Markov Fibrations

Eigil Fjeldgren Rischel

Department of Computer and Information Sciences
University of Strathclyde

2025

Signed, Eigil Fjeldgren Rischel, February 17, 2026.

Abstract

In the theory of open games, there has been a longstanding search for a good theory of "dependent optics", that is a common generalization of Riley's theory of optics and the dependent lenses of Spivak. We develop such a theory in the case of Markov categories, by introducing a novel object which we call *Markov fibrations*. Given an indexed family of objects $(X_i)_{i \in I}$ and a function $f : J \to I$, we can construct a new family $(X_{f(j)})_{j \in J}$. This structure is captured by the classical notion of Grothendieck fibration. If f is instead a probability kernel, this reindexing no longer makes sense, yet there is still a relation between stochastic maps between indexing sets and indexed stochastic maps between the indexed families. Markov fibrations generalize Grothendieck fibrations in a way which captures this relationship, and they admit a notion of fiberwise opposite which gives a useful notion of stochastic optic. In addition to introducing this concept and proving its basic properties, we also give applications to the theory of open games, and to Myers' categorical theory of systems. Along the way we also give a detailed treatment of the so-called Para-construction, including how to generalize it to arbitrary 2-categories.

Contents

Chapter 1

Introduction

In applying category theory to fields as diverse as game theory [Gha+18a; BHZ19; LHS21], machine learning [Cru+21; Gav24], dynamical systems [Mye23; Lyn+25], and server design [VC22], people have found it useful to study categories whose morphisms describe processes or functions that take place in two "stages", where the second is dependent on the first, but composes in the other direction—so-called *lenses*. In many of these domains, stochastic phenomena play a role. There is a useful generalization of lenses to categories of stochastic maps—the so-called *optics* of Riley, [Ril18]—but this does not accommodate another useful generalization, so called *dependent lenses* where not only the backwards process but the set it takes values in is indexed over the base. It has been a long-standing problem to develop a suitable common generalization, "dependent optics", of these two ideas. In this thesis, we solve this problem by developing a theory of *Markov fibrations* (chapter 3).

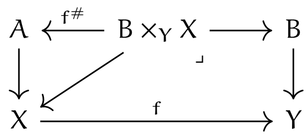
A Markov fibration is a weakening of the notion of (Grothendieck) fibration to include (subject to some assumptions) Markov categories of indexed families of objects (given by deterministic functions $E \to X$) and compatible stochastic maps (given by commutative squares). The main point of Markov fibrations is that they admit *fiberwise opposites*, which generalize the fiberwise opposites of ordinary categories. Just as the fiberwise opposite of the codomain fibration of a finitely complete category describes dependent lenses (see section 2.3), the fiberwise opposite of these codomain Markov fibrations give a good notion of stochastic lens (see Theorem 3.7.2).

We will apply our theory chiefly to two problems. First, we will use them to generalize *open games* (Hedges, [Hed16]) to a larger class of interfaces (namely, indexed families of sets) while at the same time considering possibly-stochastic maps. This will allow us to construct the so-called *external choice* operator on open games, which describes branching.

Secondly, we will generalize Myers' categorical dynamical systems theory to allow for stochastic maps in the base. When combined with another generalization of these systems, to general parametrized lenses, this gives a more natural way of modeling certain stochastic dynamical systems, such as those associated to the training dynamics of machine learning models, see eg Example 6.5.8.

1.1 A brief introduction to lenses

If $\mathcal{C}$ is a category with finite products, the category $\mathsf{Lens}(\mathcal{C})$ of *lenses* in $\mathcal{C}$ has objects pairs $\binom{A}{X}$ of objects in $\mathcal{C}$, and morphisms $\binom{A}{X} \leftrightarrows \binom{B}{Y}$ given by pairs $f : X \to Y, f^{\#} : B \times X \to A$. (Note that it is of course the morphisms, not the objects, that are called lenses). One way to generalize this is to ask for $\mathcal{C}$ to have pullbacks, and consider an object given by a more general map $A \to X$, and let a map be $f : X \to Y, f^{\#} : B \times_Y X \to A$ (so that the triangle over X commutes). (Note that this recovers the "simple lenses" when the objects are of the form $A \times X \to X$).x



Under the interpretation of a map $A \to X$ as a family of objects A_x indexed over the elements of X, we see this as a "dependent lens" (since A is a "dependent type"). There are various variations of this idea, which ultimately all fit the pattern of taking a Grothendieck fibration $\mathcal{D} \to \mathcal{C}$ and forming the fiberwise opposite—for the dependent lenses as above, this is the codomain fibration $\mathcal{C}^{\to} \to \mathcal{C}$.

Meanwhile, another wide-ranging generalization of $\mathsf{Lens}(\mathcal{C})$ are the *optics* of Riley [Ril18]. For $\mathcal{C}$ a monoidal category, the objects of $\mathsf{Optic}(\mathcal{C})$ are again pairs $\binom{A}{X}$ of objects, but now the morphisms are elements of the coend

$$\int^{M \in \mathcal{C}} \mathcal{C}(X, M \otimes Y) \times \mathcal{C}(M \otimes B, A).$$

That is, to give an optic $\binom{A}{X} \leftrightarrows \binom{B}{Y}$ is to give an object $M \in \mathcal{C}$ and morphisms $l : X \to M \otimes Y, r : M \otimes B \to A$, up to the equivalence relation generated by, whenever $s : N \to M, l : X \to N \otimes Y, r : M \otimes B \to A$, identifying the two optics given by $(M, (s \otimes 1_Y) \circ l, r)$ and $(N, l, r \circ (s \otimes 1_B))$. One can show that, in the case where $\mathcal{C}$ is a Cartesian monoidal category, this set can be identified with the set of (simple) lenses.

A vector field on a smooth manifold X is simply a (smooth) section of the tangent bundle TX. A section is a bundle map from the trivial bundle $X \to X$. This gives the idea of considering a *parametrized* dynamical system as consisting of a map $X \to A$ to some other smooth manifold, a bundle $E \to A$ (the points of E are the parameters) and a bundle map $E \times_A X \to TX$. As we will see, such a bundle map is exactly a dependent lens in a category of bundles. Building on this idea, Myers [Mye23] described a highly abstract theory of open dynamical systems, given by lenses out of tangent bundles (where the notion of space, bundle and tangent bundle are generalized to any fibration).

It has been observed that, although a lens from the tangent bundle $TS \leftrightarrows A$ may indeed be said to describe an open dynamical system with state space S, the same can be said for a lens of type $TS \otimes A \leftrightarrows I$—although a dynamical system of a different kind (essentially, the first type are the *Moore machines*, the second the *Mealy machines*—see the introduction to chapter 4 for more on this). It is natural to consider parametrized maps $TS \otimes A \leftrightarrows B$ as a common generalization of these two concepts—and in fact, special cases of this idea, lenses parametrized by tangent bundles, have already been considered many times, for example in the semantics of gradient descent (see [Cru+21; Cap+22; FST19]).

Since parametrized maps compose in an obvious way, we now have three different notions of morphisms between bundles—lenses, charts, and "generalized systems".

These should form some sort of symmetric monoidal triple category, but the right axiomatisation of this concept is somewhat elusive.

Lenses, in their various incarnations, have been widely used in applied category theory. For example, Hedges and his collaborators have developed a compositional approach to game theory [Hed15; Gha+18a; Fre+23; LHS21]. Recent work by Hedges and Sakamoto bring this idea to reinforcement learning [HS24]. Other authors have extended this theory to cover iterated games [Gha+18b], extensive form games [Cap+23], and approximate solution concepts [Gha25]. We briefly mentioned Myers' categorical systems theory above, and we will see much more of it later. As we discussed above, there is also a literature using lenses to study gradient descent in an abstract sense. This is without even mentioning their original role in the theory of functional programming (this is the origin of the somewhat confusing term *lens*), or their prehistory in Gödel's dialectica interpretation—see eg [Hed18] for a survey of this.

1.2 Dependent Optics

An open game, in the sense of Hedges, has as its interface two objects in the category of lenses (of sets) $\binom{S}{X}, \binom{R}{Y}$ (that is, two pairs of sets). An open game with this interface consists of a set Σ of *strategies* equipped with a function $\Sigma \to \mathsf{Lens}(\mathsf{Set})(\binom{S}{X}, \binom{R}{Y})$, and a subset $\mathrm{Eq}(x,k) \subset \Sigma$ of *equilibrium strategies* for each $x \in X$ and $k : Y \to R$—note that such an x is exactly a lens $\binom{*}{*} \leftrightarrows \binom{S}{X}$, and such a k is exactly a lens $\binom{R}{Y} \to \binom{*}{*}$. This view of open games makes the composition rule much simpler to define, although we will not delve into the details here, see eg [Gha+18a]. The idea is that the open game represents the strategies and preferences of a player or set of players—$\sigma \in \Sigma$ are the possible strategies, $x \in X$ is the information revealed to the player before they make their decision of which moves to make, the resulting $y \in Y$ is the choice the player "sends" in response to x, and the $r \in R$ is the "utility", the value they eventually learn, depending on their choice y, which they have some preferences over.

The above description gives a theory of deterministic games, but of course, it is completely essential for game theory to model both random decisions (mixed strategies) and decisions made under uncertainty. This motivates the replacement of lenses in this definition with optics in a category of probability kernels (the form of the equilibrium relation must be modified as well, see [BHZ19; Cap+22])

In the abstract study of open games, it is desirable to define a "choice" operation $\oplus$ on the objects (the pairs of sets) so that maps into $\binom{R}{X} \oplus \binom{R'}{X'}$ represent players who are faced with some binary choice between X and X', after which play may proceed for a time in one of two branches. However, this is immediately problematic because the category of lenses does not have coproducts—inside the category of *dependent* lenses, however, we can form the coproduct simply as $R \times X + R' \times X' \to X + X'$—that is, we get a family which is R over X, and R' over X'. This indeed does the job, but for game theory the extension to stochastic maps—optics—is absolutely essential. This motivates the search for a theory of *dependent optics*, a common generalization of optics and dependent lenses.

There have been a number of attempts at this—see eg [Bra+21; BH22; Mil21; Ver23]. While these efforts have to some extent succeeded in describing a theory that is general enough to contain the desired examples, it is generally very ad hoc. The construction of optics as a special case of Vertechi's dependent optics [Ver23] gives them as fibre optics with the indexing bicategory being (the delooping of) a monoidal category $\mathcal{M}$, but embeds lenses using a bicategory of spans. Thus simple lenses admit two distinct encodings in this theory. Moreover, Vertechi's example of dependent monoidal optics fails to describe a supercategory of $\mathsf{Optic}(\mathcal{C})$ when applied to Markov categories, and are thus not suitable for, for example, game theory.

1.3 Categorical Cybernetics

The observation that both open games and machine learning systems can be fruitfully described using lenses, naturally led to the idea that this could be used to develop a more thorough analogy between the two. After all, a machine learning system is also a player in a kind of game (with the objective of minimizing loss). Hedges coined the phrase *categorical cybernetics* to describe this locus of ideas ([Cap+22], also used in [HS24]).

Myers' categorical systems theory also describes dynamical systems as lenses. However, where a cybernetic system in the sense of [Cap+22] is a parametrized lens $\Sigma \otimes \binom{\bar{X}}{X} \leftrightarrows \binom{\bar{Y}}{Y}$, and thus has two different inputs—$x \in X$, which is the information on which basis they are allowed to choose their decision $y \in Y$, and $y^\sharp \in \bar{Y}$, which is their "utility", the information they are allowed to care about, a dynamical system in Myers' sense only has one input and one output.

It is straightforward to describe the machine learning part of the categorical cybernetics analogy in terms of Myers' theory—a "learner" is simply a lens of the above form with Σ replaced by a tangent bundle $\mathsf{T}S$. This idea (described in these terms, although not with the analogy to machine learning) has already been described by [CLS24]. However, for game theory, it is necessary that the output of a given system can be stochastically chosen (a player must be able to randomize their strategy). At the same time, for a machine learning system, the $x \in X$ is usually some sort of sample from a training distribution—ie, it is random. Hence to really describe the training dynamics of machine learning systems using this theory, we are again naturally drawn to consider systems theories (that is, fibrations) which have stochastic maps in the base, not merely the fiber.

1.4 Overview and contributions

We begin the thesis in chapter 2 with a review of some preliminary material. This chapter can largely be skipped for readers already familiar with the material (Markov categories, double categories, lenses and optics, fibrations, and categorical dynamical systems theory). The chapter on Markov categories introduces a few novel auxiliary notions, but these can be referred back to as necessary (they are primarily regularity conditions which hold in most Markov categories of interest).

In chapter 3, we give the main technical contribution of the thesis by developing a theory of *Markov fibrations*. Their fiberwise opposites generalize both the fiberwise opposites of codomain fibrations (dependent lenses in a classical sense) and optics in Markov categories. Theorem 3.7.2 provides the statement of the latter (the former is straightforward). We also give a description of monoidal structures on Markov fibrations (which pass to monoidal structures on the resulting categories of optics).

In chapter 4, we give a construction of the double category of *parametrized morphisms* for a category action (or actegory)—in fact, we do this internally to any 2-category. This has the advantage of allowing us to construct more highly-structured versions of this double category by carrying out the construction internally to higher-structured actions. (Bicategorical versions of Para are a relatively old idea, and the double categorical version is a straightforward extension which has existed for some time as folklore, but the fully-internal construction here is novel).

In chapter 5, we apply Markov fibrations to compositional game theory, and solve a longstanding problem by giving a category of stochastic open games with a general "external choice" operator. This construction is largely abstract over the particular category, and so can be applied to different Markov fibrations to describe games defined with different notions of stochasticity. We also review some previous work with Capucci, Hedges, and Gavranovic on abstract constructions of categories of open games [Cap+22].

In chapter 6, we describe how to extend Myers' categorical systems theory to Markov fibrations, to give theories of systems which may have stochastic maps not merely in the fiber (as was already described by Myers in [Mye23]) but also in the base. We also leverage our construction from chapter 4 to give a description of a *symmetric monoidal triple category* whose three types of morphisms, are lenses, charts, and what we call *bisystems,* that is morphisms $\mathsf{TS} \otimes \mathsf{A} \leftrightarrows \mathsf{B}$—these are open dynamical systems which have two directions of interaction with the environment. We also describe a particular systems theory of smooth manifolds and smooth stochastic maps between them.

1.5 Related Work

We have attempted to provide relevant citations throughout the thesis whenever we touch on some preexisting theory. Below we give a summary of this and also try to note a few related bodies of work, discussion of which did not make it into the main body of the thesis.

