$  lying over  $\uparrow$  using reindexing.

Moreover, we have an element of  $\text{Ne}_m(\mathcal{S}[A])(\Theta.(\mu \mid A).\{\mu\})$  lying over  $\mathbf{v}_0$ , whence an element of  $\text{Tm}_m^*(\mathcal{G}[A])(\Theta.(\mu \mid A).\{\mu\})$  lying over  $\mathbf{v}_0$ . This gives us an element of  $\mathcal{G}[\Gamma.(\mu \mid A)]_0(\Theta.(\mu \mid A))$  which lies over  $\text{id} = \uparrow.\mathbf{v}_0$  by the universal property of the pullback.  $\square$

**Lemma 8.3.** *Given  $\Gamma \vdash M, N : A @ m$ , if there exists a normal form  $u$  such that  $\Theta \models u : A @ m \rightsquigarrow M$  and  $\Theta \models u : A @ m \rightsquigarrow N$ , then  $\Gamma \vdash M = N : A @ m$ .*

**Lemma 8.4.** *If  $\Theta \models u @ m \rightsquigarrow A$  and  $\Theta \models u @ m \rightsquigarrow B$ , then  $\Gamma \vdash A = B @ m$ .*

We are now in a position to prove normalization. We note that one of the conditions of the normalization algorithm (respect for definitional equivalence) is automatic: we have defined the algorithm to operate only on equivalence classes of terms, so it must always respect definitional equivalence.

**Theorem 8.5.** *There exists a pair of functions  $\mathbf{nf}_\Gamma(M, A)$  and  $\mathbf{nfty}_\Gamma(A)$  such that*

- (1) *If  $\Gamma \vdash A @ m$  and  $\Theta \text{ tele } @ m \rightsquigarrow \Gamma$  then  $\Theta \models \mathbf{nfty}_\Gamma(A) @ m \rightsquigarrow A$ .*
- (2) *If  $\Gamma \vdash M : A @ m$  and  $\Theta \text{ tele } @ m \rightsquigarrow \Gamma$  then  $\Theta \models \mathbf{nf}_\Gamma(M, A) : A @ m \rightsquigarrow M$ .*

*Proof.* We may package this in the language of  $\mathbf{PSh}(\text{Ren}_m)$ . From this viewpoint,  $\mathbf{nfty}_\Gamma(A)$  is an element of  $(\mathbf{NfTy}_m)_0$  fitting into the following triangle:

$$\begin{array}{ccc} & & (\mathbf{NfTy}_m)_0 \\ & \nearrow & \downarrow \mathbf{NfTy}_m \\ \mathbf{i}[m]^*(\mathbf{y}(\Gamma)) & \xrightarrow{[A]} & \mathbf{i}[m]^*(\mathbf{y}(\Gamma)) \end{array}$$

Similarly,  $\mathbf{nf}_\Gamma(M, A)$  is a section to the inclusion  $\mathbf{Nf}_m \longrightarrow \mathbf{i}[m]^*(\tilde{\mathcal{T}}_m)$  over  $[M]$ .

Observe from the initiality of syntax that for any  $\Gamma \vdash A @ m$ , there exists  $\mathcal{G}[A] : \mathcal{G}[\Gamma] \longrightarrow \text{Ty}_m^*$ , and bi-universality ensures that  $\circ \mathcal{G}[A] = A$ , up to an isomorphism of contexts. In particular, there exists  $A^{\mathbf{nf}} : \mathcal{G}[\Gamma] \longrightarrow (\mathbf{NfTy}_m)_0$  such that  $\circ A^{\mathbf{nf}} = A$ .

Unfolding to  $\mathbf{PSh}(\mathbf{Ren}_m)$ , we have the following commuting square:

$$\begin{array}{ccc} \mathcal{G}[\Gamma]_0 & \xrightarrow{A^{\mathbf{nf}}} & (\mathbf{NfTy}_m)_0 \\ \mathcal{G}[\Gamma] \downarrow & & \downarrow \mathbf{NfTy}_m \\ \mathbf{i}[m]^*(\mathbf{y}(\Gamma)) & \xrightarrow{\mathbf{i}[m]^*(\lfloor A \rfloor)} & \mathbf{i}[m]^*(\mathcal{T}_m) \end{array}$$

By Lemma 8.2,  $\mathbf{atoms}_\Theta \in \mathcal{G}[\Gamma]_1(\Theta)$  lies over  $\text{id}$ . We then define  $\mathbf{nfty}_\Gamma(A) = A^{\mathbf{nf}}(\mathbf{atoms}_\Theta)$ , which satisfies the required properties.

The procedure is identical for  $\mathbf{nf}_\Gamma(M, A)$ . Given  $\Gamma \vdash M : A @ m$ , we use initiality to obtain the following square:

$$\begin{array}{ccc} \mathcal{G}[\Gamma]_0 & \xrightarrow{M^{\mathbf{nf}}} & (\mathbf{Nf}_m)_0 \\ \mathcal{G}[\Gamma] \downarrow & & \downarrow \mathbf{Nf}_m \\ \mathbf{i}[m]^*(\mathbf{y}(\Gamma)) & \xrightarrow{\mathbf{i}[m]^*(\lfloor M \rfloor)} & \mathbf{i}[m]^*(\tilde{\mathcal{T}}_m) \end{array}$$

We set  $\mathbf{nf}_\Gamma(M, A) = M^{\mathbf{nf}}(\mathbf{atoms}_\Theta)$ .  $\square$

**Theorem 8.6.** *Given a telescope  $\Theta \text{ tele } @ m \rightsquigarrow \Gamma$  then the following two facts hold:*

- (1) *If  $\Theta \models^{\mathbf{nf}} u : A @ m \rightsquigarrow M$ , then  $\mathbf{nf}_\Gamma(M, A) = u$ .*
- (2) *If  $\Theta \models^{\mathbf{nf}} \tau @ m \rightsquigarrow A$ , then  $\mathbf{nfty}_\Gamma(A) = \tau$ .*

*Proof.* In order to prove these results, we show three related facts. Recall that  $\mathcal{G}[-]$  is the function sending a piece of syntax to its interpretation in the normalization model. Furthermore, recall that by Lemma 8.2 ( $\lfloor \Theta \rfloor$  element  $\mathbf{atoms}_\Theta : \mathcal{G}[\Gamma]$ ).

- (1) If  $\Theta \models^e e : A @ m \rightsquigarrow M$ , then  $\mathcal{G}[M](\mathbf{atoms}_\Theta) = \uparrow_{\mathcal{G}[A](\mathbf{atoms}_\Theta)} e$
- (2) If  $\Theta \models^{\mathbf{nf}} u : A @ m \rightsquigarrow M$ , then  $\uparrow_{\mathcal{G}[A](\mathbf{atoms}_\Theta)} \mathcal{G}[M](\mathbf{atoms}_\Theta) = u$ .
- (3) If  $\Theta \models^{\mathbf{nf}} \tau @ m \rightsquigarrow A$ , then  $\mathfrak{A}_{\mathcal{G}[A](\mathbf{atoms}_\Theta)} = \tau$ .

Here we have identified a code  $u$  (resp.  $e$ ) as an  $\lfloor \Theta \rfloor$  element of  $\mathbf{Nf}_A$  (resp.  $\mathbf{Ne}_A$ ), an abuse justified by unfolding the definition of  $\mathbf{Nf}_A$  into  $\mathbf{PSh}(\mathbf{Ren}_m)$ .

Assuming these facts hold, the result immediately follows by unfolding the definition of  $\mathbf{nf}_\Gamma(M, A)$  and  $\mathbf{nfty}_\Gamma(A)$ . We prove these facts by mutual induction. All cases follow from induction except the case of variables, so we show this here.

**Case:**

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta = \Theta_0.(\mu \mid A)}{\Theta \models^e \mathbf{v}_0^{\text{id}_\mu} : A[\{\alpha\} \circ \uparrow. \{\mu\}] @ m \rightsquigarrow \mathbf{v}_0[\{\alpha\}]}$$

(Note that the case for a general  $\mathbf{v}_k^\alpha$  follows by straightforward induction on  $k$ , but is notationally heavy.)