The central problem of Markov fibrations is that reindexing is complicated in the presence of "non-copyable" maps. A very similar problem has been studied for some time under the name *Linear Dependent Types*. There are a number of approaches to this in the literature—Markov fibrations most closely match the approach of Krishnaswami et al, [KPB18], where there is a non-linear fragment and types simply can't depend on linear variables (matching how in a Markov fibration, a type can not be reindexed by a stochastic map). See also [Vák15]. A contrasting approach is *Quantitative Type Theory* (QTT) [McB16; Atk18; Bra21], which annotates binding sites with a *quantity*, allowing a variable to be utilized for *computation* only a certain number of times, but to be freely used to index dependent types. This idea has recently been further developed by Dorè [Dor25], who embeds a linear type system inside dependent type theory, allowing multiplicities to be computed dynamically. The precise relationship between these theories and Markov fibrations is not currently known. Finally in this context we may also mention Vákár's thesis [Vák17] on combining dependent types with *effectful* programs (i.e those corresponding to Kleisli morphisms for some monad).

One of our contributions is an extension of Myers' categorical systems theory to stochastic dynamical systems. There is a large literature within theoretical computer science studying such dynamical systems using *coalgebras*. For an overview of the theory of coalgebras, see [Jac16]. For a review of their application to stochastic systems specifically, see [Dob09; Dob07]. See eg [MMS07; Sok05; Jac08; Sil+13] for some examples of this theory in practice.

In addition to the above-mentioned literature on linear dependent types, our work is also connected to the theory of *probabilistic programming*, which studies syntax and semantics of programs whose intended meaning is some sort of probability distribution or kernel. Of particular interest in this field is the correct denotational semantics of probabilistic programs, which must combine the domain theory of ordinary denotational semantics with probability theory. Kozen's paper [Koz81] was probably the earliest in this field—for more recent work, see [Heu+17] and [Sta17]. We may also mention Stein's thesis [Ste21], advancing Markov categories as the right setting for probabilistic programming.

1.6 Acknowledgements

I would like to extend a great thanks to the Mathematically Structured Programming group at the University of Strathclyde, my fellow PhD students in particular, for making my time there such a great one. A good chunk of this work was completed while I worked at the Compositional Systems Laboratory at the Tallinn University of Technology—I owe everyone there a great thanks as well.

For conversations which contributed to this work, and to my mathematical development in general, I would like to thank Dylan Braithwaite, Matteo Capucci, Elena di Lavore, Tobias Fritz, Davidad, Tomáš Gonda, Bruno Gavranović, Diana Kessler, Owen Lynch, Jade Master, David Jaz Myers, Riu Nakamura, Chad Nester, Evan Patterson, Paolo Perrone, Mario Roman, Brandon Shapiro, Toby Smithe, Pawel Sobocinski, David Spivak, and Andre Videla. A special thanks to Dylan, Matteo, and Bruno, for serving their time with me in the dependent optics mines.

To Neil Ghani, Radu Mardare, and Jules Hedges, a massive thanks for serving as my advisors during the somewhat tumultuous process that was the production of this thesis.

Finally, I thank my wonderful family. Now you can stop asking me about it.

Chapter 2

Preliminaries

2.1 Introduction

We will begin by reviewing some theory that plays a key role in this thesis. With few exceptions, nothing here is novel, but we find it useful to include this here—both for the convenience of the reader, to familiarize them with theory that we will make constant reference to, but also to set the stage for our contributions.

First, the theory of (Grothendieck) fibrations. There is far too much to say about these for such a brief space, so we will limit ourselves to what we need for the rest of the thesis, especially for the section on Markov fibrations.

Second, we will give an overview of *optics* and *lenses*, in a bit more detail than the introduction. We have already given a review of the sources there, but we will find it useful to put this on proper footing.

Next, we will review the theory of Markov categories. This is a synthetic approach to probability theory, introduced by Fritz [Fri20], and since developed further by many collaborators, including the author. Here we note the only exceptions to the claim that nothing in this chapter is novel. First, the notion of *representable Markov category* Definition 2.4.3 was introduced by Fritz, Gonda, Perrone, and the author in [Fri+23b]. We will not give a thorough treatment here, but since representable Markov categories are so ubiquitous, we will frequently note how different properties or structure on a Markov category relates to representability. We will also introduce a few new concepts which play a role in the theory of Markov fibrations in chapter 3. These are of no great independent interest, as far as we can tell, nor are they difficult, but this seemed the best place to put them.

Finally, we give a brief review of Myers' *categorical systems theory*. This will mainly be to set the stage for chapter 6, where we develop a triple categorical version of the theory.

2.2 Fibrations

In category theory, there are many families of categories indexed by the objects of some other category. For example, for each commutative ring, we have the category $\mathsf{Mod}(R)$ of modules over R. Given a ring homomorphism $\phi : R \to S$, there is an induced *restriction of scalars* functor $\phi^* : \mathsf{Mod}(S) \to \mathsf{Mod}(R)$ (given simply by composing the module structure by ϕ), and this is (contravariant) functorial, assembling into a functor $\mathsf{Mod}(-) : \mathsf{CRing}^{\mathrm{op}} \to \mathsf{Cat}$.

In most cases, one can't expect strict functoriality as above. From an abstract point of view, it makes sense that one should really ask only for a natural isomorphism $\phi^*\psi^* \simeq (\psi\phi)^*$, up to some coherence conditions. This assembles into a so-called *pseudofunctor* into the 2-category Cat.

From a concrete point of view, there are many natural families of categories which arise as pseudofunctors. For example, restriction of scalars always has a left adjoint (extension of scalars, given by $M \mapsto M \otimes_R S$, viewing S as an R-module via the map ϕ)—since adjoints compose (that is, if $F \vdash G$ and $F' \vdash G'$, then $FF' \vdash G'G$) this must be functorial up to natural isomorphism, but this is the best we can promise.

To avoid the higher categorical algebra involved in working with pseudofunctors, Grothendieck introduced the notion of *fibration* in [Gro60].

Definition 2.2.1. Let $p : \mathcal{D} \to \mathcal{C}$ be a functor. Given $X \in \mathcal{C}$, write $\mathcal{D}_X$ for the (strict) pullback $\{X\} \times_{\mathcal{C}} \mathcal{D}$. Explicitly, this consists of the objects in $\mathcal{D}$ with $p(A) = X$ and the morphisms with $p(f) = 1_X$.

Let $f : X \to Y \in \mathcal{C}$ be a morphism.

1. A map $\bar{f} : \bar{X} \to \bar{Y}$ with $p(\bar{f}) = f$ is *locally Cartesian* if for each $\bar{X}'$ with $p(\bar{X}') = X$, postcomposition with $\bar{f}$ induces a bijection
$$\{g : \bar{X}' \to \bar{X} \mid p(g) = 1_X\} \xrightarrow{\sim} \{g' : \bar{X}' \to \bar{Y} \mid p(g') = f\}$$
2. A map is *Cartesian* if for every $g : Z \to X$ and $\bar{Z}$ with $p(\bar{Z}) = Z$, there is a bijection
$$\{\bar{g} : \bar{Z} \to \bar{X} \mid p(\bar{g}) = g\} \to \{\bar{g}' : \bar{Z} \to \bar{Y} \mid p(\bar{g}') = fg\},$$
note that every Cartesian map is locally Cartesian (take $g = 1_X$)
3. p is a *Grothendieck fibration* (or just fibration) if, for every $\bar{Y} \in \mathcal{D}$ such that $p(\bar{Y}) = Y$, there exists a Cartesian map $\bar{f} : \bar{X} \to \bar{Y}$ (for some $\bar{Y}$) so that $p(\bar{f}) = f$

Lemma 2.2.2. $p : \mathcal{D} \to \mathcal{C}$ *is a Grothendieck fibration if and only if every* f *admits a locally Cartesian lift, and the class of locally Cartesian morphisms in* $\mathcal{D}$ *is stable under composition.*

We will often abuse notation and simply denote a fibration $\mathcal{D} \to \mathcal{C}$ by the domain category $\mathcal{D}$. We say $\mathcal{D}$ is *fibred* over $\mathcal{C}$ if there is some implicitly-understood functor $\mathcal{D} \to \mathcal{C}$ which is a fibration.

Theorem 2.2.3. *Let* $\mathcal{D} \to \mathcal{C}$ *be a Grothendieck fibration. For every* $f : X \to Y, \bar{Y} \in \mathcal{D}_Y$, *select a Cartesian lift* $f^*\bar{Y} \to \bar{Y}$ *of* f. *Then there is a unique extension of* f^* *to a functor* $\mathcal{D}_Y \to \mathcal{D}_X$ *so that the squares*

$$\begin{array}{ccc} f^*\bar{Y} & \longrightarrow & \bar{Y} \\ \downarrow & & \downarrow \\ f^*\bar{Y}' & \longrightarrow & \bar{Y}' \end{array}$$

commute. With this, the assignment $X \mapsto \mathcal{D}_X, f \mapsto f^*$ *assembles into a pseudofunctor* $\mathcal{C}^{\mathrm{op}} \to \mathsf{Cat}$.

Remark 2.2.4. If $\mathcal{D} \to \mathcal{C}$ is such that each morphism admits merely a locally Cartesian lift (but these do not compose,) it is called a *prefibration*. Note that this is unrelated to our notion of Markov prefibration (reading ahead a bit, the "Cartesian" maps in a Markov prefibration do compose, but they enjoy the unique lifting property only for a subset of morphisms). This clash of terminology is perhaps unfortunate, but other potential prefixes seemed inferior (quasi-, pseudo-, semi-).

Example 2.2.5. Let $\mathcal{C}$ be any category, and let $\mathcal{C}^{\to}$ denote the arrow category. Then the codomain functor $\mathcal{C}^{\to} \to \mathcal{C}$ is a fibration if and only if $\mathcal{C}$ admits all pullbacks, and in this case the functors $f^* : \mathcal{C}_Y \to \mathcal{C}_X$, given $f : X \to Y$, are given by pullback along f.

Because of this, in a general fibration, the functors f^* are sometimes referred to as *pullback*, a convention we generally adopt. They are also sometimes called *base-change* functors.

When $f : X \to Y$ and $A \in \mathcal{D}_Y$, we may write A_X for the object f^*A if there is no chance of confusion. (Compare that the choice of f is also suppressed in the notation $A \times_Y X$ for a pullback)

We will not go into a comprehensive description of the theory of fibrations, but simply give a few basic results. We will give some examples in the next section. For a textbook treatment, see eg. [Jac99, Chapters 1, 9], or [Bor94, Chapter 8]. Note that we will not give a formal definition of the term "pseudofunctor" here. See eg [Jac99, def. 1.4.4]. See [JY20] for a more thorough discussion of the higher-categorical considerations, and [Buc14] for the theory of fibrations between higher categories. There have been a number of variations on the concept of fibration, see eg [LR20] for a review.

Definition 2.2.6. Let $\mathcal{A} : \mathcal{C}^{op} \to \mathsf{Cat}$ be a pseudofunctor. Then there is a category $\int_{X \in \mathcal{C}} \mathcal{A}(X)$ defined as follows:

1. The objects are pairs $\binom{\bar{X} \in \mathcal{A}(X)}{X \in \mathcal{C}}$
2. The morphisms $\binom{\bar{X}}{X} \to \binom{\bar{Y}}{Y}$ are pairs $f : X \to Y, f^{\#} : \bar{X} \to \mathcal{A}(f)(\bar{Y}) \in \mathcal{A}(X)$
3. Composition is given by the "chain rule" $(f, f^{\#}) \circ (g, g^{\#}) = (fg, \mathcal{A}(g)(f^{\#})g^{\#})$

There is an obvious forgetful functor $\int_X \mathcal{A}(X) \to \mathcal{C}$.

The category $\int_X \mathcal{A}(X)$ is known as the *Grothendieck construction of* $\mathcal{A}$

Proposition 2.2.7. $\int_X \mathcal{A}(X) \to \mathcal{C}$ *is a Grothendieck fibration, and the pseudofunctor it induces is equivalent to* $\mathcal{A}$

Proof. Given $f : X \to Y$ and $\bar{Y} \in \mathcal{A}(Y)$, it is clear that the map $\binom{\mathcal{A}(f)(\bar{Y})}{X} \to \binom{\bar{Y}}{Y}$ given by $f, 1_{\mathcal{A}(f)(\bar{Y})}$ is locally Cartesian—the required bijection is the definition of maps in the Grothendieck construction. But it's straightforward to see that these compose. □

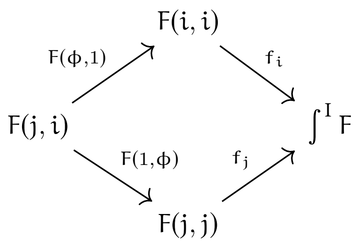
Given a pseudofunctor $\mathcal{A} : \mathcal{C}^{op} \to \mathsf{Cat}$, it is obvious that the assignment $\mathcal{A}(-)^{op}$ is pseudofunctorial as well (the required natural isomorphisms are just the opposites of the ones for $\mathcal{A}$). Applying this through the equivalence of fibrations and pseudofunctors leads to the *fiberwise opposite* of a fibration. Explicitly:

Corollary 2.2.8. *Let* $p : \mathcal{D} \to \mathcal{C}$ *be a fibration. Then there exists a category* $\mathcal{D}^{fop}$*, called the* fiberwise opposite of $\mathcal{D}$*, whose objects are the same as* $\mathcal{D}$*, and where a morphism* $X \to Y$ *is a tuple* $(f : p(X) \to p(Y), f^{\#} : f^*Y \to X \in \mathcal{D}_{p(X)})$.

2.3 Optics and Lenses

We have already given somewhat of an account of lenses, optics, and their applications in the introduction. We briefly review the theory here, especially to normalize the notation and definitions. The best general source for this material is still [Ril18].

The definition of optic relies on the notion of *coend*, which we briefly recall, see [Lor21] for a textbook account. If $F : \mathcal{J}^{op} \times \mathcal{J} \to \mathcal{C}$ is a functor, the *coend* $\int^{i \in I} F(i, i)$ is defined as the initial object receiving a map $f_i : F(i, i) \to \int^{i \in I} F(i, i)$ for each $i \in I$, so that for each $\phi : i \to j$, the square



commutes.

If $\mathcal{C}$ has enough colimits, we may express the coend as the coequalizer of the diagram

$$\coprod_{f : i \to j \in \mathcal{J}} F(i, j) \rightrightarrows \coprod_{i} F(i, i),$$

where the two maps are given on each component by $F(i, f) : F(i, j) \to F(i, i)$ and $F(f, j) : F(i, j) \to F(j, j)$, respectively. (This is Remark 1.2.4 in [Lor21]). Note that in particular this implies that if $\mathcal{C}$ has all (small) colimits, then it also has all (small) coends (that is, coends where J is small).

Definition 2.3.1 (Optic)**.** Let $\mathcal{M}$ be a monoidal category which acts on two categories $\mathcal{C}, \mathcal{D}$. Then the category of *optics* $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})$ has

1. Objects pairs $\binom{A \in \mathcal{D}}{X \in \mathcal{C}}$
2. The set of morphisms $\binom{A}{X} \to \binom{B}{Y}$ given by the coend

$$\int^{M \in \mathcal{M}} \mathcal{C}(X, M \cdot Y) \times \mathcal{D}(M \cdot B, A)$$

3. Given two optics with representatives $(M, f : X \to M \cdot Y, g : M \cdot B \to A), (N, f' : Y \to N \cdot Z, g' : N \cdot C \to B)$, their composite is given by $(M \otimes N, (1_M \cdot f')f, g(1_M \cdot g'))$, where we omit coherence morphisms.

When $\mathcal{C}$ is a monoidal category acting on itself by tensor, we write $\mathsf{Optic}_{\mathcal{C}}(\mathcal{C}, \mathcal{C}) =: \mathsf{Optic}(\mathcal{C})$

Note that if $\mathcal{M}, \mathcal{C}, \mathcal{D}$ are symmetric monoidal and these actions are symmetric, $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})$ inherits a symmetric monoidal structure given by $\binom{A}{X} \otimes \binom{B}{Y} = \binom{A \otimes B}{X \otimes Y}$. (Also given a braiding one can induce a non-symmetric monoidal structure, but this almost never comes up).

Remark 2.3.2 (On notation)**.** Objects and morphisms in the category of optics have two parts—one going "forwards", in the same direction as the optic, and one going "backwards". In [Ril18] the objects are written (X, A), where X is the forwards part. Hedges' work on open games used the binomial notation $\binom{X}{A}$, but wrote the forwards part on *top*.

To make the connection to fibrations more natural, we instead write the forwards part on the *bottom*, $\binom{A}{X}$. It is the backwards part which depends on the forwards part, hence the forwards part is the base of the fibration (when one exists)—and every part of the language of fibrations is built around a mental model where the base is at the bottom and the fibers are over it (including the word "base"). When reading the references, this may cause some confusion, but hopefully this can be overcome.

While we're at it, let us note that when talking about optics we will freely use terms like "the forwards part", "the backwards object", and so on—the meaning of this should now be clear. Of course, once we get to cooptics/charts, this would be more than a little confusing, since in that case both components are in the same direction. In those cases we will speak of either the primary (forwards) part or the secondary (backwards) part, or use the language of fibrations and speak of the map or object "in the base" and "in the fiber".

As noted above, the coend $\int^M \mathcal{C}(X, M \cdot Y) \times \mathcal{D}(M \cdot B, A)$ consists of triples $(M, f : X \to M \cdot Y, g : M \cdot B \to A)$ up to the equivalence relation which, for every $\phi : M \to M', f : X \to M \cdot Y, g : M' \cdot B \to A$, identifies the two tuples $(M', (\phi \cdot 1_Y)f, g)$ and $(M, f, g(\phi \cdot 1_B))$. Note that this relation is not assumed to be inherently an equivalence relation—one takes the transitive-symmetric closure as usual.

We call this relation the *sliding* relation (because we slide the map ϕ from the backwards part to the forwards part).

Remark 2.3.3. Suppose $\mathcal{M}, \mathcal{C}, \mathcal{D}$ are small categories. Then the coend defining $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})\left(\binom{A}{X}, \binom{B}{Y}\right)$ is a small colimit of small sets, hence again small. Since clearly the set of objects $\mathbf{ob}\,\mathcal{C} \times \mathbf{ob}\,\mathcal{D}$ is small, $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})$ is again a small category.

However, if $\mathcal{M}, \mathcal{D}, \mathcal{C}$ are merely assumed to be *locally* small, we can not guarantee the same is true of $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})$, since the hom-sets are now defined by a coend/colimit with large indexing category. However, in many special cases, it can still be seen to be locally small, such as in the Cartesian case (where $\mathsf{Lens}(\mathcal{C})$ is clearly locally small).

In this thesis, we will not delve further into this subtlety, simply working inside some universe where all our categories are small.

Of course, we can also let the arrows in $\mathcal{D}$ go in the same direction as $\mathcal{C}$. This does not seem to have played any role in the literature, but we will give this a name, as it is a useful example to have in mind for Markov fibrations (where we will construct our dependent optics, conceptually, as a fiberwise opposite)

Definition 2.3.4 (Co-Optics)**.** Given $\mathcal{M}$ acting on $\mathcal{C}, \mathcal{D}$, the category of *co-optics*, $\mathsf{coOptic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D})$ has objects pair $\binom{A \in \mathcal{D}}{X \in \mathcal{C}}$, and morphisms given by the coend

$$\int^M \mathcal{C}(X, M \cdot Y) \times \mathcal{D}(M \cdot A, B)$$

Remark 2.3.5. For any monoidal category $\mathcal{C}$, $\mathsf{Optic}_{\mathcal{C}}\left(\binom{A}{X}, \binom{I}{I}\right) = \mathcal{C}(X, A)$. One way to think of an optic $\binom{A}{X} \leftrightarrows \binom{B}{Y}$ is as a string diagram $X \to A$, but which has a hole with space for a morphism $B \to Y$. One inserts such a morphism by composing the optic with the optic $\binom{B}{Y} \leftrightarrows \binom{I}{I}$ representing it. This idea of "open diagrams" has been developed in much more detail by Román [Rom21].

Proposition 2.3.6. *Suppose $\mathcal{M}$ is semicartesian—in other words, that $I \in \mathcal{M}$ is terminal. Then there is a functor $\mathsf{Optic}_{\mathcal{M}}(\mathcal{C}, \mathcal{D}) \to \mathcal{C}$, which takes a $\binom{A}{X}$ to X, and pair $\langle f : X \to M \cdot Y, g\rangle$ to the composite $X \to M \cdot Y \to I \cdot Y \cong Y$.*

In the case of $\mathsf{Optic}(\mathcal{M})$, the map $\mathsf{Optic}(\mathcal{M})(\binom{I}{I}, \binom{A}{X}) \to \mathcal{M}(I, X)$ is a bijection.

The fact that in $\mathsf{Optic}(\mathcal{M})$, states (maps from the monoidal unit) on $\binom{A}{X}$ are given by states on X, while costates are given by maps $X \to A$ plays an important role in the use of optics to describe open games. See chapter 5 for more on this.

Let us say a few things about lenses.

Proposition 2.3.7. *If $\mathcal{C}$ is Cartesian monoidal,*

$$\mathsf{Optic}(\mathcal{C})\left(\begin{pmatrix}A\\X\end{pmatrix}, \begin{pmatrix}B\\Y\end{pmatrix}\right) \cong \mathcal{C}(X, Y) \times \mathcal{C}(X \times B, A)$$

Proof.

$$\int^{M} \mathcal{C}(X, M \times Y) \times \mathcal{C}(M \times B, A) \cong \int^{M} \mathcal{C}(X, Y) \times \mathcal{C}(X, M) \times \mathcal{C}(M \times B, A)$$
$$\cong \mathcal{C}(X, Y) \times \mathcal{C}(X \times B, A).$$

Here we use a general fact about coends, that $\int^{M} \mathcal{C}(X, M) \times F(M) \cong F(X)$—this has been called the *ninja Yoneda lemma*, see [Lor21]. □

In this case we sometimes write $\mathsf{Lens}(\mathcal{C})$ for the category $\mathsf{Optic}(\mathcal{C})$. We have the following fact:

Proposition 2.3.8. *The functor $\binom{A}{X} \mapsto X$, $\mathsf{Lens}(\mathcal{C}) \to \mathcal{C}$, is a fibration. The fiber $\mathsf{Lens}(\mathcal{C})_X$ has the following description:*

1. *Its objects are the objects of $\mathcal{C}$.*
2. *A map $A \to B \in \mathsf{Lens}(\mathcal{C})_X$ is a map $X \times B \to A \in \mathcal{C}$.*
3. *The composite of $X \times B \to A$, $X \times C \to B$ is given by composing the two into $X \times X \times C \to A$, then using the diagonal.*
4. *Given $f : X \to Y$, the pullback functor $\mathsf{Lens}(\mathcal{C})_Y \to \mathsf{Lens}(\mathcal{C})_X$ is given by precomposing by f.*

Proposition 2.3.9. *If $\mathcal{C}$ moreover admits pullbacks, there is a fibred functor $\mathsf{Lens}(\mathcal{C})^{\mathrm{fop}} \to \mathcal{C}^{\to}$, over $\mathcal{C}$, which carries an object $\binom{A}{X}$ to $X \times A \xrightarrow{\pi_X} X$, and a morphism $f : X \times A \to B$ to the map $X \times A \xrightarrow{\langle \pi_X, F\rangle} X \times B$ over X. This is fully faithful.*

This is the first way to see the maps of $(\mathcal{C}^{\to})^{\mathrm{fop}}$ as dependent lenses—they receive the category of lenses as a full subcategory.

Remark 2.3.10. The squares

$$\begin{array}{ccc} A \times X & \longrightarrow & A \times Y \\ \downarrow & & \downarrow \\ X & \longrightarrow & Y \end{array}$$

are pullbacks in any category with products, even if it does not admit pullbacks in general. It follows that the full subcategory of $\mathcal{C}^{\to}$ spanned by product projections is always fibred over $\mathcal{C}$, which is isomorphic to $\mathsf{coOptic}(\mathcal{C})$—the fiberwise dual is isomorphic to $\mathsf{Lens}(\mathcal{C})$.

2.4 Markov categories

Recall that *Markov categories* (Fritz, [Fri20]) are semicartesian symmetric monoidal categories equipped with a suitably coherent system of comonoids $(\mathrm{copy}_X : X \to X \otimes X, \mathrm{del}_X : X \to I)$ for each object $X \in \mathcal{C}$. These have been used as the basis for the categorical analysis of probability theory for a number of years, see eg [FR20; Fri+23a; Fri+25; MP23]. See [Fri20] for the basic theory, and [Fri+23b] for the theory of representable Markov categories. We assume familiarity with this theory in general, but we will review some theory below, as well as a few new minor results which are helpful for the theory of Markov fibrations.

The basic idea of a Markov category is to interpret morphisms $X \to Y$ as "stochastic processes" or *kernels*—that is, functions valued in probability measures. A morphism $P \to X \otimes Y$ is a parametrized joint probability measure—the comonoid structure allows us to build a canonical such given parametrized measures $P \to X, P \to Y$ (by precomposing their tensor with the copy_P map). This is the product measure of the two—the fact that in general not every map has this form (because $\mathcal{C}$ is not necessarily Cartesian) allows us to express the probabilistic dependence—as in, non-independence—of one variable on another.

A morphism is called *deterministic* if it is a comonoid (co)homomorphism, which amounts to the claim that $\mathrm{copy}_Y f = (f \otimes f)\mathrm{copy}_X$—in other words, that running two independent copies of the kernel (with the same input) is equivalent to running one and copying the output. The deterministic morphisms form a Cartesian monoidal subcategory which is denoted $\mathcal{C}_{\mathrm{det}} \subseteq \mathcal{C}$.

We first note the following alternative characterization of Markov categories in terms of their deterministic morphisms.

Definition 2.4.1 (Premarkov structure)**.** Let $\mathcal{C}$ be a semicartesian symmetric monoidal category. A *premarkov structure* on $\mathcal{C}$ is a wide symmetric monoidal subcategory $\mathcal{C}' \subseteq \mathcal{C}$—that is, a class of morphisms which contains all identities and structural isomorphisms, and is stable under composition and monoidal products—so that the monoidal category $\mathcal{C}'$ is Cartesian.

Proposition 2.4.2. *Let $\mathcal{C}$ be a symmetric semicartesian monoidal category.*

1. *Given a premarkov structure $\mathcal{C}' \subseteq \mathcal{C}$, there is a unique Markov structure on $\mathcal{C}$ so that each morphism in $\mathcal{C}'$ is deterministic.*

2. *Given a Markov structure, $\mathcal{C}_{\mathrm{det}} \subseteq \mathcal{C}$ is a premarkov structure.*

3. *A premarkov structure has the form $\mathcal{C}_{\mathrm{det}}$ for some Markov structure if and only if it is maximal.*

Proof. The existence part of the first claim is clear: $\mathcal{C}'$ acquires a unique Markov structure since it is Cartesian, and the inclusion of that Markov structure into $\mathcal{C}$ gives a Markov structure on $\mathcal{C}$. To prove uniqueness, suppose $\mathcal{C}$ is a Markov category and $\mathcal{C}' \subseteq \mathcal{C}_{\mathrm{det}}$ is a class of deterministic morphisms which is still Cartesian. This means the projections $X \otimes Y \to X, Y$ still exhibit $X \otimes Y$ as a product in $\mathcal{C}'$. But since the pairing are the unique map lifting two given maps $A \to X, Y$, the pairing must be preserved by the inclusion $\mathcal{C}' \to \mathcal{C}_{\mathrm{det}}$. Since the canonical Markov structure is given as a pairing, this means the Markov structure induced by $\mathcal{C}' \to \mathcal{C}$ must agree with the one given by $\mathcal{C}_{\mathrm{det}}$, which is just the original one.

The second claim is straightforward—see eg [Fri20] remark 10.13.

Now suppose $\mathcal{C}_{\mathrm{det}} \subseteq \mathcal{C}' \subseteq \mathcal{C}$, where $\mathcal{C}'$ is some larger premarkov structure. By the argument above, they must generate the same Markov structure on $\mathcal{C}$. But again, this implies that every map in $\mathcal{C}'$ is deterministic for this Markov structure, so we have $\mathcal{C}_{\mathrm{det}} = \mathcal{C}'$. This finishes the proof. □

Note that given a merely monoidal category, we can ask whether it is Cartesian, and if it is, it admits a unique symmetry induced by the universal property of the product. Hence if $\mathcal{C}' \subseteq \mathcal{C}$ is a Cartesian wide monoidal subcategory, we may attempt to define a symmetry on $\mathcal{C}$ simply using the one from $\mathcal{C}'$. However, it is not automatic that this symmetry is natural for all the morphisms in $\mathcal{C}$.

This basic idea was already noted by Fritz (and goes back to Golubtsov's work in [Gol02], an important part of the prehistory of Markov categories), although the precise statement above appears to be novel. Since the coherence conditions required of the comonoids in a Markov structure can be somewhat hard to remember, this characterization may be easier to understand.

Definition 2.4.3 (Representable Markov category)**.** A Markov category is *representable* if the inclusion $\mathcal{C}_{\mathrm{det}} \hookrightarrow \mathcal{C}$ admits a right adjoint. In this case we denote the right adjoint P and call the object PX for $X \in \mathcal{C}$ a *distribution object* for X. Observe that $\mathcal{C}(X, Y) = \mathcal{C}_{\mathrm{det}}(X, PY)$ (by definition,) and hence $\mathcal{C} = \mathrm{Kl}(P)$ where we denote the induced monad on $\mathcal{C}_{\mathrm{det}}$ P by an abuse of notation.