First, we recall that  $\mathcal{G}[\mathbf{v}_0]$  defined by projection from the context. Therefore,  $\mathcal{G}[\mathbf{v}_0](\mathbf{atoms}_\Theta) = \uparrow_{\mathcal{G}[A](\mathbf{atoms}_\Theta)} \mathbf{v}_0^{\text{id}_\mu}$  by unfolding the construction of  $\mathbf{atoms}_\Theta$  in Lemma 8.2.

**Case:**

$$\frac{\Theta. \{\mu\} \models^{\mathbf{nf}} \tau @ n \rightsquigarrow A}{\Theta \models^{\mathbf{nf}} \langle \mu \mid \tau \rangle @ m \rightsquigarrow \langle \mu \mid A \rangle}$$

In this case, by unfolding the definitions we have

$$\mathfrak{A}_{\mathcal{G}[\langle \mu | A \rangle](\text{atoms}_{\Theta, \{\mu\}})} = \mathbf{Mod}_{\mu}(\mathfrak{A}_{A(\text{atoms}_{\Theta})})$$

By induction hypothesis, we have  $\mathfrak{A}_{A(\text{atoms}_{\Theta})} = \tau$ , so the conclusion follows.

**Case:**

$$\frac{\Theta.\{\mu\} \models^{\text{nf}} u : A @ n \rightsquigarrow M}{\Theta \models^{\text{nf}} \text{mod}_{\mu}(u) : \langle \mu | A \rangle @ m \rightsquigarrow \text{mod}_{\mu}(M)}$$

In this case, by unfolding the definitions we have

$$\begin{aligned} & \downarrow_{\mathcal{G}[\langle \mu | A \rangle](\text{atoms}_{\Theta, \{\mu\}})} \mathcal{G}[\text{mod}_{\mu}(M)](\text{atoms}_{\text{Theta}, \{\mu\}}) \\ &= \mathbf{mod}_{\mu} \downarrow_{\mathcal{G}[A](\text{atoms}_{\Theta})} \mathcal{G}[M](\text{atoms}_{\Theta}) \\ &= \mathbf{mod}_{\mu}(u) \end{aligned}$$

The last step follows from the induction hypothesis.  $\square$

**Corollary 8.7.** *Each term and type in MTT has a unique normal form.*

**Corollary 8.8.** *The conversion problem in MTT is equivalent the conversion problem of normal forms.*

*Proof.* **Theorem 8.5** and **Lemma 8.3** imply that  $\Gamma \vdash M = N : A @ m$  is equivalent to  $\mathbf{nf}_{\Gamma}(M, A) = \mathbf{nf}_{\Gamma}(N, A)$ . Similarly, **Theorem 8.5** and **Lemma 8.4** reduce  $\Gamma \vdash A = B @ m$  to  $\mathbf{nfty}_{\Gamma}(A) = \mathbf{nfty}_{\Gamma}(B)$ . As the proof given is constructively valid, these reductions are effective.  $\square$

**Corollary 8.9.** *If  $\Gamma \vdash (\mu | A_0) \rightarrow B_0 = (\mu | A_1) \rightarrow B_1 @ m$ , then  $\Gamma.\{\mu\} \vdash A_0 = A_1 @ m$  and  $\Gamma.(\mu | A_0) \vdash B_0 = B_1 @ m$ .*

*Proof.* By inspection on the definition of  $\mathbf{nfty}_{\Gamma}((\mu | A_i) \rightarrow B_i)$ , we observe that both are interpreted using **Lemma 7.5**. Accordingly, we have the following:

$$\begin{aligned} \mathbf{nfty}_{\Gamma}((\mu | A_0) \rightarrow B_0) &= u = (\mu | \mathbf{nfty}_{\Gamma.\{\mu\}}(A_0)) \rightarrow \mathbf{nfty}_{\Gamma.(\mu | A_0)}(B_0) \\ \mathbf{nfty}_{\Gamma}((\mu | A_1) \rightarrow B_1) &= v = (\mu | \mathbf{nfty}_{\Gamma.\{\mu\}}(A_1)) \rightarrow \mathbf{nfty}_{\Gamma.(\mu | A_1)}(B_1) \end{aligned}$$

From  $\Gamma \vdash (\mu | A_0) \rightarrow B_0 = (\mu | A_1) \rightarrow B_1 @ m$  and **Corollary 8.8**, we obtain  $u = v$ . From this, we use inversion to conclude that  $\mathbf{nfty}_{\Gamma.\{\mu\}}(A_0) = \mathbf{nfty}_{\Gamma.\{\mu\}}(A_1)$  and  $\mathbf{nfty}_{\Gamma.(\mu | A_0)}(B_0) = \mathbf{nfty}_{\Gamma.(\mu | A_1)}(B_1)$ . The result now follows from **Corollary 8.8**.  $\square$

**Corollary 8.10.** *If modalities and 2-cells enjoy decidable equality, typechecking MTT is decidable.*

#### ACKNOWLEDGMENTS

I am thankful for discussions with Carlo Angiuli, Martin Bidlingmaier, Lars Birkedal, Thierry Coquand, Alex Kavvos, Christian Sattler, and Jonathan Sterling. I would like to express particular gratitude to Jonathan Sterling for conversations about Synthetic Tait Computability and to Lars Birkedal and Alex Kavvos for feedback on this proof.

## APPENDIX A. NEUTRAL AND NORMAL FORMS

$$\boxed{\Theta \text{ tele } @ m \rightsquigarrow \Gamma}$$

$$\frac{}{1 \text{ tele } @ m \rightsquigarrow 1} \quad \frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n}{\Theta.(\mu \mid A) \text{ tele } @ m \rightsquigarrow \Gamma.(\mu \mid A)} \quad \frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma}{\Theta.\{\mu\} \text{ tele } @ m \rightsquigarrow \Gamma.\{n\}}$$

$$\boxed{\Theta \vdash \psi : \Psi @ m \rightsquigarrow \gamma}$$

$$\frac{}{\Theta \vdash ! : 1 @ m \rightsquigarrow !} \quad \frac{\Theta \text{ tele } @ n \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n}{\Theta.(\mu \mid A) \vdash \uparrow : \Theta @ m \rightsquigarrow \uparrow} \quad \frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma}{\Theta \vdash \text{id} : \Theta @ m \rightsquigarrow \text{id}}$$

$$\frac{\Gamma, \Delta, \Xi \text{ tele } @ m \rightsquigarrow \Theta, \Psi, \Phi \quad \Theta \vdash r : \Psi @ m \rightsquigarrow \gamma \quad \Psi \vdash s : \Phi @ m \rightsquigarrow \delta}{\Theta \vdash s \circ r : \Phi @ m \rightsquigarrow \delta \circ \gamma}$$

$$\frac{\Gamma, \Delta \text{ tele } @ m \rightsquigarrow \Theta, \Psi \quad \Theta \vdash r : \Psi @ m \rightsquigarrow \delta}{\Theta.\{\mu\} \vdash r.\{\mu\} : \Psi.\{\mu\} @ n \rightsquigarrow \delta.\{\mu\}}$$

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \mu, \nu : n \rightarrow m \quad \alpha : \nu \rightarrow \mu}{\Theta.\{\mu\} \vdash \{\alpha\}_\Theta : \Theta.\{\nu\} @ n \rightsquigarrow \{\alpha\}_\Gamma}$$

$$\frac{\Theta \vdash r : \Psi @ m \rightsquigarrow \delta \quad \Gamma, \Delta \text{ tele } @ m \rightsquigarrow \Theta, \Psi \quad \Delta.\{\mu\} \vdash A @ n \quad \Gamma.\{\mu\} \models^e \mathbf{v}_k^\alpha : A[\delta.\{\mu\}] @ n \rightsquigarrow M}{\Theta \vdash r.\mathbf{v}_k^\alpha : \Psi.(\mu \mid A) @ m \rightsquigarrow \delta.M}$$

$$\boxed{\Theta \models^f \tau @ m \rightsquigarrow A}$$

$$\frac{}{\Theta \models^f \text{bool} @ m \rightsquigarrow \text{bool} \quad \Theta \models^f \mathbf{U} @ m \rightsquigarrow \mathbf{U}}$$

$$\frac{\Theta.\{\mu\} \models^f \tau @ n \rightsquigarrow A \quad \Theta.(\mu \mid A) \models^f \sigma @ m \rightsquigarrow B}{\Theta \models^f (\mu \mid \tau) \rightarrow \sigma @ m \rightsquigarrow (\mu \mid A) \rightarrow B}$$

$$\frac{\Theta \models^f \tau @ n \rightsquigarrow A \quad \Theta.(\text{id}_m \mid A) \models^f \sigma @ m \rightsquigarrow B}{\Theta \models^f \tau \times \sigma @ m \rightsquigarrow A \times B} \quad \frac{\Theta.\{\mu\} \models^f \tau @ n \rightsquigarrow A}{\Theta \models^f \langle \mu \mid \tau \rangle @ m \rightsquigarrow \langle \mu \mid A \rangle}$$

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \mathbf{U} @ m \rightsquigarrow A}{\Theta \models^f \text{El}(u) @ m \rightsquigarrow A}$$

$$\boxed{\Theta \models^e e : A @ m \rightsquigarrow M \quad \Theta \models^f u : A @ m \rightsquigarrow M}$$

$$\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta(k) = (\mu \mid A) \quad \text{mods}(\Theta, k) = \nu \quad \alpha : \mu \rightarrow \nu}{\Theta \models^e \mathbf{v}_k^\alpha : A[\{\alpha\} \circ (\uparrow.\{\nu_{k-1}\}) \cdots \circ (\uparrow.\{\nu_0\})] @ m \rightsquigarrow \mathbf{v}_0[\{\alpha\} \circ (\uparrow.\{\nu_{k-1}\}) \cdots \circ (\uparrow.\{\nu_0\})]}$$



$$\begin{array}{c}
\frac{}{\Theta \models^f \text{tt} : \text{bool} @ m \rightsquigarrow \text{tt}} \quad \frac{}{\Theta \models^f \text{ff} : \text{bool} @ m \rightsquigarrow \text{ff}} \quad \frac{\Theta \models^e e : \text{bool} @ m \rightsquigarrow M}{\Theta \models^f \text{up}(e) : \text{bool} @ m \rightsquigarrow M} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.(\text{id}_m \mid \text{bool}) \models^f \tau @ m \rightsquigarrow A \quad \Theta \models^e u : \text{bool} @ m \rightsquigarrow M \quad \Theta \models^f v_1 : A[\text{id.tt}] @ m \rightsquigarrow N_1 \quad \Theta \models^f v_2 : A[\text{id.ff}] @ m \rightsquigarrow N_2}{\Theta \models^e \text{if}(\tau; u; v_1; v_2) : A[\text{id}.M] @ m \rightsquigarrow \text{if}(A; M; N_1; N_2)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta.(\mu \mid A) \models^f u : B @ m \rightsquigarrow M}{\Theta \models^f \lambda(u) : (\mu \mid A) \rightarrow B @ m \rightsquigarrow \lambda(M)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta \models^e e : (\mu \mid A) \rightarrow B @ m \rightsquigarrow M \quad \Theta \models^f u : A @ m \rightsquigarrow N}{\Theta \models^e e(u) : B[\text{id}.N] @ m \rightsquigarrow M(N)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta \models^f u : A @ m \rightsquigarrow M \quad \Theta \models^f v : B[\text{id}.M] @ m \rightsquigarrow N}{\Theta \models^f (u, v) : A \times B @ m \rightsquigarrow (M, N)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Gamma.(\mu \mid A) \vdash B @ m \quad \Theta \models^e u : A \times B @ m \rightsquigarrow M}{\Theta \models^e \text{pr}_1(u) : A @ m \rightsquigarrow \text{pr}_1(M) \quad \Theta \models^e \text{pr}_2(u) : B[\text{id}.\text{pr}_1(M)] @ m \rightsquigarrow \text{pr}_2(M)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Theta.\{\mu\} \models^f u : A @ n \rightsquigarrow M}{\Theta \models^f \text{mod}_\mu(u) : \langle \mu \mid A \rangle @ m \rightsquigarrow \text{mod}_\mu(M)} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A @ n \quad \Theta \models^e e : \langle \mu \mid A \rangle @ m \rightsquigarrow M}{\Theta \models^f \text{up}(e) : \langle \mu \mid A \rangle @ m \rightsquigarrow M} \\
\\
\frac{\nu : o \rightarrow n \quad \mu : n \rightarrow m \quad \Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\}.\{\nu\} \vdash A @ o \quad \Theta.\{\mu\} \models^e u : \langle \nu \mid A \rangle @ n \rightsquigarrow M \quad \Theta.(\mu \mid \langle \nu \mid A \rangle) \models^f \tau @ m \rightsquigarrow B \quad \Theta.(\mu \circ \nu \mid A) \models^f u : B[\uparrow.\text{mod}_\nu(\mathbf{v}_0)] @ m \rightsquigarrow}{\Theta \models^e \text{letmod}(\mu; \nu; \tau; e; u) : B[\text{id}.M] @ m \rightsquigarrow \text{let}_\nu \text{mod}_\mu(\_) \leftarrow M \text{ in } N} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Theta \models^e e : \mathbf{U} @ m \rightsquigarrow A}{\Theta \models^f \text{up}(e) : \mathbf{U} @ m \rightsquigarrow A} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Theta.\{\mu\} \models^f u : \mathbf{U} @ n \rightsquigarrow A \quad \Theta.(\mu \mid \text{El}(A)) \models^f v : \mathbf{U} @ m \rightsquigarrow B}{\Theta \models^f (\mu \mid u) \rhd v : \mathbf{U} @ m \rightsquigarrow (\mu \mid A) \rhd B} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \mathbf{U} @ m \rightsquigarrow A \quad \Theta.(\text{id}_m \mid \text{El}(A)) \models^f v : \mathbf{U} @ m \rightsquigarrow B}{\Theta \models^f u \widehat{\times} v : \mathbf{U} @ m \rightsquigarrow (\mu \mid A) \rhd B} \\
\\
\frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma \quad \Theta.\{\mu\} \models^f u : \mathbf{U} @ m \rightsquigarrow A}{\Theta \models^f \widehat{\langle \mu \mid u \rangle} : \mathbf{U} @ m \rightsquigarrow \widehat{\langle \mu \mid A \rangle}} \quad \frac{\Theta \text{tele} @ m \rightsquigarrow \Gamma}{\Theta \models^f \widehat{\text{bool}} : \mathbf{U} @ m \rightsquigarrow \widehat{\text{bool}}}
\end{array}$$