Example 2.4.4 (Commonly used Markov Categories)**.** Here are some important Markov categories:

1. Stoch ([Fri20, Section 4]) is the category whose objects are measurable spaces, whose morphisms are Markov kernels, with composition given by the Chapman-Kolmogorov equation and monoidal structure given by product measures.

2. $\mathsf{BorelStoch} \subseteq \mathsf{Stoch}$ ([Fri20, Section 4]) is the full subcategory of Stoch spanned by the *standard Borel spaces,* that is by those measurable spaces arising as the Borel σ-algebra on separable, complete metric space.

3. $\mathsf{FinStoch}$ is the subcategory of Stoch spanned by finite sets in the powerset σ-algebra. A morphism $X \to Y$ in $\mathsf{FinStoch}$ is equivalently a matrix $f_{xy} : x \in X, y \in Y$ with entries in $\mathbb{R}_{\geqslant 0}$ and with $\sum_y f_{xy} = 1$ for each x (this is what is called a *stochastic matrix*).

4. Let $\Delta : \mathsf{Set} \to \mathsf{Set}$ be the monad which assigns to $X \in \mathsf{Set}$ the set $\Delta(X)$ of countably-supported probability measures. Then the Kleisli category $\mathrm{Kl}(\Delta)$ is a Markov category, sometimes called the category of *discrete* probability.

5. TychStoch ([Fri+23a, Example A.1.4]) is the category of *Tychonoff* topological spaces and kernels which are valued in Radon probability measures, and where the measure $f(- \mid x)$ varies continuously in $x \in X$ with respect to the weak topology—in other words, given any continuous function $u \in C(Y)$, the resulting function on X given by $E_{y \sim f(-|x)} u(y)$ is continuous. (A space X is *Tychonoff* if it is Hausdorff and, given $K \subset X$ closed and x_0 not in K, there exists continuous $f : X \to [0,1]$ with $f(x_0) = 0, f(k) = 1$ for $k \in K$. Every locally compact Hausdorff space is Tychonoff).

Example 2.4.5 (Diagram Markov Categories)**.** If $\mathcal{C}$ is a Markov category and I is any ordinary category, there is a Markov category $\mathsf{Fun}(I, \mathcal{C})$ whose objects are functors $I \to \mathcal{C}_{\mathrm{det}}$, and whose morphisms are natural transformations between these considered as functors into $\mathcal{C}$ (i.e natural transformations with stochastic components). The monoidal and Markov structure is defined simply component-wise.

In particular, taking $I = \to = \{0 \to 1\}$ the walking arrow, we obtain a Markov category of deterministic arrows $\mathsf{Fun}(\to, \mathcal{C})$. We will denote this category simply $\mathcal{C}^{\to}$. Again, the objects of this category are the deterministic morphisms of $\mathcal{C}$, while the morphisms are the commutative squares with not-necessarily-deterministic sides

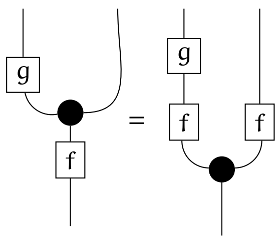
Remark 2.4.6 (A note on terminology)**.** When we speak of a *stochastic* map in a Markov category, we always mean a map which is *not necessarily deterministic*—that is, a general map of $\mathcal{C}$.

Given a map $p : P \to X \otimes Y$ into a tensor product, the composites with the projections $P \to X, P \to Y$ are called the *marginals*. Given two maps $f : P \to X, g : P \to Y$, we refer to any map $P \to X \otimes Y$ with those marginals as a *pairing* of the two. Note that there is always a canonical pairing given by $(f \otimes g)\mathrm{copy}_P$. We write $\langle f, g \rangle : P \to X \otimes Y$ for this canonical pairing. For deterministic maps this is just the usual pairing using the universal property of the product. (We mostly use this in cases where one map is deterministic, so that the pairing is unique assuming positivity). We write $\pi_X = 1_X \otimes \mathrm{del}_Y : X \otimes Y \to Y$ for the projection maps. When $p = \langle \pi_X p, \pi_Y p \rangle$, we say $X \perp Y \mid P$, and say the two coordinates are *independent given* P (or just independent).

Recall that the codomain functor $\mathcal{C}^{\to} \to \mathcal{C}$ is a fibration if and only if $\mathcal{C}$ admits pullbacks. Since Markov categories have terminal objects but not, in general, products, they clearly cannot be expected to have pullbacks. However, if $\mathcal{C}$ is *positive* ([Fri20, Definition 11.22]—and see Proposition 2.4.8 below), given $f : P \to X$ *deterministic*, every $g : P \to Y$ has a unique pairing $P \to X \otimes Y$ with f. Computing pullbacks $X \otimes_Z Y$ in the subcategory $\mathcal{C}_{\mathrm{det}}$, we usually have a similar property. This will be the basic idea behind Markov prefibrations.

By the following straightforward proposition, this property of having unique deterministic pairings is in fact equivalent to positivity.

Definition 2.4.7. A Markov category $\mathcal{C}$ is *positive* if, whenever $f : X \to Y, g : Y \to Z$ are such that gf is deterministic, we also have the equation



Proposition 2.4.8 (Characterization of positivity)**.** *Let $\mathcal{C}$ be a Markov category. The following are equivalent:*

1. *In $\mathcal{C}$, if $f : P \to X \otimes Y$ has deterministic marginal $P \to X$, then $X \perp Y \mid P$. In other words, anything is independent of a deterministic variable.*

2. *Given a deterministic morphism $P \to X$ and any morphism $P \to Y$, there is a unique $P \to X \otimes Y$ with those marginals.*

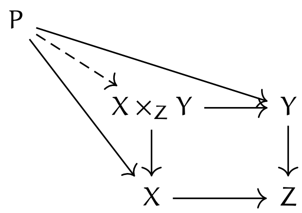
3. *$\mathcal{C}$ is positive.*

Proof. Clearly 1 and 2 are equivalent, since independence just means f is the independent pairing of the marginals—if it is uniquely determined by its marginals, it must be equal to the independent pairing, and conversely if it is necessarily independent, it is determined by its marginals. Now let us show this is equivalent to positivity.

First suppose $\mathcal{C}$ has unique pairings in this sense. Let f, g be as in the definition of positivity. Then the two maps $X \to Y \otimes Z$ indicated are pairings of gf and f, and since gf is deterministic, they are identical by hypothesis.

Finally, the implication $3 \Rightarrow 1$ is precisely [Fri20], prop. 12.14. □

Definition 2.4.9 (Pullback-positive)**.** Let $\mathcal{C}$ be a Markov category. We say $\mathcal{C}$ is *pullback-positive* if $\mathcal{C}_{\mathrm{det}}$ admits pullbacks and, given a diagram of this form where $P \to X$ is deterministic, there is a unique map $P \to X \times_Z Y$ making the squares commute.



Note that a pullback-positive category is in particular positive by taking $Z = I$.

Remark 2.4.10. The pullbacks appearing in Definition 2.4.9 are of course pullbacks in $\mathcal{C}_{\mathrm{det}}$, not $\mathcal{C}$, analogously to how $X \otimes Y$ is a product in $\mathcal{C}_{\mathrm{det}}$, not $\mathcal{C}$. We will still use the notation $X \times_Z Y$ for these pullbacks, which should not lead to any confusion. We may occasionally write $X \times Y$ instead of $X \otimes Y$, if we are carrying out a construction which primarily involves $\mathcal{C}_{\mathrm{det}}$. Since essentially no Markov categories have Cartesian products (except when the tensor product is Cartesian), this should also not lead to any ambiguity.

Lemma 2.4.11.

1. *Suppose any category $\mathcal{C}'$ admits products and* intersections*—that is, pullbacks $U \times_X V$ whenever U, V are subobjects of X. Then it admits all finite limits.*

2. *Suppose $\mathcal{C}_{\mathrm{det}}$ admits and $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ preserves pullbacks along monomorphisms. Suppose further $\mathcal{C}$ is positive. Then it is pullback-positive.*

Proof. To see the first point, let $f : X \to Z, g : Y \to Z$ be arbitrary maps. Note that $X \times_Z Y = (X \times Y) \times_{X \times Y \times Z \times Z} X \times Y \times Z$, where the horizontal map is given by $(x, y) \mapsto (x, y, f(x), g(y))$, and the top by $(x, y, z) \mapsto (x, y, z, z)$—in the sense that the universal property of this intersection is exactly the universal property of the given pullback.

$$\begin{array}{ccc} X \times_Z Y & \longrightarrow & X \times Y \times Z \\ \downarrow & \lrcorner & \downarrow \\ X \times Y & \longrightarrow & X \times Y \times Z \times Z \end{array}$$

To see this, observe that by the universal property of the product, to give maps into the latter pullback is to give two maps into X, two into Y, and one into Z, so that two possible maps into the product $X \times Y \times Z \times Z$ agree. These two maps agree if and only if they agree coordinate-wise. Their X-coordinates agree if the two given maps into X agree, and similarly Y. Thus write x, y, z for the maps into X, Y, Z. The first Z-coordinate agrees if $f(x) = z$, and the second Z-coordinate agrees if $g(x) = z$. This concludes the argument.

Now suppose $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ preserves this intersection, and $\mathcal{C}$ is positive. (Note that it doesn't follow that the inclusion preserves the pullback $X \times_Z Y$, because it doesn't preserve the products). This amounts to the claim that a map $P \to X \otimes Y$ lifts to the pullback if and only if the composite $P \to X \otimes Y \otimes Z \otimes Z$ lifts over the map $1_X \otimes 1_Y \otimes \mathrm{copy}_Z$ (note that the pullback of a monomorphism is a monomorphism).

This implies in particular such a lift is always unique. Since if $\mathcal{C}$ is positive, given $P \to X, P \to Y$, if the latter is deterministic there is a unique pairing $P \to X \otimes Y$, this implies there is at most one map $P \to X \times_Z Y$ pairing the two. On the other hand, since the composite $P \to Y \to Z$ is deterministic and equal to the composite $P \to X \to Z$, they are both deterministic, and hence the pairing $P \to Z \otimes Z$ factors over the diagonal. Applying positivity again, to the tensor product $(X \otimes Y) \otimes (Z \otimes Z)$, since the latter component is deterministic, this pairing is independent, and hence the map $P \to X \otimes Y \to X \otimes Y \otimes Z \otimes Z$ does indeed lift, finishing the proof. □

The idea here is that a distribution on a subobject $X' \subseteq X$ defined by some condition $f(x \in X) \in U \subseteq Y$ is simply a distribution so that the condition is satisfied with probability 1. This is a natural condition which holds in many Markov categories.

The assumption that $\mathcal{C}$ is pullback-positive will play a key role in the development of the theory of Markov fibrations. Although the theory could possibly be developed without assuming the base category has deterministic pullbacks, positivity seems to be an essential part.

Lemma 2.4.12. *If $\mathcal{C}$ is representable and positive, $\mathcal{C}_{\mathrm{det}}$ has finite limits, and the monad P preserves intersections, then $\mathcal{C}$ is pullback-positive.*

Proof. This is clear, since if the monad P preserves a given limit, so does the inclusion into the Kleisli category (for completely abstract reasons), hence by Lemma 2.4.11 we are done. □

We will need the notion of *support* in a Markov category, introduced in [Fri+23a], for certain examples, so we briefly record the definition and a few of its properties here.

Definition 2.4.13. Let $p : A \to X$ and $f, g : X \otimes Y \to Z$ be morphisms in a Markov category. Then we say f, g are p-almost surely equal if

$$(X \otimes f)(\mathrm{copy}_X\, p \otimes Y) = (X \otimes f)(\mathrm{copy}_X\, p \otimes Y),$$

in other words, (for every y and a) the joint distribution of x, z does not depend on whether we use f or g.

This recovers the usual definition of almost sure equality.

Definition 2.4.14 (Support of a morphism)**.** Let $p : X \to Y, q : A \to Y$ be two morphisms in a Markov category. We say q *is absolutely continuous with respect to* p and write $q << p$ if, whenever two maps $Y \to Z$ are p-almost surely equal, they are also q-almost surely equal.

Let $p : X \to Y$ be a morphism in a Markov category. The *support* of p, if it exists, is an object which represents the functor $\mathcal{C}(-, Y)_{<<p}$ of morphisms into Y which are absolutely continuous with respect to p.

Proposition 2.4.15.

1. *The support of* $p : X \to A$ *is equipped with a canonical deterministic monomorphism* $S_p \hookrightarrow A$*, so that a map into* A *is absolutely continuous with respect to* p *if and only if it factors over the support.*

2. *Two maps* $A \otimes W \to B$ *are* p*-almost surely equal if and only if they are strictly equal on the support of* p.

2.5 Double Categories

Definition 2.5.1 (Double Category)**.** A (strict) *double category* is a category internal to the category Cat of categories. Concretely, it consists of:

1. A set of objects $\mathbf{ob}\,\mathbb{C}$

2. A collection of *vertical morphisms* forming a category $\mathbb{C}_v$ with $\mathbf{ob}\,\mathbb{C}_v = \mathbf{ob}\,\mathbb{C}$

3. A collection of *horizontal morphisms* forming a category $\mathbb{C}_h$ with $\mathbf{ob}\,\mathbb{C}_h = \mathbf{ob}\,\mathbb{C}$

4. A collection of *squares*. Each square has a left and right boundary given by vertical morphisms l, r, and top and bottom boundary given by horizontal morphisms t, b so that $\mathrm{dom}\, t = \mathrm{dom}\, l, \mathrm{cod}\, t = \mathrm{dom}\, r$ and so on:

$$\begin{array}{ccc} X & \xrightarrow{t} & Y \\ \big\downarrow{\scriptstyle l} & & \big\downarrow{\scriptstyle r} \\ A & \xrightarrow{b} & B \end{array}$$

   The squares compose horizontally and vertically in the obvious way, each of which form a category (in particular, there are *identity squares* for each vertical and horizontal map).

A double functor is a mapping on objects, vertical and horizontal morphisms, and squares, which preserves all the identities and composition.

There is also a notion of *pseudo double category*, which weakens the horizontal composition to only be associative and unital up to a coherent system of squares. We will not go into the details here, see chapter 4 for more on this.

Example 2.5.2. There is a pseudo double category $\mathbb{C}\mathsf{at}$ where the objects are categories, the vertical maps are functors, the horizontal maps are *profunctors* (functors $\mathcal{C} \times \mathcal{D}^{op} \to \mathsf{Set}$, sometimes called bimodules), and the squares are natural transformations.