$$\begin{array}{c}
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^e e : \text{El}(\widehat{\text{bool}}) @ m \rightsquigarrow M}{\Theta \models^e \text{dec}^\triangleright(e) : \text{bool} @ m \rightsquigarrow \text{dec}^\triangleright(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \text{bool} @ m \rightsquigarrow M}{\Theta \models^f \text{dec}^\triangleleft(u) : \text{El}(\widehat{\text{bool}}) @ m \rightsquigarrow \text{dec}^\triangleleft(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^e e : \text{U} @ m \rightsquigarrow A \quad \Theta \models^e f : \text{El}(e) @ m \rightsquigarrow M}{\Theta \models^f \text{up}(f) : \text{El}(e) @ m \rightsquigarrow M} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \text{U} @ m \rightsquigarrow A \quad \Theta.(\text{id}_m \mid \text{El}(A)) \models^f v : \text{U} @ m \rightsquigarrow B \quad \Theta \models^e e : \text{El}(A \hat{\times} B) @ m \rightsquigarrow M}{\Theta \models^e \text{dec}^\triangleright(e) : \text{El}(A) \times \text{El}(B) @ m \rightsquigarrow \text{dec}^\triangleright(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \text{U} @ m \rightsquigarrow A \quad \Theta.(\text{id}_m \mid \text{El}(A)) \models^f v : \text{U} @ m \rightsquigarrow B \quad \Theta \models^f w : \text{El}(A) \times \text{El}(B) @ m \rightsquigarrow M}{\Theta \models^f \text{dec}^\triangleleft(w) : \text{El}(A \hat{\times} B) @ m \rightsquigarrow \text{dec}^\triangleleft(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \text{U} @ n \rightsquigarrow A \quad \Theta.(\mu \mid \text{El}(A)) \models^f v : \text{U} @ m \rightsquigarrow B \quad \Theta \models^e e : \text{El}((\mu \mid A) \hat{\rhd} B) @ m \rightsquigarrow M}{\Theta \models^e \text{dec}^\triangleright(e) : (\mu \mid \text{El}(A)) \rightarrow \text{El}(B) @ m \rightsquigarrow \text{dec}^\triangleright(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Theta \models^f u : \text{U} @ n \rightsquigarrow A \quad \Theta.(\mu \mid \text{El}(A)) \models^f v : \text{U} @ m \rightsquigarrow B \quad \Theta \models^f w : (\mu \mid \text{El}(A)) \rightarrow \text{El}(B) @ m \rightsquigarrow M}{\Theta \models^f \text{dec}^\triangleleft(w) : \text{El}((\mu \mid A) \hat{\rhd} B) @ m \rightsquigarrow \text{dec}^\triangleleft(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A : \text{U} @ n \quad \Theta \models^e e : \text{El}(\widehat{\langle \mu \mid A \rangle}) @ m \rightsquigarrow M}{\Theta \models^e \text{dec}^\triangleright(e) : \langle \mu \mid \text{El}(A) \rangle @ m \rightsquigarrow \text{dec}^\triangleright(M)} \\
\\
\frac{\Theta \text{ tele } @ m \rightsquigarrow \Gamma \quad \Gamma.\{\mu\} \vdash A : \text{U} @ n \quad \Theta \models^f u : \langle \mu \mid \text{El}(A) \rangle @ m \rightsquigarrow M}{\Theta \models^f \text{dec}^\triangleleft(u) : \text{El}(\widehat{\langle \mu \mid A \rangle}) @ m \rightsquigarrow \text{dec}^\triangleleft(M)}
\end{array}$$

## APPENDIX B. A FULL DEFINITION OF AN INTERNAL MTT COSMOS

We present the full definition of an internal **MTT** cosmos. The fact that these constants correspond to the exist in e.g. the syntactic **MTT** cosmos follows from unfolding the internal language this signature is presented in and observing that it is identical to [Definition 3.4](#) in cosmoi rich enough to use **MTT** as an internal language.

## B.1. Term and type sorts.

$$\begin{array}{l}
\text{Ty}_m : \text{U} \\
\text{Tm}_m : \text{Ty}_m \rightarrow \text{U}
\end{array}$$

**B.2. Dependent sums.**

$$\begin{aligned} \text{Sig} &: \left[ \sum_{A: \text{Ty}_m} \text{Tm}_m(A) \rightarrow \text{Ty}_m \right] \rightarrow \text{Ty}_m \\ \alpha_{\text{Sig}} &: (A : \text{Ty}_m)(B : \text{Tm}_m(A) \rightarrow \text{Ty}_m) \rightarrow \left[ \text{Tm}_m(\text{Sig}(A, B)) \cong \sum_{a: \text{Tm}_m(A)} \text{Tm}_m(B(a)) \right] \end{aligned}$$

**B.3. Dependent products.**

$$\begin{aligned} \text{Prod} &: \left[ \sum_{A: \langle \mu | \text{Ty}_n \rangle} \text{let mod}_\mu(A) \leftarrow A \text{ in } [\langle \mu | \text{Tm}_n(A) \rangle \rightarrow \text{Ty}_m] \right] \rightarrow \text{Ty}_m \\ \alpha_{\text{Prod}} &: (\mu | A : \text{Ty}_n) \rightarrow (B : \langle \mu | \text{Tm}_n(A) \rangle \rightarrow \text{Ty}_m) \rightarrow \\ &\quad \left[ \text{Tm}_m(\text{Prod}(\text{mod}_\mu(A), B)) \cong \prod_{a: \langle \mu | \text{Tm}_n(A) \rangle} \text{Tm}_m(B(a)) \right] \end{aligned}$$

**B.4. Booleans.**

$$\begin{aligned} \text{Bool} &: \text{Ty}_m \\ \text{true, false} &: \text{Tm}_m(\text{Bool}) \\ \text{if} &: (A : \text{Tm}_m(\text{Bool}) \rightarrow \text{Ty}_m) \rightarrow \\ &\quad \text{Tm}_m(A(\text{true})) \rightarrow \text{Tm}_m(A(\text{false})) \rightarrow \\ &\quad (b : \text{Tm}_m(\text{Bool})) \rightarrow \text{Tm}_m(A(b)) \\ - &: (A : \text{Tm}_m(\text{Bool}) \rightarrow \text{Ty}_m)(t : \text{Tm}_m(A(\text{true}))) (f : \text{Tm}_m(A(\text{false}))) \rightarrow \\ &\quad (\text{if}(A, t, f, \text{true}) = t) \times (\text{if}(A, t, f, \text{false}) = f) \end{aligned}$$

**B.5. Modal types.**

$$\begin{aligned} \text{Mod}[\mu] &: (\mu | A : \text{Ty}_n) \rightarrow \text{Ty}_m \\ \mathbf{m}[\mu] &: (\mu | A : \text{Ty}_n)(\mu | a : \text{Tm}_n(A)) \rightarrow \text{Tm}_m(\text{Mod}[\mu](A)) \\ \text{letmod}[\mu; \nu] &: (\nu \circ \mu | A : \text{Ty}_n) [B : \langle \mu | \text{Tm}_n(\text{Mod}[\mu](A)) \rangle \rightarrow \text{Ty}_o] \rightarrow \\ &\quad [(\nu \circ \mu | x : \text{Tm}_n(A)) \rightarrow \text{Tm}_o(B(\mathbf{m}[\mu](A, x)))] \rightarrow \\ &\quad (\nu | a : \text{Tm}_m(\text{Mod}[\mu](A))) \rightarrow \\ &\quad \text{Tm}_o(B(a)) \\ - &: (\nu \circ \mu | A : \text{Ty}_n) [B : \langle \mu | \text{Tm}_n(\text{Mod}[\mu](A)) \rangle \rightarrow \text{Ty}_o] \rightarrow \\ &\quad [b : (\nu \circ \mu | x : \text{Tm}_n(A)) \rightarrow B(\mathbf{m}[\mu](A, x))] \rightarrow \\ &\quad (\nu \circ \mu | a : \text{Tm}_n(A)) \rightarrow \text{letmod}[\mu; \nu](A, B, b, \mathbf{m}[\mu](A, a)) = b(a) \end{aligned}$$

**B.6. Universe á la Tarski.**

$$\begin{aligned} \text{Uni} &: \text{Ty}_m \\ \text{El} &: \text{Tm}_m(\text{Uni}) \rightarrow \text{Ty}_m \\ \widehat{\text{Sig}} &: \left[ \sum_{A: \text{Tm}_m(\text{Uni})} \text{Tm}_m(\text{El}(A)) \rightarrow \text{Tm}_m(\text{Uni}) \right] \rightarrow \text{Tm}_m(\text{Uni}) \\ \widehat{\text{Prod}} &: \left[ \sum_{A: \langle \mu | \text{Tm}_n(\text{Uni}) \rangle} \text{let mod}_\mu(A) \leftarrow A \text{ in } [\langle \mu | \text{Tm}_n(\text{El}(A)) \rangle \rightarrow \text{Ty}_m] \right] \rightarrow \text{Tm}_m(\text{Uni}) \\ \widehat{\text{Bool}} &: \text{Tm}_m(\text{Uni}) \\ \widehat{\text{Mod}} &: \langle \mu | \text{Tm}_n(\text{Uni}) \rangle \rightarrow \text{Tm}_m(\text{Uni}) \\ \text{dec}_{\widehat{\text{Sig}}} &: (A : \text{Tm}_m(\text{Uni}))(B : \text{Tm}_m(\text{El}(A)) \rightarrow \text{Tm}_m(\text{Uni})) \rightarrow \end{aligned}$$

$$\begin{aligned}
& \mathsf{Tm}_m(\mathsf{El}(\widehat{\mathsf{Sig}}(A, B))) \cong \mathsf{Tm}_m(\mathsf{Sig}(\mathsf{El}(A), \mathsf{El} \circ B)) \\
\mathsf{dec}_{\widehat{\mathsf{Prod}}} : (\mu \mid A : \mathsf{Tm}_m(\mathsf{Uni}))(B : (\mu \mid \mathsf{Tm}_n(\mathsf{El}(A))) & \rightarrow \mathsf{Tm}_m(\mathsf{Uni})) \rightarrow \\
& \mathsf{Tm}_m(\mathsf{El}(\widehat{\mathsf{Prod}}(A, B))) \cong \mathsf{Tm}_m(\mathsf{Prod}(\mathsf{El}(A), \mathsf{El} \circ B)) \\
\mathsf{dec}_{\widehat{\mathsf{Bool}}} : \mathsf{Tm}_m(\mathsf{El}(\widehat{\mathsf{Bool}})) & \cong \mathsf{Tm}_m(\mathsf{Bool}) \\
\mathsf{dec}_{\widehat{\mathsf{Mod}}} : (\mu \mid A : \mathsf{Tm}_m(\mathsf{Uni})) & \rightarrow \mathsf{Tm}_m(\mathsf{El}(\widehat{\mathsf{Mod}}(A))) \cong \mathsf{Tm}_m(\mathsf{Mod}[\mu](\mathsf{El}(A)))
\end{aligned}$$

**B.7. Internal and external MTT cosmoi coincide.** Let us fix a cosmos  $F : \mathcal{V}$  (see [Definition 3.3](#)) such that  $F$  supports a model of MTT where types are interpreted (up to equivalence) by families of objects and a  $\langle \mu \mid - \rangle$  is interpreted by  $F(\mu)$ . For instance, the syntactic MTT cosmos defined by  $F(m) = \mathbf{PSh}(\mathbf{Cx}_m)$ .

**Theorem B.1.** *The cosmos  $F$  is an MTT cosmos (see [Definition 3.4](#)) precisely when the internal language of  $F$  supports the constants of an internal MTT cosmos.*

*Proof.* First, observe that unfolding the constants  $\mathsf{Ty}_m$  and  $\mathsf{Tm}_m$  into the model in  $F$  gives an object  $\mathcal{T}_m = \llbracket \mathsf{Ty}_m \rrbracket$  and a family over  $\mathcal{T}_m$  given by  $\tau_m = \llbracket \mathsf{Tm}_m \rrbracket$ . We call write  $\dot{\mathcal{T}}_m$  for the domain of  $\tau_m$ . Moreover, this transformation is a bijection: every family  $\tau_m$  gives rise to a pair of  $\mathsf{Ty}_m$  and  $\mathsf{Tm}_m$ .

Showing that the remain constants induce the structure of an MTT cosmos is a standard exercise in unfolding the interpretation of MTT into a presheaf topos. We show only the representative case of modalities.

First, observe that the constant  $\mathsf{Mod}[\mu] : (\mu \mid A : \mathsf{Ty}_n) \rightarrow \mathsf{Ty}_m$  is precisely determined by a morphism  $M : F(\mu)(\mathcal{T}_n) \rightarrow \mathcal{T}_m$ , using the fact that modalities in the model of MTT are interpreted by  $F(\mu)$ . Similarly,  $\mathsf{m}[\mu] : (\mu \mid A : \mathsf{Ty}_n)(\mu \mid a : \mathsf{Tm}_n(A)) \rightarrow \mathsf{Tm}_m(\mathsf{Mod}[\mu](A))$  is precisely equivalent to a map  $m : F(\mu)(\dot{\mathcal{T}}_n) \rightarrow \dot{\mathcal{T}}_m$  such that  $\tau_m \circ m = M \circ F(\mu)(\tau_n)$ .

The final equivalence between the elimination constant  $\mathsf{letmod}[\mu; \nu]$  and the lifting structure from [Definition 3.4](#) is identical similar: unfolding  $\mathsf{letmod}[\mu; \nu]$  and its equation amounts to an internal lifting structure.  $\square$

## APPENDIX C. NEUTRAL AND NORMAL FORMS, INTERNALLY

We require that all normal and neutral forms become equal to their counterparts in the syntactic internal MTT cosmos ([Appendix B](#) and [Section 6](#)) under the assumption  $z : \mathbf{syn}$ . We avoid repeatedly stating this in the specifications of normals and neutrals that follows.

### C.1. Normal types.

$$\begin{aligned}
\mathsf{Prod} : (\mu \mid A : \mathsf{NfTy}_n)(B : (\mu \mid \mathsf{V}_n(A)) & \rightarrow \mathsf{NfTy}_m) \rightarrow \mathsf{NfTy}_m \\
\mathsf{Sum} : (A : \mathsf{NfTy}_m)(B : \mathsf{V}_m(A) & \rightarrow \mathsf{NfTy}_m) \rightarrow \mathsf{NfTy}_m \\
\mathsf{Bool} : \mathsf{NfTy}_m \\
\mathsf{Mod}_\mu : \langle \mu \mid \mathsf{NfTy}_n \rangle & \rightarrow \mathsf{NfTy}_m
\end{aligned}$$

### C.2. Dependent products.

$$\begin{aligned}
\mathsf{lam} : (\mu \mid A : \circ_z \mathsf{Ty}_n(z))(B : (\mu \mid \circ_z \mathsf{Tm}_n(z, A(z))) & \rightarrow \circ_z \mathsf{Ty}_m z) \\
& \rightarrow ((\mu \mid a : \mathsf{V}_n(A)) \rightarrow \mathsf{Nf}_m(B(a))) \\
& \rightarrow \mathsf{Nf}_m(\mathsf{Prod}(A, B))
\end{aligned}$$

$$\begin{aligned}
\mathbf{app} : (\mu \mid A : \circ_z \mathbf{Ty}_n(z)) (B : (\mu \mid \circ_z \mathbf{Tm}_n(z, A(z))) \rightarrow \circ \mathbf{Ty}_m) \\
\rightarrow \mathbf{Ne}_m(\mathbf{Prod}(A, B)) \\
\rightarrow (a : \mathbf{Nf}_m(A)) \\
\rightarrow \mathbf{Ne}_m(B(\eta(a)))
\end{aligned}$$

### C.3. Dependent sums.

$$\begin{aligned}
\mathbf{pair} : (A : \circ_z \mathbf{Ty}_m(z)) (B : \circ_z \mathbf{Tm}_m(z, A(z)) \rightarrow \circ_z \mathbf{Ty}_m(z)) \\
\rightarrow (a : \mathbf{Nf}_m(A)) \rightarrow \mathbf{Nf}_m(B(\eta(a))) \\
\rightarrow \mathbf{Nf}_m(\mathbf{Sig}(A, B)) \\
\mathbf{proj}_0 : (A : \circ_z \mathbf{Ty}_m(z)) (B : \circ_z \mathbf{Tm}_m(z, A(z)) \rightarrow \circ_z \mathbf{Ty}_m(z)) \\
\rightarrow \mathbf{Ne}_m(\mathbf{Sig}(A, B)) \\
\rightarrow \mathbf{Ne}_m(A) \\
\mathbf{proj}_1 : (A : \circ_z \mathbf{Ty}_m(z)) (B : \circ_z \mathbf{Tm}_m(z, A(z)) \rightarrow \circ_z \mathbf{Ty}_m(z)) \\
\rightarrow (p : \mathbf{Ne}_m(\mathbf{Sig}(A, B))) \\
\rightarrow \mathbf{Ne}_m(B(\eta(\mathbf{proj}_0(p))))
\end{aligned}$$