For any category with pullbacks $\mathcal{C}$, there is a pseudo double category $\mathbb{S}\mathsf{pan}(\mathcal{C})$ with $\mathcal{C}$ as the vertical category, spans as the horizontal morphisms, and commutative diagrams as the squares.

There is a double category $\mathbb{R}\mathsf{el}$ of sets, functions, and relations.

For any Markov category $\mathcal{C}$, there is a double category with $\mathbb{C}_v = \mathcal{C}_{\mathrm{det}}$ and $\mathbb{C}_h = \mathcal{C}$.

For any category at all, there is a double category $\mathrm{Sq}(\mathcal{C})$ with $\mathrm{Sq}(\mathcal{C})_v = \mathrm{Sq}(\mathcal{C})_h = \mathcal{C}$ and commutative squares as the pullback squares.

Definition 2.5.3. A double category is *thin* if, for each compatible square of vertical and horizontal morphisms, there is at most one square filling it. In other words, such a square either commutes or doesn't.

Sometimes the two classes of morphism are instead called *loose* and *tight*, especially in cases where the composition of the loose class is not associative, or if the loose class is a superset of the tight class.

Although the explicit description above is probably the best way to think about the data of a double category, on a technical level it is often useful to think in terms of internal categories. An *internal category* in a category $\mathcal{C}$ is a pair of objects C_0, C_1, maps $d, c : C_1 \to C_0, i : C_0 \to C_1$ (the domain, codomain and identity), and a map $m : C_1 \times_{C_0} C_1 \to C_1$ (the multiplication), satisfying the usual laws of a category. A double category in the above sense is then an internal category in the category of categories.

In the terms of Definition 2.5.1, C_0 is $\mathbb{C}_v$, and C_1 is the category whose *objects* are horizontal arrows, and whose morphisms are squares (composed vertically). Obviously, we could just as easily have oriented things horizontally, but this is the convention usually adopted.

Given a double category $\mathbb{C}$, there is another double category $\mathbb{C}^\top$ called the *transpose* of $\mathbb{C}$, which has the same objects but exchanges the horizontal and vertical morphisms. In many cases it is not clear which of $\mathbb{C}$ and $\mathbb{C}^\top$ is the "correct" one to work with, and we may have to pass back and forth between them. We try to stick to the convention that the horizontal morphisms are the "loose" ones. To avoid confusion, we will also simply specify the classes of morphisms directly, speaking for example of "the lenses" or "the charts" when working with the double category $\mathbb{A}\mathsf{rena}(\mathcal{A})$ (Definition 2.6.3).

The concept of double category goes back to [Ehr63]. See [JY20, Section 12.3] for a textbook treatment. They have been widely used in applied category theory. An early application is in [BC20], which constructed a "structured" version of the double category of cospans. Here the philosophy is that the horizontal morphisms are the "systems" (in a general sense) which are being wired together by horizontal composition, while the vertical morphisms (and squares) are "structure-preserving maps between systems". This is similar to the our double category of "bisystems" (section 6.4). Recent work by Lambert and Patterson [LP24] applies double categories to categorical algebra, with many applications to systems modeling, see [Car24] (as well as to classical category theory). In the next section we'll see their application to categorical systems theory, and later see them applied to parametrized morphisms.

2.6 Review of Categorical Systems Theory

We will give a review of the framework called *categorical systems theory*, due to Myers [Mye23].

Definition 2.6.1 (Theory of Dynamical Systems)**.** A *theory of dynamical systems* is an indexed category $\mathcal{A} : \mathcal{C}^{\mathrm{op}} \to \mathsf{Cat}$ equipped with a section of its Grothendieck construction $T : \mathcal{C} \to \int \mathcal{A}$.

It is called *monoidal* if $\mathcal{C}$ is a monoidal category, and $\mathcal{A}$ is a lax monoidal functor. It is further called symmetric monoidal if $\mathcal{C}$ is a symmetric monoidal category and $\mathcal{A}$ is a symmetric lax monoidal functor. Note that this is equivalent to requiring $\int \mathcal{A} \to \mathcal{C}$ to be a (symmetric) monoidal fibration in the sense of [Shu09] (see also [MV20]).

Definition 2.6.2 (Lenses and charts)**.** Let $\mathcal{A} : \mathcal{C}^{\mathrm{op}} \to \mathsf{Cat}$ be an indexed category. A *chart* is simply a morphism in the Grothendieck construction $\int \mathcal{A}$. A *lens* is a morphism in the fiberwise opposite $\int \mathcal{A}(-)^{\mathrm{op}}$.

Definition 2.6.3 (Arenas)**.** Given an indexed category $\mathcal{A} : \mathcal{C}^{\mathrm{op}} \to \mathsf{Cat}$, the double category $\mathbb{A}\mathsf{rena}$ of *arenas* has

1. Objects the objects of $\int \mathcal{A}$—note that these are the same as the objects of $\int \mathcal{A}(-)^{\mathrm{op}}$
2. Vertical morphisms the lenses
3. Horizontal morphisms the charts
4. The double category is thin. Given a square of this form

$$\begin{array}{ccc} \binom{\bar{A}_1}{A_1} & \longrightarrow & \binom{\bar{B}_1}{B_1} \\ \downarrow & & \downarrow \\ \binom{\bar{A}_2}{A_2} & \longrightarrow & \binom{\bar{B}_2}{B_2}, \end{array}$$

we can first project it to a square in $\mathcal{C}$. If this commutes, we can pull the lenses and charts back to a square in $\mathcal{A}(A_1)$. The above square of lenses and charts will be said to commute if both of these squares commute.

We will write arenas either as $\binom{\bar{A} \in \mathcal{A}(A)}{A \in \mathcal{C}}$, or, for brevity when there is no need to discuss the two levels separately, simply with a symbol A. We will also write both $\binom{TS}{S}$ and TS as the situation calls for—there should be no confusion resulting from this.

Definition 2.6.4 (Dynamical system with interface A)**.** Let A be an arena. A *dynamical system* with interface A is an object $S \in \mathcal{C}$ and a lens $TS \leftrightarrows A$. A morphism of systems is a map $S \to S'$ so that this square commutes in $\mathbb{A}\mathsf{rena}$:

$$\begin{array}{ccc} TS & \longrightarrow & A \\ \downarrow & & \| \\ TS' & \longrightarrow & A \end{array}$$

The category of systems with interface A is denoted $\mathsf{Sys}(A)$

Proposition 2.6.5. *Given a lens* $A \leftrightarrows A'$, *postcomposition gives a functor* $\mathsf{Sys}(A) \to \mathsf{Sys}(A')$. *Given a chart* $A \rightrightarrows A'$, *there is a* profunctor $\mathsf{Sys}(A) \nrightarrow \mathsf{Sys}(A')$, *where the set over* $TS \leftrightarrows A$ *and* $TS' \leftrightarrows A'$ *is the set of maps* $S \to S'$ *so that this square commutes:*

$$\begin{array}{ccc} TS & \longrightarrow & A \\ \downarrow & & \downarrow \\ TS' & \longrightarrow & A' \end{array}$$

This defines a double functor $\mathsf{Sys} : \mathbb{A}\mathsf{rena} \to \mathbb{C}\mathsf{at}$.

The fibration $\int \mathcal{A} \to \mathcal{C}$ encodes what sort of information can be "indexed over a space", while the section $\mathsf{T} : \mathcal{C} \to \int \mathcal{A}$ tells us what sort of information (such as a next step, or a gradient vector) must be produced to give a dynamical system. In this respect, the theory is very similar to the coalgebraic approach to dynamical systems, or systems that "do something", and indeed we have the following:

Example 2.6.6. Let $\mathsf{F} : \mathsf{Set} \to \mathsf{Set}$ be a functor. Then there is a dynamical systems theory where $\mathcal{C} = \mathsf{Set}$, $\mathcal{A}$ is constant at Set, and $\mathsf{T}(\mathsf{X}) = (\mathsf{X}, \mathsf{F}(\mathsf{X}))$. It is easy to see that the category of closed dynamical systems in this theory is equivalently the category of F-coalgebras.

Of course, coalgebras can already encode systems with input and output—the point of the CST framework is to separate out the input and output of systems so that they can be acted on in a compositional manner.

Example 2.6.7 (The theory of discrete dynamical systems)**.** There is a theory of dynamical systems with category of spaces $\mathcal{C} = \mathsf{Set}$, category of bundles $\mathcal{A}(\mathsf{X}) = \mathsf{Set}_{/\mathsf{X}}$ (with pullbacks for reindexing), and tangent bundle $\mathsf{TX} = \mathsf{X} \times \mathsf{X} \xrightarrow{\pi_0} \mathsf{X}$. We call this the theory of *discrete dynamical systems*. (In the sense that they are both discrete-time and discrete-space).

In another direction, we have the following comparison result:

Example 2.6.8. In the theory of discrete dynamical systems:

1. The category of lenses $\int \mathcal{A}(-)^{\mathrm{op}}$ is equivalently the category of *polynomial functors* and natural transformations between them.
2. Using this equivalence, the category of dynamical systems with interface p is exactly the category of p-coalgebras and homomorphisms

The first part here (which works, suitably formulated, for any locally Cartesian closed category), is a classical part of the theory of polynomial functors, see eg [GK13]. The second part is due to Spivak, see [Spi20]. There is an extensive body of work on the description of interacting (discrete, deterministic) dynamical systems in terms of polynomial functors, see e.g. also [SS23].

Example 2.6.9 (The theory of stochastic discrete dynamical systems)**.** There is a systems theory where $\mathcal{C} = \mathsf{Set}, \mathcal{A}(\mathsf{X}) = \mathsf{Kl}(\Delta)^{\mathsf{X}}$ (with the evident reindexing maps) and $\mathsf{T}(\mathsf{X}) = \mathsf{X} \times \mathsf{X}$.

In this theory, a closed dynamical system is a set equipped with a map $\mathsf{X} \to \Delta(\mathsf{X})$. An open dynamical system has stochastic *update* function, but output which depends deterministically on the current system state.

In fact this example works for any monad on the category of sets.

Example 2.6.10 (The theory of smooth dynamical systems)**.** There is a dynamical systems theory where $\mathcal{C} = \mathsf{SmMfd}$ is the category of smooth manifolds, $\mathcal{A}_\mathsf{X}$ is the category of fiber bundles over X, with $\mathsf{f}^* : \mathcal{A}_\mathsf{Y} \to \mathcal{A}_\mathsf{X}$ given by pullback along $\mathsf{f} : \mathsf{X} \to \mathsf{Y}$ (note that pullbacks of bundles always exist), and with TX being the tangent bundle of X in the ordinary sense. In this case, closed dynamical systems are manifolds equipped with (smooth) vector fields, which is what is classically thought of as a smooth dynamical system.

Example 2.6.11 (Clock system)**.** Consider the theory of discrete-time, discrete dynamical systems from Example 2.6.7. Let $\xi : \mathsf{T}S \leftrightarrows \binom{I}{O}$ be a system. Consider the system $c : \mathsf{T}\mathbb{N} \leftrightarrows \binom{*}{\mathbb{N}}$ given by $n \mapsto n$ in the forwards direction, and $(n, *) \mapsto n + 1$ in the backwards direction. (To be clear, the object denoted $\binom{*}{\mathbb{N}}$ is the map $\mathbb{N} \to \mathbb{N}$—it is a singleton *in each fiber*). Then a morphism of systems $c \to \xi$ consists of the following data:

1. A function $x : \mathbb{N} \to S$.
2. Another function $o : \mathbb{N} \to O$.
3. A third function $i : \mathbb{N} \to I$ so that $i(n) \in I_{o(n)}$.
4. So that $o(n) = \xi(x(n))$ and $\xi^{\#}(x(n), i(n)) = x(n + 1)$

That is, it is a choice of a sequence x_n of points in the state-space, and a sequence of inputs i_n compatible with the outputs, so that this sequence *obeys the dynamics*. In other words, it is a *trajectory of the system.*

Because of this, one views a generic chart map as a *generalized trajectory*, of a type given by the domain system. As another example, taking the state space to be $\{1, 2, \ldots, n\}$ with an update map that carries n to 1, one finds a system which classifies n-*periodic* trajectories. Similarly, in the smooth case, the system $(\mathbb{R}, d/dt : \mathbb{R} \to \mathsf{T}\mathbb{R})$ classifies solutions of a smooth differential equation (those which extend to infinity). Recent work by Lynch, Myers, Staton, and the author [Lyn+25] constructs clock systems for theories of stochastic, discrete-time doctrines, although we will not delve into this here.

Because this thesis is about Markov fibrations, we have chosen to present these ideas in terms of fibrations equipped with sections. Myers' book [Mye23] actually prefers the presentation in terms of indexed categories. Similarly, we construct a double category of systems which is fibred (in a certain sense) over the double category of arenas. Myers instead displays this as a *doubly indexed category* $\mathbb{A}\mathsf{rena} \to \mathbb{C}\mathsf{at}$, which carries lenses to functors and charts to profunctors.

In a recent paper [LM25], Myers and Libkind further develop the category theory of what they term *double operadic categorical systems theory,* which again concerns notions of "composable system" which are described in terms of such doubly indexed category (although for technical reasons, they use the language of *right modules* in that paper and a somewhat different presentation, the concept is the same.)

Chapter 3

Markov Fibrations and Stochastic Modules

3.1 Introduction

Let $\mathsf{Kl}(\Delta)$ denote the Kleisli category of the discrete (countable) distribution monad on Set. This is a simple setting for working with probability theory—sufficient for many applications. In order to study compositional Bayesian game theory [BHZ19; Cap+22] one studies the category $\mathsf{Optic}(\mathsf{Kl}(\Delta))$ of optics in $\mathsf{Kl}(\Delta)$. These have the right level of expressivity to talk about players taking random actions, and about payoff which depends stochastically on players' decisions.

In $\mathsf{Optic}(\mathsf{Kl}(\Delta))$, $\binom{R}{X} + \binom{R}{Y} \cong \binom{R}{X+Y}$. Games with this codomain naturally describe the situation of a player who has a binary choice between X or Y. We call coproducts of this form the "good" coproducts—note that also $\binom{*}{*} + \binom{\emptyset}{*} = \binom{\emptyset}{2}$, but this is considered somewhat pathological, since it relies on the nonexistence of any morphism $\binom{\emptyset}{X} \to \binom{A}{Y}$ when A and X are nonempty.

In order to describe this structure on $\mathsf{Optic}(\mathsf{Kl}(\Delta))$, it would be useful if it had all coproducts. Unfortunately this is not the case. $\mathsf{Optic}(\mathsf{Set}) = \mathsf{Lens}(\mathsf{Set})$ has a well-known extension with all coproducts, given by the fiberwise opposite of the fibration $\mathsf{Fam}(\mathsf{Set}) \to \mathsf{Set}$ (this is Proposition 2.3.9 in the case $\mathcal{C} = \mathsf{Set}$). Extending this to stochastic maps would be the obvious way of constructing such a category of "dependent optics".