### C.4. Booleans.

$$\begin{aligned}
\mathbf{up} : \mathbf{Ne}_m(\mathbf{Bool}) \rightarrow \mathbf{Nf}_m(\mathbf{Bool}) \\
\mathbf{tt}, \mathbf{ff} : \mathbf{Nf}_m(\mathbf{Bool}) \\
\mathbf{if} : (A : \circ_z \mathbf{Tm}_m(z, \mathbf{Bool}(z)) \rightarrow \circ_z \mathbf{Ty}_m(z)) \\
\rightarrow \mathbf{Nf}_m(A(\mathbf{true})) \rightarrow \mathbf{Nf}_m(A(\mathbf{false})) \rightarrow (b : \mathbf{Ne}_m(\mathbf{Bool})) \rightarrow \mathbf{Ne}_m(A(\eta(b)))
\end{aligned}$$

### C.5. Modal types.

$$\begin{aligned}
\mathbf{up} : (\mu \mid A : \circ_z \mathbf{Ty}_n(z)) \rightarrow \mathbf{Ne}_m(\mathbf{Mod}[\mu]A) \rightarrow \mathbf{Nf}_m(\mathbf{Mod}[\mu]A) \\
\mathbf{mod}_\mu : (\mu \mid A : \circ_z \mathbf{Ty}_n(z)) \rightarrow \langle \mu \mid \mathbf{Nf}_n(A) \rangle \rightarrow \mathbf{Nf}_{\mathbf{Mod}[\mu]A} \\
\mathbf{letmod}_{\mu;\nu} : (\nu \circ \mu \mid A : \circ_z \mathbf{Ty}_n(z)) (B : (\nu \mid a : \circ_z \mathbf{Tm}_m(z, \mathbf{Mod}[\mu](z, A(z)))) \rightarrow \circ_z \mathbf{Ty}_o(z)) \\
\rightarrow ((\mu \circ \nu \mid a : \mathbf{V}_n(A)) \rightarrow \mathbf{Nf}_o(B(\mathbf{m}[\mu](A, \eta(a)))) \\
\rightarrow (\nu \mid a : \mathbf{Ne}_m(\mathbf{Mod}[\mu](A))) \rightarrow \mathbf{Ne}_o(B(\eta(a)))
\end{aligned}$$

### C.6. Universe á la Tarski.

$$\begin{aligned}
\mathbf{Uni} : \mathbf{NfTy}_m \quad \mathbf{El} : \mathbf{Nf}_m(\mathbf{Uni}) \rightarrow \mathbf{NfTy}_m \\
\mathbf{up} : \mathbf{Ne}_m(\mathbf{Uni}) \rightarrow \mathbf{Nf}_m(\mathbf{Uni}) \\
\widehat{\mathbf{Prod}} : (\mu \mid A : \mathbf{Nf}_n(\mathbf{Uni})) (B : (\mu \mid \mathbf{V}_n(\mathbf{El}(A))) \rightarrow \mathbf{Nf}_m(\mathbf{Uni})) \rightarrow \mathbf{Nf}_m(\mathbf{Uni}) \\
\widehat{\mathbf{Sum}} : (A : \mathbf{Nf}_m(\mathbf{Uni})) (B : \mathbf{V}_m(\mathbf{El}(A)) \rightarrow \mathbf{Nf}_m(\mathbf{Uni})) \rightarrow \mathbf{Nf}_m(\mathbf{Uni}) \\
\widehat{\mathbf{Bool}} : \mathbf{Nf}_m(\mathbf{Uni}) \\
\widehat{\mathbf{Mod}}_\mu : \langle \mu \mid \mathbf{Nf}_n(\mathbf{Uni}) \rangle \rightarrow \mathbf{Nf}_m(\mathbf{Uni}) \\
\mathbf{dec}_{\widehat{\mathbf{Prod}}}^\triangleright : (\mu \mid A : \mathbf{Nf}_n(\mathbf{Uni})) (B : (\mu \mid \mathbf{V}_n(\mathbf{El}(A))) \rightarrow \mathbf{Nf}_m(\mathbf{Uni})) \\
\rightarrow \mathbf{Nf}_m(\mathbf{Prod}(\mathbf{El}(A), \mathbf{El}(B))) \rightarrow \mathbf{Nf}_m(\mathbf{El}(\widehat{\mathbf{Prod}}(A, B))) \\
\mathbf{dec}_{\widehat{\mathbf{Sum}}}^\triangleright : (A : \mathbf{Nf}_m(\mathbf{Uni})) (B : \mathbf{V}_m(\mathbf{El}(A)) \rightarrow \mathbf{Nf}_m(\mathbf{Uni}))
\end{aligned}$$

$$\begin{aligned}
& \rightarrow \text{Nf}_m(\mathbf{Sum}(\mathbf{El}(A), \mathbf{El}(B))) \rightarrow \text{Nf}_m(\mathbf{El}(\widehat{\mathbf{Sum}}(A, B))) \\
\text{dec}_{\mathbf{Prod}}^{\triangleleft} & : (\mu \mid A : \text{Nf}_n(\text{Uni}))(B : (\mu \mid \text{V}_n(\mathbf{El}(A)))) \rightarrow \text{Nf}_m(\text{Uni}) \\
& \rightarrow \text{Ne}_m(\mathbf{El}(\widehat{\mathbf{Prod}}(A, B))) \rightarrow \text{Ne}_m(\mathbf{Prod}(\mathbf{El}(A), \mathbf{El}(B))) \\
\text{dec}_{\mathbf{Sum}}^{\triangleleft} & : (A : \text{Nf}_m(\text{Uni}))(B : \text{V}_m(\mathbf{El}(A))) \rightarrow \text{Nf}_m(\text{Uni}) \\
& \rightarrow \text{Ne}_m(\mathbf{El}(\widehat{\mathbf{Sum}}(A, B))) \rightarrow \text{Ne}_m(\mathbf{Sum}(\mathbf{El}(A), \mathbf{El}(B))) \\
\text{dec}_{\mathbf{Bool}}^{\triangleright} & : \text{Nf}_m(\text{Bool}) \rightarrow \text{Nf}_m(\mathbf{El}(\widehat{\mathbf{Bool}})) \\
\text{dec}_{\mathbf{Mod}_\mu}^{\triangleright} & : (\mu \mid A : \text{Nf}_n(\text{Uni})) \rightarrow \text{Nf}_m(\text{Mod}[\mu](A)) \rightarrow \text{Nf}_m(\mathbf{El}(\widehat{\mathbf{Mod}}(A))) \\
\text{dec}_{\mathbf{Bool}}^{\triangleleft} & : \text{Ne}_m(\mathbf{El}(\widehat{\mathbf{Bool}})) \rightarrow \text{Ne}_m(\text{Bool}) \\
\text{dec}_{\mathbf{Mod}_\mu}^{\triangleleft} & : (\mu \mid A : \text{Nf}_n(\text{Uni})) \rightarrow \text{Ne}_m(\mathbf{El}(\widehat{\mathbf{Mod}}(A))) \rightarrow \text{Ne}_m(\text{Mod}[\mu](A))
\end{aligned}$$

**C.7. Interpreting neutral and normal forms in the glued model.** Substantiating these constants in the glued model (see [Theorem 6.1](#)) relies on unfolding the constants described above into this model and showing that the expected normal or neutral form from [Section 5](#) can be used. These proofs follow Hofmann [\[Hof99\]](#), and again the main subtlety is showing that binders in constants such as **Prod** are correctly interpreted.