Consider a category $\mathsf{Kl}(\Delta)^{\rightarrow}$ defined as follows. Its objects are the objects of $\mathsf{Set}^{\rightarrow}$—that is, they are indexed families of sets. A map in $\mathsf{Kl}(\Delta)^{\rightarrow}$ is a stochastic map in the base $f : X \to Y \in \mathsf{Kl}(\Delta)$, and a stochastic map on the total spaces $f' : \bar{X} \to \bar{Y}$ which is compatible with it. Note that if $\bar{X} \to X$ is surjective, the map on the base is fully determined by the map on the fibers, which must merely satisfy the condition that the distribution of the indexing point in Y depends only on the indexing point in X, not the specific point in the fiber $\bar{X}_x$.

$$\begin{array}{ccc} X \times_Z Y & \longrightarrow & X \times Y \times Z \\ \big\downarrow & \lrcorner & \big\downarrow \\ X \times Y & \longrightarrow & X \times Y \times Z \times Z \end{array}$$

$$\begin{array}{ccc} \bar{X} & \xrightarrow{f'} & \bar{Y} \\ \big\downarrow & & \big\downarrow \\ X & \xrightarrow{f} & Y \end{array}$$

We claim $\mathrm{Kl}(\Delta)^{\rightarrow}$ is a reasonable notion of "stochastic charts". Recall that by "chart" we mean something like "lenses where both maps go forward". If stochastic *lenses* are supposed to include optics as the full subcategory spanned by the "non-dependent" objects, then the charts should include "co-optics"—that is, maps between $X' \otimes X \to X$ and $Y' \otimes Y \to Y$ should be given by the coend $\int^M \mathrm{Hom}(X, M \otimes Y) \times \mathrm{Hom}(X' \otimes M, Y')$

And in fact this is the case: Clearly there is a map from this coend to maps in $\mathrm{Kl}(\Delta)^{\rightarrow}$. By taking $M = X \otimes Y$ and conditioning on Y, we see this is surjective. Finally, by restricting to the support of the forwards part inside $X \otimes Y$, we obtain a representative for each element of the coend which is uniquely determined by $X \to Y$ and $X' \times X \to Y'$ (since the conditional is well-defined on the support).

Note: This relies both on the fact that $\mathrm{Kl}(\Delta)$ has conditionals, and on the existence of supports. We've previously seen these defined in abstract Markov categories (Definition 2.4.14). Supports in $\mathrm{Kl}(\Delta)$ of a morphism $p : A \to B$ are simply given by those b so that $p(b|a) > 0$. Note that the existence of *both* conditionals and supports is a very strong assumption—the only categories we are aware of with both properties are those whose probability distributions have a discrete character, like $\mathrm{Kl}(\Delta)$ and $\mathsf{FinStoch}$. Neither will be essential to the theory, but both will play a role in certain theorems—we will see more of this later.

The goal of the theory of Markov fibrations is to give a notion of "fiberwise opposite" which can be applied to the codomain functor $\mathrm{Kl}(\Delta)^{\rightarrow} \to \mathrm{Kl}(\Delta)$ to give a reasonable notion of "stochastic lenses". In particular, we should recover the usual category of optics in the previous case.

It is clear that the codomain functor is not a (Grothendieck) fibration, since this would require $\mathrm{Kl}(\Delta)$ to have pullbacks, which can only hold for a Cartesian Markov category. However, we can do some things. Namely, given a Cartesian (pullback) square in Set

$$\begin{array}{ccc} \bar{A} & \longrightarrow & \bar{B} \\ \big\downarrow & & \big\downarrow \\ A & \longrightarrow & B \end{array}$$

and a map $\binom{\bar{X}}{X} \to \binom{\bar{B}}{B}$ where the base map $X \to B$ is deterministic, for each deterministic factorizing map $X \to A$ there is a unique lift $\bar{X} \to \bar{A}$. In other words, the pullback over $\mathsf{Set} \hookrightarrow \mathrm{Kl}(\Delta)$ is a fibration—in fact, it is simply the family fibration $\mathsf{Fam}(\mathsf{Kl}(\Delta))$.

Furthermore, if $\bar{X} \to \bar{B}$ is itself deterministic, there is such a unique lift even without assuming that the factorization $X \to A$ is deterministic.

Moreover, we can factor any map in $\mathsf{Kl}(\Delta)^{\to}$ as such an induced map followed by a map over a deterministic base, as follows:

$$\begin{array}{ccccc} \bar{X} & \leftarrow\!\dashrightarrow & \bar{X}\otimes Y & \xrightarrow{\phi} & \bar{Y} \\ \downarrow & & \downarrow & & \downarrow \\ X & \underset{p}{\leftarrow\!\dashrightarrow} & X\otimes Y, & \xrightarrow{p'} & Y \end{array}$$

This gives us a hope that we can, in some way, control the category $\mathsf{Kl}(\Delta)^{\to}$ using the pullback over Set, which *is* a fibration, and some information somehow given by these extra maps. Note also that the diagram above is equivalent to giving: a span $X \leftarrow M \to Y$ and a section $X \to M$, which all lives in the base, and a map $p^*\bar{X} \to p'^*\bar{Y}$ in the fiber over M. Thus it would seem to be very amenable to fiberwise dualization.

Analogous to our argument above that "co-optics" are equivalent to maps in $\mathsf{Kl}(\Delta)^{\to}$, we can do the following:

Suppose given two tuples $(M_0, p_0, p'_0, s_0, \phi_0), (M_1, p_1, p'_1, s_1, \phi_1)$ as above. Suppose there exists a map $f : M_0 \to M_1$ over X, Y, so that $fs_0 = s_1$. Then there is a canonical map $p_0^*\bar{X} \to p_1^*\bar{X}$ over f, because pullbacks commute. If the triangle

$$\begin{array}{ccc} p_0^*\bar{X} & \longrightarrow & p_1^*\bar{X} \\ & \searrow & \downarrow \\ & & \bar{Y} \end{array}$$

moreover commutes, then these two triples represent the same map in $\mathsf{Kl}(\Delta)^{\to}$

These equivalence relations correspond to "sliding" for *deterministic* morphisms $M \to M'$. Note that the condition here can be checked just on the fibration $\mathsf{Kl}(\Delta)^{\to} \times_{\mathsf{Kl}(\Delta)} \mathsf{Set} \to \mathsf{Set}$.

To obtain the full set of sliding equations, we will need to use stochastic maps $M \to M'$, and thus leave that fibration behind. However, we are tantalizingly close to realizing $\mathsf{Kl}(\Delta)^{\to}$ as being presented by some sort of additional structure on the fibration $\mathsf{Fam}(\mathsf{Kl}(\Delta))$. (For a general monad T acting on C, the category $\mathsf{Optic}_{\mathcal{C}}(\mathsf{Kl}(T), \mathsf{Kl}(T))$, of effectful optics up to sliding of pure morphisms, was studied by Riley in [Ril18, Section 4.9], and by Hedges in [Hed24])

In this chapter we will indeed provide such a structure, and analyze it. In section 3.2, we'll axiomatise the lifting property of $\mathsf{Kl}(\Delta)^{\to}$ discussed above into a property we call a *Markov prefibration* (Definition 3.2.2). In section 3.3, we exhibit a free Markov prefibration associated to a fibration—its morphisms are precisely spans of the form seen above. Naturally, given a prefibration $\mathcal{D} \to \mathcal{C}$, its underlying fibration on $\mathcal{C}_{\mathrm{det}}$ becomes an algebra for the monad of this adjunction. We term such algebras *stochastic module fibrations* (or merely *stochastic modules*). These consist essentially of a fibration plus an operation which, given a parameterized map $M \times A \to B$ (a map between indexed spaces) and a distribution on M (a map in the base category), produces a composite map $A \to B$.

Given any adjunction $L : \mathcal{A} \rightleftharpoons \mathcal{B} : R$, and an algebra for the associated monad, there is a canonical way to build an object of $\mathcal{B}$. In particular, for any stochastic module fibration, there is a "presented" category over $\mathcal{C}$. Applying this to the fiberwise opposite stochastic module, we obtain a useful construction of *stochastic lenses*. Using this notion, we can prove (see Definition 3.4.6):

Theorem 3.1.1. *There exists a (strong) symmetric monoidal functor* $\mathsf{Optic}(\mathsf{Kl}(\Delta)) \to \mathsf{SLens}(\mathsf{Kl}(\Delta)^{\to})$ *which*

1. *Is fully faithful.*
2. *Preserves the good coproducts.*
3. *So that the image has all coproducts, and every object of the image is a coproduct of objects in* $\mathsf{Optic}(\mathsf{Kl}(\Delta))$
4. *And so that moreover the monoidal structure on the image is distributive*

In addition, some Markov prefibrations are presented by the stochastic module they induce—these we term the *Markov fibrations*.

Following this, we review a few properties of the theory of stochastic modules and Markov fibrations, in particular, their stability under limits (section 3.5), and induced monoidal structures (section 3.6).

3.2 Markov Prefibrations

Remark 3.2.1. Markov categories generally do not have pullbacks, for the same reason that they usually don't have products. This issue generally hinders the construction of fibrations, in the ordinary sense, of Markov categories. However, we can go part of the way. The idea of the following definitions is that given a pullback in the deterministic category, say $A \times_Y X$, a map $P \to A \times_Y X$ *where the* X*-coordinate is deterministic* should be uniquely determined by a choice of (deterministic) map $P \to X$ and (stochastic) $P \to A$ such that the square commutes—as we claimed above (and will see below), this holds for the Markov category of discrete probability $\mathsf{Kl}(\Delta)$. The analogous statement for products—that a map $P \to A \otimes X$ with deterministic X-component is uniquely determined by the projections (or marginals) $P \to X, P \to A$—is a consequence of positivity (see Proposition 2.4.8), and hence holds in most Markov categories of interest.

Definition 3.2.2 (Markov prefibration)**.** Let $\mathcal{C}$ be a Markov category, and let $p : \mathcal{D} \to \mathcal{C}$ be a functor into it. Then we call p a *Markov prefibration* if the following two conditions hold:

1. The pullback $\mathcal{D} \times_{\mathcal{C}} \mathcal{C}_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$ is a (Grothendieck) fibration
2. Given maps $f : A \to C$, $g : B \to C$ in $\mathcal{D}$, such that $p(f), p(g)$ are deterministic and f, g are Cartesian for the above fibration, p induces a bijection between maps $h : A \to B \in \mathcal{C}$ such that $gh = f$, and maps $h' : p(A) \to p(B)$ so that $p(g)h' = p(f)$. Note that when restricted to those maps where $p(h)$ is deterministic, this being a bijection is the defining property of g being Cartesian (for any f, not necessarily a Cartesian one).

Given a Markov prefibration $\mathcal{D}$, we write $\mathcal{D}|_{\mathrm{det}}$ for $\mathcal{D} \times_{\mathcal{C}} \mathcal{C}_{\mathrm{det}}$. We will refer to this as the *deterministic part* of $\mathcal{D}$—note that this does have the potential for confusion, as when $\mathcal{D}$ is itself a Markov category, this is not necessarily the same as the deterministic subcategory of $\mathcal{D}$. When $f \in \mathcal{D}$ lies inside $\mathcal{D}|_{\mathrm{det}}$, and is Cartesian for that fibration, we will simply refer to it as a Cartesian map in $\mathcal{D}$ (there are no other types of Cartesian maps, so this should not lead to confusion). A *morphism of Markov prefibrations* is a functor $\mathcal{D} \to \mathcal{D}'$ over $\mathcal{C}$ which preserves Cartesian maps. The category of Markov prefibrations over $\mathcal{C}$ thus defined is denoted $\mathsf{MarkPreFib}(\mathcal{C})$. Taking the deterministic part defines a functor $(-)|_{\mathrm{det}} : \mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$.

Remark 3.2.3 (On 2-category theory)**.** Since we will shortly be working with a number of functors between categories whose objects are themselves categories with some structure, it may be thought that we should give some consideration to the strictness of our constructions—for example, we will shortly construct a left adjoint to $(-)|_{\mathrm{det}} : \mathsf{MarkPreFib}(\mathcal{C}_{\mathrm{det}}) \to \mathsf{Fib}(\mathcal{C})$, and it may well be asked how strict this adjoint is, whether we need to consider the definition of *pseudomonad* when we get so far, et cetera.

However, we can largely avoid this issue. The key observation is that none of our functors will alter the *objects* of the underlying category (since $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ is identity on objects,). Hence, all the natural transformations that we would ordinarily ask to be *equivalences* of categories will instead be isomorphisms, and we can largely ignore considerations of higher category theory—similarly, all our functors will be strictly functorial. As a simple example of this, observe that the deterministic part functor $(-)|_{\mathrm{det}}$ is automatically strict—it simply consists in restriction to a subset of the morphisms in $\mathcal{D}$ (which is automatically closed under composition), and thus clearly preserves composition strictly.

Definition 3.2.4. In a Markov prefibration $p : \mathcal{D} \to \mathcal{C}$ (as previously noted), a morphism $f : \bar{X} \to \bar{Y}$ in $\mathcal{D}$ is called *Cartesian* if $p(f)$ is deterministic and f is Cartesian in the fibration $\mathcal{D}|_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$. It is called *vertical* if $p(f)$ is an identity. It is called a *stochastic-Cartesian* if there exists a Cartesian map $r : \bar{Y} \to \bar{X}$ so that $rf = 1_{\bar{X}}$ (recall that in this case f is uniquely determined by r and $p(f)$).

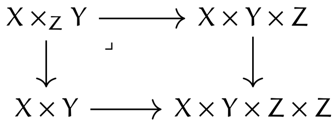
Note that if f is stochastic-Cartesian and $p(f)$ is deterministic, then f is Cartesian.

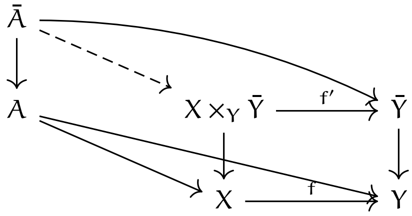
The introduction to this chapter contains the argument that $\mathsf{Kl}(\Delta)$ is a Markov prefibration. This is a key motivating example.

Example 3.2.5 (Markov prefibrations over Cartesian base)**.** Let $\mathcal{C}$ be a Markov category which is Cartesian (that is, one where all morphisms are deterministic). Then a Markov prefibration over $\mathcal{C}$ is simply a Grothendieck fibration.