For each normal and neutral form, the requirement that constant lie over the appropriate constant from the syntactic cosmos forces the “syntactic portion” of each constant. More precisely, consider again **Prod** and set  $I = (\mu \mid A : \text{NfTy}_n) \times \text{V}_n(A)$  and  $I' = (\mu \mid A : \text{Ty}_n) \times \text{Tm}_n(A)$ . The definition of **Prod** is a commuting square of the following shape in  $\mathbf{PSh}(\text{Ren}_m)$ :

$$\begin{array}{ccc}
\llbracket I \rrbracket^1 & \xrightarrow{\llbracket \mathbf{Prod} \rrbracket^1} & \llbracket \text{NfTy}_m \rrbracket^1 \\
\llbracket I \rrbracket \downarrow & & \downarrow \llbracket \text{NfTy}_m \rrbracket \\
\mathbf{i}[m]^*(\llbracket I' \rrbracket) & \xrightarrow{\llbracket \mathbf{Prod} \rrbracket^0} & \mathbf{i}[m]^*(\llbracket \text{Ty}_m \rrbracket)
\end{array}$$

The bottom map of this square  $\llbracket \mathbf{Prod} \rrbracket^0$ , moreover, must be  $\mathbf{i}[m]^*(\text{Prod})$  in order to ensure that  $\mathbf{Prod} = \text{Prod}(z)$  with  $z : U$ . It remains to define the top arrow of this diagram. This is complicated by the difference between the interpretation of dependent products in  $\mathcal{G}\llbracket m \rrbracket$  and  $\mathbf{PSh}(\text{Ren}_m)$ . In particular, while the “downstairs” portion of a dependent product in  $\mathcal{G}\llbracket m \rrbracket$  is a dependent product in  $\mathbf{PSh}(\text{Cx}_m)$ , the upstairs portion is not a dependent product in  $\mathbf{PSh}(\text{Ren}_m)$ .

In order to define  $\llbracket \mathbf{Prod} \rrbracket^1$ , it suffices to fix a morphism  $x : \mathbf{y}(\Theta) \rightarrow \llbracket \text{NfTy}_m \rrbracket^1$  and define the action of  $\llbracket \mathbf{Prod} \rrbracket^1$  on it. First, observe that we can extend  $x$  to the

following commuting square:

$$\begin{array}{ccc}
 \mathbf{y}(\Theta) & \xrightarrow{x} & \llbracket I \rrbracket^0 \\
 \downarrow & & \downarrow \llbracket I \rrbracket \\
 \mathbf{i}[m]^*(\mathbf{y}(\mathbf{i}[m](\Theta))) & \xrightarrow{\mathbf{i}[m]^*(\tilde{x})} & \mathbf{i}[m]^*(\llbracket I' \rrbracket)
 \end{array}$$

By universal property, therefore, we can decompose  $x$  into a pair:

$$\begin{array}{ccc}
 \mathbf{y}(\Theta.\{\mu\}) & \xrightarrow{x_0} & \llbracket \mathbf{NfTy}_n \rrbracket^1 \\
 \downarrow & & \downarrow \llbracket \mathbf{NfTy}_n \rrbracket \\
 \mathbf{i}[n]^*(\mathbf{y}((\mathbf{i}[m](\Theta)).\{\mu\}))) & \xrightarrow{\mathbf{i}[n]^*(\tilde{x}_0)} & \mathbf{i}[n]^*(\llbracket \mathbf{Ty}_n \rrbracket) \\
 \\ 
 \mathbf{y}(\Theta.x_0) & \xrightarrow{x_1} & \llbracket \mathbf{NfTy}_m \rrbracket^1 \\
 \downarrow & & \downarrow \llbracket \mathbf{NfTy}_m \rrbracket \\
 \mathbf{i}[m]^*(\mathbf{y}(\mathbf{i}[m](\Theta.(\mu \mid \tilde{x}_0)))) & \xrightarrow{\mathbf{i}[m]^*(\tilde{x}_1)} & \mathbf{i}[m]^*(\llbracket \mathbf{Ty}_m \rrbracket)
 \end{array}$$

We may now define  $\mathbf{Prod}(x) = (\mu \mid x_0) \rightarrow x_1$ . The naturality of this assignment, as well as the fact that it commutes appropriately, are direct computations.

#### REFERENCES

- [Abe13] Andreas Abel. “Normalization by Evaluation: Dependent Types and Impredicativity”. Habilitation. Ludwig-Maximilians-Universität München, 2013 (cit. on p. 6).
- [Alt+01] T. Altenkirch, P. Dybjer, M. Hofmann, and P. Scott. “Normalization by Evaluation for Typed Lambda Calculus with Coproducts”. In: *Proceedings of the 16th Annual IEEE Symposium on Logic in Computer Science*. LICS '01. Washington, DC, USA: IEEE Computer Society, 2001, pp. 303–. URL: <http://dl.acm.org/citation.cfm?id=871816.871869> (cit. on p. 6).
- [AHS95] Thorsten Altenkirch, Martin Hofmann, and Thomas Streicher. “Categorical reconstruction of a reduction free normalization proof”. In: *Category Theory and Computer Science*. Ed. by David Pitt, David E. Rydeheard, and Peter Johnstone. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995, pp. 182–199. ISBN: 978-3-540-44661-3 (cit. on pp. 3, 6, 7).



- [AK16] Thorsten Altenkirch and Ambrus Kaposi. “Normalisation by Evaluation for Dependent Types”. In: *1st International Conference on Formal Structures for Computation and Deduction (FSCD 2016)*. Ed. by Delia Kesner and Brigitte Pientka. Vol. 52. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016, 6:1–6:16. ISBN: 978-3-95977-010-1. DOI: [10.4230/LIPIcs.FSCD.2016.6](https://doi.org/10.4230/LIPIcs.FSCD.2016.6). URL: <http://drops.dagstuhl.de/opus/volltexte/2016/5972> (cit. on p. 3).
- [AGV72] Michael Artin, Alexander Grothendieck, and Jean-Louis Verdier. *Théorie des topos et cohomologie étale des schémas*. Séminaire de Géométrie Algébrique du Bois-Marie 1963–1964 (SGA 4), Dirigé par M. Artin, A. Grothendieck, et J.-L. Verdier. Avec la collaboration de N. Bourbaki, P. Deligne et B. Saint-Donat, Lecture Notes in Mathematics, Vol. 269, 270, 305. Berlin: Springer-Verlag, 1972 (cit. on pp. 7, 14, 18).
- [Awo18] Steve Awodey. “Natural models of homotopy type theory”. In: *Mathematical Structures in Computer Science* 28.2 (2018), pp. 241–286. ISSN: 09601295. DOI: [10.1017/S0960129516000268](https://doi.org/10.1017/S0960129516000268). eprint: [1406.3219](https://arxiv.org/abs/1406.3219) (cit. on pp. 7, 19).
- [Bir+20] Lars Birkedal, Ranald Clouston, Bassel Manna, Rasmus Ejlers Møgelberg, Andrew M. Pitts, and Bas Spitters. “Modal dependent type theory and dependent right adjoints”. In: *Mathematical Structures in Computer Science* 30.2 (2020), pp. 118–138. DOI: [10.1017/S0960129519000197](https://doi.org/10.1017/S0960129519000197). eprint: [1804.05236](https://arxiv.org/abs/1804.05236) (cit. on pp. 2, 13).
- [CJ95] Aurelio Carboni and Peter Johnstone. “Connected limits, familial representability and Artin glueing”. In: *Mathematical Structures in Computer Science* 5.4 (1995), pp. 441–459. DOI: [10.1017/S0960129500001183](https://doi.org/10.1017/S0960129500001183) (cit. on pp. 7, 14, 18).
- [Coq19] Thierry Coquand. “Canonicity and normalization for dependent type theory”. In: *Theoretical Computer Science* 777 (2019), pp. 184–191. DOI: [10.1016/j.tcs.2019.01.015](https://doi.org/10.1016/j.tcs.2019.01.015) (cit. on pp. 3, 7).
- [Fio02] Marcelo Fiore. “Semantic Analysis of Normalisation by Evaluation for Typed Lambda Calculus”. In: *Proceedings of the 4th ACM SIGPLAN International Conference on Principles and Practice of Declarative Programming*. PPDP ’02. Pittsburgh, PA, USA: ACM, 2002, pp. 26–37. ISBN: 1-58113-528-9. DOI: [10.1145/571157.571161](https://doi.org/10.1145/571157.571161) (cit. on pp. 3, 6, 7).
- [Gra+20a] Daniel Gratzer, G.A. Kavvos, Andreas Nuyts, and Lars Birkedal. *Multimodal Dependent Type Theory*. Extended version of the LICS paper by the same name. 2020 (cit. on pp. 1–3, 7, 8, 17–19).
- [Gra+20b] Daniel Gratzer, G.A. Kavvos, Andreas Nuyts, and Lars Birkedal. “Multimodal Dependent Type Theory”. In: *Proceedings of the 35th Annual ACM/IEEE Symposium on Logic in Computer Science*. LICS ’20. ACM, 2020. DOI: [10.1145/3373718.3394736](https://doi.org/10.1145/3373718.3394736) (cit. on p. 2).
- [GS20] Daniel Gratzer and Jonathan Sterling. *Syntactic categories for dependent type theory: sketching and adequacy*. 2020. arXiv: [2012.10783](https://arxiv.org/abs/2012.10783) [cs.LG] (cit. on pp. 4, 7).