Proposition 3.2.6. *Let $\mathcal{C}$ be a Markov category. Then $\mathcal{C}^{\to} \to \mathcal{C}$ is a Markov prefibration with Cartesian maps given by pullback squares, if and only if $\mathcal{C}$ is pullback-positive.*

Proof. Assume first $\mathcal{C}$ is pullback-positive. It is clear that computing pullbacks in $\mathcal{C}_{\mathrm{det}}$ gives the required Cartesian lifts—pullback positivity is precisely the claim that lifts exist uniquely in the definition of a Cartesian morphism (since the map to the base leg of the pullback is always deterministic in this case). Since Cartesian lifts can be taken to be deterministic (we have just constructed deterministic Cartesian lifts, and such lifts are unique up to unique isomorphism), the second condition also follows from this assumption, simply taking the other leg to be deterministic.





Conversely, suppose $\mathcal{C}^{\to} \to \mathcal{C}$ is a Markov prefibration and suppose the Cartesian maps are given by the deterministic pullback squares. Let $Y \to Z$ be an object of $\mathcal{C}^{\to}$ and let $X \to Z$ be a deterministic map, and form the pullback $X \times_Z Y$, which is Cartesian. Let $P \to X$ be deterministic and let $P \to Y$ be any map. Expanding the latter into a map from the object $P \to P$ to $Y \to Z$, the Cartesian property of the square implies there is a unique pairing $P \to X \times_Z Y$ □

Remark 3.2.7. In a general Markov category, not every isomorphism is necessarily deterministic. This means that, in general, fibres over isomorphic objects in a Markov prefibration are not necessarily isomorphic or even equivalent as categories. This seemingly immoral situation is, in fact, in accordance with other results indicating that *deterministic isomorphism* is really the proper notion of identification in a Markov category. See eg [FR20, Section 4], for further discussion of this point. (Since the basic idea of a Markov category involves objects equipped with some structure which is not preserved by all the morphisms, it is not so paradoxical that an isomorphism in this situation should be insufficient to render two objects identical).

Under very weak assumptions on the Markov category $\mathcal{C}$, such as positivity, all isomorphisms are deterministic. This implies that all Markov prefibrations over $\mathcal{C}$ are isofibrations, and thus rules out any sort of behavior like the above. As noted, we are only concerned with positive Markov categories.

Relatedly, in the proof of Proposition 3.2.6, a careless prover may have erroneously concluded after the first step that all Cartesian lifts are deterministic—but since Cartesian lifts are characterized only up to *isomorphism*, not necessarily *deterministic* isomorphism, this does not automatically follow. (But of course replacing a nondeterministic lift with an isomorphic deterministic one in this situation cannot alter the unique existence of the factorization, so it does not matter).

Proposition 3.2.8. *Let $p : \mathcal{D} \to \mathcal{C}$ be a Markov prefibration, let $\mathcal{C}' \subseteq \mathcal{C}, \mathcal{D}' \subseteq \mathcal{D}$ be full subcategories so that $p(\mathcal{D}')$ is contained in $\mathcal{C}'$, and suppose $\mathcal{C}'$ is a monoidal subcategory (which is then automatically a sub-Markov category). Suppose finally $\mathcal{D}'$ is stable under pullback along deterministic morphisms in $\mathcal{C}'$. Then $\mathcal{D}' \to \mathcal{C}'$ is again a Markov prefibration.*

Proof. By assumption, given $Y \in \mathcal{D}'$ and $f : X \to p(Y) \in \mathcal{C}'$, the Cartesian lift $X' \to Y$ is again in $\mathcal{D}'$. The fullness of the subcategory inclusions suffices to prove the existence and uniqueness of the required lifts so that this is still Cartesian after restricting. For the same reason, since we have just observed that the Cartesian lifts are the same as in $\mathcal{D} \to \mathcal{C}$, the second part of the definition also holds. □

Proposition 3.2.9. *The codomain functor* $\mathsf{BorelStoch}^{\to} \to \mathsf{BorelStoch}$ *is a Markov prefibration.*

Proof. Standard Borel spaces are known to be stable under products and measurable subsets (see eg. [Sri08] propositions 3.1.23 and 3.3.15). Hence the category of standard Borel spaces and measurable maps admits pullbacks. By d t suffices to verify that the Giry monad on BorelStoch preserves intersections. But clearly a probability measure is supported on $A \cap B$ if and only if it is supported on A and on B, which implies that it does. This concludes the proof. □

Corollary 3.2.10. $\mathsf{FinStoch}^{\to} \to \mathsf{FinStoch}$ *is a Markov prefibration.*

Example 3.2.11. The functor

$$\mathsf{Optic}(\mathsf{BorelStoch}) \to \mathsf{BorelStoch}$$

is not a Markov prefibration, although its pullback over $\mathsf{BorelStoch}_{\mathrm{det}}$ is a Grothendieck fibration.

To see this, first consider the deterministic pullback. An optic $\binom{A}{X} \to \binom{B}{Y}$ with deterministic base can be identified with a map $X \otimes B \to A$ (and the base deterministic map $X \to Y$). To see this, first observe that the subset of $\mathsf{BorelStoch}(X, Y \otimes M)$ with the marginal $X \to Y$ deterministic is in bijection with $\mathsf{BorelStoch}_{\mathrm{det}}(X, Y) \times \mathsf{BorelStoch}(X, M)$, since BorelStoch is positive. Hence we can calculate

$$\int^{M} \mathsf{BorelStoch}_{\mathrm{det}}(X, Y) \times \mathsf{BorelStoch}(X, M) \times \mathsf{BorelStoch}(M \otimes B, A)$$

$$\cong \mathsf{BorelStoch}_{\mathrm{det}}(X, Y) \times \mathsf{BorelStoch}(X \times B, A),$$

using the ninja yoneda lemma as in Proposition 2.3.7.

Hence this part is a fibration with the fiber over X being the coKleisli category of the $X \times -$ monad, and the pullback functors given by reindexing these parametrized maps. The Cartesian lift of a map $X \to Y$ at $\binom{B}{Y}$ is given by the optic $\binom{B}{X} \to \binom{B}{Y}$ with unit residual and identity backwards component.

Now, let $g : I \to \mathbb{R}$ denote the standard Gaussian distribution, let $f : \mathbb{R} \otimes \mathbb{R} \to \mathbb{R}$ be the function given by $f(x, y) = 0$ if $x = y$ and y otherwise, and consider the two optics $\binom{\mathbb{R}}{*} \to \binom{\mathbb{R}}{\mathbb{R}}$ given by $(I, g : I \to \mathbb{R}, 1_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}), (\mathbb{R}, \mathrm{copy}_{\mathbb{R}} g : I \to \mathbb{R} \otimes \mathbb{R}, f : \mathbb{R} \otimes \mathbb{R} \to \mathbb{R})$ (where we recall that the first argument is the residual). They cannot be equal, as postcomposition with the optic $\binom{\mathbb{R}}{\mathbb{R}} \to \binom{*}{*}$ given by the identity $\mathbb{R} \to \mathbb{R}$ yields, for the former, the standard Gaussian $g : I \to \mathbb{R}$, and for the latter, the constant zero map. But postcomposition with the projection $\binom{\mathbb{R}}{\mathbb{R}} \to \binom{\mathbb{R}}{*}$ does give the same optic (the identity), because, for every fixed $y \in \mathbb{R}$, when x is normally distributed, $f(x, y) = y$ with probability one. Hence the unique lifting of Cartesian maps over Cartesian maps cannot hold.

This counterexample indicates that, although $\mathsf{BorelStoch}^{\to}$ is a Markov prefibration, we can not expect a dual version of this prefibration—in fact, since over deterministic maps Optic(BorelStoch) *is* the fiberwise dual of (the restriction to trivially-indexed objects of) $\mathsf{BorelStoch}^{\to}$, this example shows that there is *no Markov prefibration whose deterministic part is the fiberwise dual of* BorelStoch, since this is the only property we used. Moreover, as the example indicates, this is not a mere technical issue, but an unavoidable fact about optics in general measurable spaces—even up to behavioral equivalence, they simply don't satisfy the conditions of being a Markov prefibration. (But see Theorem 3.7.2)

On the other hand, we have:

Proposition 3.2.12. *Let $\mathcal{C}$ be a positive Markov category with supports. Then* $\mathsf{Optic}(\mathcal{C}) \to \mathcal{C}$ *is a Markov prefibration.*

Proof. As above, we see that the deterministic part is a fibration, so take $X \to Y \leftarrow Z$ deterministic maps, and let $\binom{A}{X} \to \binom{A}{Z}$ be an optic so that the induced triangle with the two Cartesian lifts to $\binom{A}{Y}$ commutes. Let the two parts be $f : X \to M \otimes Y, g : M \otimes A \to A$. The implication is that $X \to M \otimes Y \to M$ -almost surely, g is equal to the projection to A (and $X \to M \otimes Y \to Y$ renders the triangle in $\mathcal{C}$ commutative). Note that f factors over the support of this map, hence we can assume the marginal $X \to M$ has full support. Hence up to sliding equivalence, g is strictly equal to the projection. This implies the lift is uniquely determined as desired. □

The existence of supports rules out the pathological behaviour. Essentially, in the presence of supports, we can sensibly reason about "the points of measure zero" and exclude them from consideration—and Proposition 2.4.15 implies that the independent pairing of two measures always has the *least* "points of measure zero", and so that what can be proven equivalent under the assumption of independence will always be equivalent. By contrast, the map $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ from Example 3.2.11 satisfies $f(x, y) = y$ for almost all x when x is normally distributed, for all y, but this does not imply that for all measures on x, y with this marginal, $f(x, y)$ is distributed as the marginal of y.

One point of view is that the map f is simply pathological, and we should restrict our attention to maps that are continuous in some sense (from the point of view of computer science, one argument for this is that *computable* maps are necessarily continuous). The category $\mathsf{TychStoch}$ of Tychonoff spaces and weakly continuous kernels does indeed have supports. However, since it lacks *conditionals,* it is still not ideal from our point of view.

3.3 Stochastic Modules

Because every map in $\mathsf{Kl}(\Delta)^{\to}$ factors into arrows which are "induced" from arrows in $\mathsf{Kl}(\Delta)^{\to}|_{\mathrm{det}}$ and the Markov prefibration property, it may initially be hoped that $\mathsf{Kl}(\Delta)^{\to}$ is in some sense "free" on the data of the fibration $\mathsf{Kl}(\Delta)^{\to}|_{\mathrm{det}} \to \mathsf{Set}$ and the inclusion $\mathsf{Set} \to \mathsf{Kl}(\Delta)$. If that was true, we may further hope that taking the fiberwise opposite of the fibration and applying the same free generation principle would generate a good notion of stochastic lens.

Unfortunately, this is not the case. We will see that the free prefibration is given by gadgets which look a bit like an indexed version of optics, up to a sliding equivalence for *deterministic* maps on the residual. This prompts us to look for some extra structure on the fibration $\mathcal{D}|_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$ which describes sliding equivalences for stochastic maps on the residual. In the next section, we will see that this is exactly the structure of an Eilenberg-Moore algebra for the free prefibration monad on $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$.

In this section, we will give a description of the free Markov prefibration on a fibration $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}}$ (assuming $\mathcal{C}$ is pullback positive). There is a fairly simple description of the hom-sets, but their composition is a bit tricky, and verifying associativity even more so. Hence we will employ a technical trick: by characterizing the hom-sets as "freely generated" in a certain sense from the hom-sets in $\mathcal{D}_0$, we can identify them with sets of natural transformations using a Yoneda-type argument, and infer composition and associativity from there.

Definition 3.3.1 (Indexed copresheaf)**.** Let $p : \mathcal{D} \to \mathcal{C}$ be a functor. An *indexed copresheaf* on $\mathcal{D}$ is a tuple $(X \in \mathcal{C}, F : \mathcal{D} \to \mathsf{Set}, \alpha : F(-) \to \mathcal{C}(X, p(-)))$ consisting of a copresheaf, an object of $\mathcal{C}$, and a natural transformation α as indicated. We say the indexed copresheaf is *over* X, and we will abuse the terminology by referring to F itself as an indexed copresheaf, leaving the transformation α implicit (for example, "let F be an indexed copresheaf over X").

We denote the subset $\alpha^{-1}(\{f\}) \subseteq F(\bar{A})$, for $f : X \to p(\bar{A})$ by $F(\bar{A})_f$.

A map of indexed copresheaves $(X, F, \alpha) \to (Y, G, \beta)$ is a natural transformation $F \to G$ and a map $Y \to X \in \mathcal{C}$ so that the obvious square of natural transformations commutes. We denote the category of indexed copresheaves by $\mathsf{IcoPSh}(\mathcal{C}/\mathcal{D})$. Note that there is an obvious forgetful functor $\mathsf{IcoPSh}(\mathcal{D}/\mathcal{C})^{op} \to \mathcal{C}$

Observe that for each object $A \in \mathcal{D}$, there is a corepresentable copresheaf $(p(A), \mathcal{D}(A, -), p)$. Maps between these obey the Yoneda lemma, in the sense that they are in bijection with maps between the underlying objects in $\mathcal{D}$. This defines a fully faithful functor $\mathcal{D} \to \mathsf{IcoPSh}(\mathcal{D}/\mathcal{C})^{op}$ over $\mathcal{C}$.

Of course, there is a dual notion of indexed presheaf, but this will not interest us.

Proposition 3.3.2. *Let $p : \mathcal{D} \to \mathcal{C}$ be any functor and let $\mathcal{C}_0 \to \mathcal{C}$ be an identity-on-objects functor. Write $\mathcal{D}_0 = \mathcal{D} \times_{\mathcal{C}} \mathcal{C}_0$ for the pullback. If $F : \mathcal{D} \to \mathsf{Set}$ is a copresheaf indexed over $X \in \mathcal{C}$, the pullback $\bar{A} \mapsto F(\bar{A}) \times_{\mathcal{C}(X, p\bar{A})} \mathcal{C}_0(X, p\bar{A})$ is a copresheaf on $\mathcal{D}_0$ indexed over X again in a unique way. This defines a functor* $\mathsf{IcoPSh}(\mathcal{D}/\mathcal{C}) \to \mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C}_0)$*. Moreover, this functor preserves the corepresentable copresheaves (since $\mathcal{C}_0 \to \mathcal{C}$ is identity on objects, so is $\mathcal{D}_0 \to \mathcal{D}$, so this statement makes sense).*

Corollary 3.3.3. *Let $\mathcal{D}_0 \to \mathcal{C}_0$ be a functor and let $\mathcal{C}_0 \to \mathcal{C}$ be identity-on-objects and faithful. Then the pullback of the composite $\mathcal{D}_0 \to \mathcal{C}$ and the inclusion $\mathcal{C}_0 \to \mathcal{C}$ is identical to $\mathcal{D}_0$. Therefore, Proposition 3.3.2 gives a functor* $\mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C}) \to \mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C}_0)$.