- [GSB19a] Daniel Gratzer, Jonathan Sterling, and Lars Birkedal. “Implementing a Modal Dependent Type Theory”. In: *Proc. ACM Program. Lang.* 3 (ICFP 2019). DOI: [10.1145/3341711](https://doi.org/10.1145/3341711) (cit. on p. 3).
- [GSB19b] Daniel Gratzer, Jonathan Sterling, and Lars Birkedal. *Normalization-by-Evaluation for Modal Dependent Type Theory*. Technical Report for the ICFP paper by the same name. 2019. URL: <https://jozefg.github.io/papers/2019-implementing-modal-dependent-type-theory-tech-report.pdf> (cit. on p. 3).
- [Hof99] Martin Hofmann. “Semantical Analysis of Higher-Order Abstract Syntax”. In: *Proceedings of the 14th Annual IEEE Symposium on Logic in Computer Science*. LICS ’99. Washington, DC, USA: IEEE Computer Society, 1999, pp. 204–. ISBN: 0-7695-0158-3. URL: <http://dl.acm.org/citation.cfm?id=788021.788940> (cit. on pp. 19, 35).
- [HS97] Martin Hofmann and Thomas Streicher. “Lifting Grothendieck Universes”. Unpublished note. 1997. URL: <https://www2.mathematik.tu-darmstadt.de/~streicher/NOTES/lift.pdf> (cit. on p. 13).
- [KHS19] Ambrus Kaposi, Simon Huber, and Christian Sattler. “Gluing for type theory”. In: *Proceedings of the 4th International Conference on Formal Structures for Computation and Deduction (FSCD 2019)*. Ed. by Herman Geuvers. Vol. 131. 2019 (cit. on p. 7).
- [KKA19] Ambrus Kaposi, András Kovács, and Thorsten Altenkirch. “Constructing Quotient Inductive-inductive Types”. In: *Proc. ACM Program. Lang.* 3.POPL (Jan. 2019), 2:1–2:24. ISSN: 2475-1421. DOI: [10.1145/3290315](https://doi.org/10.1145/3290315) (cit. on p. 9).
- [KPT99] Yoshiaki Kinoshita, John Power, and Makoto Takeyama. “Sketches”. In: *Journal of Pure and Applied Algebra* 143.1 (1999), pp. 275–291. ISSN: 0022-4049. DOI: [10.1016/S0022-4049\(98\)00114-5](https://doi.org/10.1016/S0022-4049(98)00114-5) (cit. on p. 10).
- [Lac09] Stephen Lack. “A 2-Categories Companion”. In: *The IMA Volumes in Mathematics and its Applications* (Sept. 2009), pp. 105–191. ISSN: 0940-6573. DOI: [10.1007/978-1-4419-1524-5\\_4](https://doi.org/10.1007/978-1-4419-1524-5_4) (cit. on p. 10).
- [LS16] Daniel R. Licata and Michael Shulman. “Adjoint Logic with a 2-Category of Modes”. In: *Logical Foundations of Computer Science*. Ed. by Sergei Artemov and Anil Nerode. Springer International Publishing, 2016, pp. 219–235. DOI: [10.1007/978-3-319-27683-0\\_16](https://doi.org/10.1007/978-3-319-27683-0_16) (cit. on p. 2).
- [MS93] John C. Mitchell and Andre Scedrov. “Notes on sconing and relators”. In: *Computer Science Logic*. Ed. by E. Börger, G. Jäger, H. Kleine Büning, S. Martini, and M. M. Richter. Springer Berlin Heidelberg, 1993, pp. 352–378. DOI: [10.1007/3-540-56992-8\\_21](https://doi.org/10.1007/3-540-56992-8_21) (cit. on p. 5).
- [OP18] Ian Orton and Andrew M. Pitts. “Axioms for Modelling Cubical Type Theory in a Topos”. In: *Logical Methods in Computer Science* 14.4 (2018). DOI: [10.23638/LMCS-14\(4:23\)2018](https://doi.org/10.23638/LMCS-14(4:23)2018). arXiv: [1712.04864](https://arxiv.org/abs/1712.04864) (cit. on p. 18).
- [Red20] The RedPRL Development Team. *cooltt*. 2020. URL: <http://www.github.com/RedPRL/cooltt> (cit. on p. 8).
- [RSS20] Egbert Rijke, Michael Shulman, and Bas Spitters. “Modalities in homotopy type theory”. In: *Logical Methods in Computer Science* 16.1 (2020). eprint: [1706.07526](https://arxiv.org/abs/1706.07526) (cit. on p. 13).

- [Shu15] Michael Shulman. “Univalence for inverse diagrams and homotopy canonicity”. In: *Mathematical Structures in Computer Science* 25.5 (2015), pp. 1203–1277. DOI: [10 . 1017 / S0960129514000565](https://doi.org/10.1017/S0960129514000565). eprint: [1203.3253](https://arxiv.org/abs/1203.3253) (cit. on p. 7).
- [Ste21] Jonathan Sterling. “First Steps in Synthetic Tait Computability”. Thesis draft. 2021 (cit. on pp. 1, 3, 5, 7, 8).
- [SA21] Jonathan Sterling and Carlo Angiuli. “Normalization for Cubical Type Theory”. In: *Proceedings of the 36th Annual ACM/IEEE Symposium on Logic in Computer Science*. LICS ’21. New York, NY, USA: ACM, 2021 (cit. on p. 5).
- [SG20] Jonathan Sterling and Daniel Gratzer. *Lecture Notes on Synthetic Tait Computability*. Subsumed by Sterling’s forthcoming thesis. 2020 (cit. on p. 5).
- [SH20] Jonathan Sterling and Robert Harper. “Logical Relations As Types: Proof-Relevant Parametricity for Program Modules”. Draft paper. 2020. arXiv: [2010.08599](https://arxiv.org/abs/2010.08599). URL: <https://arxiv.org/abs/2010.08599> (cit. on pp. 1, 5, 18).
- [Str98] Thomas Streicher. “Categorical intuitions underlying semantic normalisation proofs”. In: *Preliminary Proceedings of the APPSEM Workshop on Normalisation by Evaluation*. Ed. by O. Danvy and P. Dybjer. Department of Computer Science, Aarhus University, 1998 (cit. on p. 6).
- [Uem19] Taichi Uemura. “A General Framework for the Semantics of Type Theory”. In: (Apr. 2019). eprint: [1904.04097](https://arxiv.org/abs/1904.04097) (math.CT). URL: <https://arxiv.org/abs/1904.04097> (cit. on p. 4).