The idea of our construction of the free Markov prefibration is to give a certain monad on $\mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C})$ and consider the Kleisli maps between the representable copresheaves.

At this point, the notion of a deterministic map $M \to X$ equipped with a stochastic (ie not *necessarily* deterministic) section begins playing a key role. The phrase "stochastic section" will always carry an implicit "of a deterministic map". In most cases the map that the section is a section *of* will be clear from the context.

Note that, given a stochastic section $Y \to M$ and a deterministic map $X \to Y$, if $\mathcal{C}$ is pullback-positive, there is a unique lifting of this to a section of the projection $Y \times_X M \to Y$. We will use this fact several times.

Definition 3.3.4 (Stochastic Module)**.** Let $\mathcal{C}$ be a Markov category, and let $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}}$ be a fibration. Then a *stochastic module* consists of

1. An indexed copresheaf $F = (A, F, \rho) \in \mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C})$
2. For each $\bar{X} \in \mathcal{D}_0$, Cartesian morphism $\bar{a} : \bar{X}_M \to \bar{X}$ lying over $a : M \to X$, and stochastic section $\alpha : X \to M$, a function $\alpha_* : F(\bar{X}) \to F(\bar{X}_M)$, which acts on the underlying morphisms in $\mathcal{C}$ as composition with α

3. Satisfying, whenever given a commutative square of Cartesian morphisms:

$$\begin{array}{ccc} \bar{Y}_M & \xrightarrow{\bar{f}} & \bar{Y}_N \\ \downarrow{\scriptstyle \bar{a}} & & \downarrow{\scriptstyle \bar{b}} \\ \bar{Y}_X & \xrightarrow{\bar{g}} & \bar{Y} \end{array}$$

   and stochastic sections α, β lying over $a = p(\bar{a}), b = p(\bar{b})$, so that we have a digram in $\mathcal{C}$:

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ {\scriptstyle \alpha}\uparrow\downarrow{\scriptstyle a} & & {\scriptstyle \beta}\uparrow\downarrow{\scriptstyle b} \\ X & \xrightarrow{g} & Y \end{array}$$

   Where the maps except α, β are deterministic, α and β are sections of a and b, and both the square of deterministic maps and the square involving α, β commute, the condition that the square

$$\begin{array}{ccc} F(\bar{Y}_M) & \xrightarrow{F(\bar{f})} & F(\bar{Y}_N) \\ {\scriptstyle \alpha_*}\uparrow & & {\scriptstyle \beta_*}\uparrow \\ F(\bar{Y}_X) & \xrightarrow{F(\bar{g})} & F(\bar{Y}) \end{array}$$

   commutes.

4. And satisfying furthermore, for every two stochastic sections $\alpha : X \to M$, $\beta : M \to N$, the equation $(\beta\alpha)_* = \beta_*\alpha_*$ (note that this makes sense because pullbacks compose).

A morphism of stochastic modules is an indexed natural transformation which preserves the operations α_*. The category of stochastic modules is denoted $\mathsf{SMod}(\mathcal{D}_0/\mathcal{C})$. There is an apparent forgetful functor $\mathsf{SMod}(\mathcal{D}_0/\mathcal{C}) \to \mathsf{IcoPSh}(\mathcal{D}_0/\mathcal{C})$.

Remark 3.3.5. Note that α_* of course depends on α, not just α.

If $f : M \to X$ is a deterministic map, $\alpha : X \to M$ is a stochastic section, and $\bar{X}_M, \bar{X}'_M \to \bar{X}$ are two Cartesian lifts of f to $\bar{X} \in \mathcal{D}_X$, then applying the commutativity axiom for stochastic modules implies that the triangle

$$\begin{array}{ccc} F(\bar{X}) & \xrightarrow{\alpha_*} & F(\bar{X}_M) \\ & \searrow^{\alpha_*} & \downarrow{\scriptstyle \wr} \\ & & F(\bar{X}'_M) \end{array}$$

commutes. In what follows, we will simply write $f^*\bar{X}$ for some choice of cartesian lift, and speak of $\alpha_* : F(\bar{X}) \to F(f^*\bar{X})$. The above shows that this is a harmless abuse—the actions α_* are preserved by the identification of different Cartesian lifts. In particular, we will often make arguments as if pullbacks compose strictly, although in general they only compose up to isomorphism. The above triangle means this is harmless.

The term "stochastic module" is not very good, but this is mostly a nonce definition in any case, so we won't worry too much about it.

Stochastic modules over a given $X \in \mathcal{C}$ can be seen to be monadic over the category of indexed copresheaves over that X. However, the compatibility of these local left adjoints with the structure of the rest of the category is somewhat subtle. But we do have free stochastic modules on *representable* indexed copresheaves, as we will soon see.

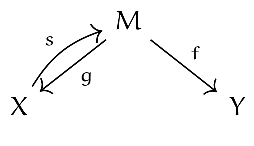
Proposition 3.3.6. *Let $\mathcal{D} \to \mathcal{C}$ be a Markov prefibration, and let $\mathcal{D}_0 = \mathcal{D} \times_{\mathcal{C}} \mathcal{C}_{\mathrm{det}}$. Then the corepresentable copresheaf $\mathcal{D}(\bar{A}, -)$, restricted to $\mathcal{D}_0$, (but not pulled back—that is, we remember the whole set $\mathcal{D}(\bar{A}, \bar{X})$, even the part over stochastic f, but only the* composition *with maps in $\mathcal{D}_0$) is a stochastic module in a canonical way, with $\alpha_* : F(\bar{X}) \to F(f^*\bar{X})$ given by composition with the unique induced lift of α. Moreover, any morphism of Markov prefibrations $\phi : \mathcal{D} \to \mathcal{D}'$ induces a homomorphism of stochastic modules $\mathcal{D}(\bar{A}, -) \to \mathcal{D}'(\phi(\bar{A}), -)$*

Proof. Given $f : M \to X$ and a stochastic section α, it's clear that composition with the unique lift $\bar{X} \to f^*\bar{X}$ is a map of the right type, so we just have to verify the equations. For the first equation (item 3 in the definition of stochastic module), we are comparing two maps $F(g^*\bar{X}) \to F(b^*\bar{X})$. These are given by composition with two maps, let's call them $\phi, \psi : g^*\bar{X} \to b^*\bar{X}$. These two maps are lifts of $f\alpha$ and βg, but by assumption these two are equal. Hence by the uniqueness property of Markov prefibrations, $\phi = \psi$, and we have our equation. The other equation follows in a completely analogous way.

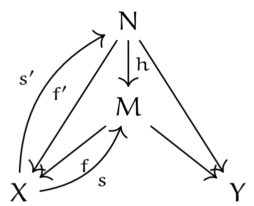
To prove the homomorphism property, note that a morphism of prefibrations preserves Cartesian morphisms, and hence (by uniqueness) must preserve the unique lifts of stochastic sections. Then by functoriality it must preserve composition with these, which finishes the proof. □

Lemma 3.3.7. *1. Let $\bar{X} \in \mathcal{D}_0$ be an object. Then there is a free stochastic module $T\mathcal{D}_0(\bar{X}, -)$ on its representable copresheaf, in the sense that if F is another stochastic module, homomorphisms $T\mathcal{D}_0(\bar{X}, -) \to F$ are in bijection with indexed natural transformations $\mathcal{D}_0(\bar{X}, -) \to F$*

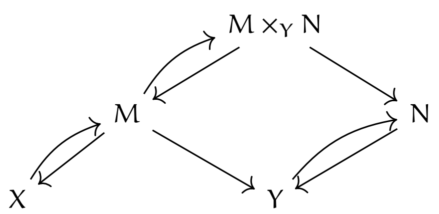
2. The free stochastic module on a corepresentable presheaf is given as follows: an element of $T\mathcal{D}_0(\bar{X}, -)(\bar{Y})$ consists of a diagram in $\mathcal{C}$ of the form



(where $X = p(\bar{X}), Y = p(\bar{Y})$,) where $gs = 1_X$ and f, g are deterministic, plus a map $f^*\bar{X} \to \bar{Y}$ lying over f. This is up to the equivalence relation which, given some other such tuple, identifies them whenever there exists deterministic $h : N \to M$ as in this diagram:

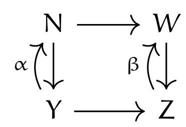


so that the two deterministic triangles commute, $hs' = s$, *and so that the unique Cartesian map* $f'^*\bar{X} \to f^*\bar{X}$ *over* h *forms a commutative triangle with the two maps to* $\bar{Y}$. *(Note that we do not claim the relation just described is inherently an equivalence relation, rather we form the equivalence relation generated by this). It is clear how a map* $\bar{Y} \to \bar{Z}$ *acts on this to make it a copresheaf. It is indexed by taking a tuple as above to the composite* $X \to M \to Y$. *Given* $N \to Y$ *with a stochastic section* $s' : Y \to N$,*, and an element of* $T\mathcal{D}_0(\bar{X}, -)(\bar{Y})$, *the induced element in* $T\mathcal{D}_0(\bar{X}, -)(s'^*\bar{Y})$ *is given by forming the pullback* $M \times_Y N$, *taking the pullback of the map over* g *to one lying over the projection* $M \times_Y N \to N$, *and composing the section* s *with the induced lift* $M \to M \times_Y N$



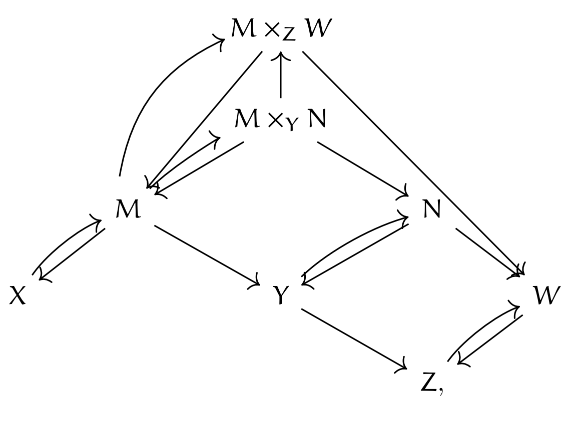
3. *The underlying copresheaf of this respects Cartesian maps in* $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}}$, *in the sense that given a Cartesian map* $\bar{A} \to \bar{B}$ *and an element* $\phi \in T\mathcal{D}_0(\bar{X}, -)(\bar{B})$ *lying over a deterministic map* $X \to B$, *the natural map from lifts* $\psi \in T\mathcal{D}_0(\bar{X}, -)(\bar{A})$ *to lifts* $X \to A$ *is a bijection.*

Proof. First, we have to verify the equations of a stochastic module for $T\mathcal{D}_0(\overline{X}, -)$. For the first case, suppose we are given a square



with all but the upwards maps deterministic. Now let $\bar{Z}$ be some object over Z and suppose we are given an element of $T\mathcal{D}_0(\bar{X}, \bar{Z}_Y)$, represented by a span $X \leftarrow M \to Y$, a section $X \to M$ and a map $\phi : \bar{X}_M \to \bar{Z}_M$ over M.

Consider then the below diagram:

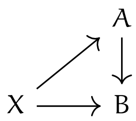


By definition, the first of the two possible elements of $T\mathcal{D}_0(\bar{Z}_W)$ are given by either forming the pullback $M \times_Y N$, taking the lift of α to $M \to M \times_Y N$ and composing to get a section $X \to M \times_Y N$, and pulling back ϕ along the projection to get a map $\bar{X}_{M\times_Y N} \to \bar{Z}_{M\times_Y N}$, then taking the span $X \leftarrow M \times_Y N \to W$. The second is given by first composing with the map $Y \to Z$, then applying the above procedure with the pullback $M \times_Z W$. The induced map $M \times_Y N \to M \times_Z W$ exhibits the equality of these two under the equivalence relation defining $T\mathcal{D}_0(\bar{X}, -)$ (commutativity of the bottom-right square implies that triangle of stochastic sections commutes.)

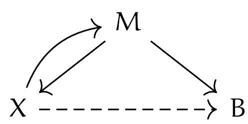
Given some other stochastic module F over A with a map of indexed copresheaves $\mathcal{D}_0(\bar{X}, -) \to F$, over $A \to X$, there is at most one extension to a map of stochastic presheaves $T\mathcal{D}_0(\bar{X}, -) \to F$ over $A \to X$—given an element with representative $(s : X \to M, X \leftarrow M \to Y, \bar{X}_M \to \bar{Y}_M)$, it must go to the identity element of $F(\bar{X})$, acted on by the stochastic section s to produce an element of $F(\bar{X}_M)$, followed by F applied to the map $\bar{X}_M \to \bar{Y}$ over $M \to Y$.

But it is not hard to see that the equivalence relation imposed by $T\mathcal{D}_0(\bar{X}, -)$ is implied by the equations of a stochastic module, and so this map is well-defined, establishing the property.

Secondly, let $\bar{A} \to \bar{B}$ be a Cartesian map over $A \to B$, and take a commutative triangle



of deterministic maps. Finally take an element of $T\mathcal{D}_0(\bar{X}, -)(\bar{B})$ over the given map $X \to B$. We must show it has a unique lift to A over the given map $X \to A$. Let us take a representative given by a diagram:



First, note that the two maps $M \to X \to B$ and $M \to B$ do not necessarily agree. However, we can remedy this by replacing M by their equalizer—note that as we argued above, this gives an equivalent element of the stochastic module. (By writing their equalizer in $\mathcal{C}_{\mathrm{det}}$ as the split pullback $M \times_{B\times B} B$, we can see that the section factors over this, even if $\mathcal{C}$ does not have all equalizers in general). Hence we can assume the triangle formed by adding the dashed arrow commutes.

Since $M \to B$ now factors over X, the lift $X \to A$ gives a lift $M \to A$. Now by the Cartesian property of $\bar{A} \to \bar{B}$, there is a unique lift of $p^*\bar{X} \to \bar{B}$ to this map (here we just use the fact that $\mathcal{D}_0$ is a fibration). This gives the desired lift.

Finally, given two distinct lifts (again, we can assume their maps $M \to A$ factor over X,) clearly any map $N \to M$ witnessing an identity between their composites $\bar{X} \to \bar{Y}$ would likewise exhibit an identity between their lifts (since pullbacks compose). This proves uniqueness, and finishes the proof. □

We are now ready to prove the main proposition of this section:

Proposition 3.3.8. *Let $\mathcal{C}$ be a pullback-positive Markov category and let $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}}$ be a fibration. Consider the full subcategory of stochastic modules spanned by the free modules on the corepresentables. Denote the opposite of this category $\bar{\mathcal{D}_0}$. Clearly there is a commutative diagram*

$$\begin{array}{ccc} \mathcal{D}_0 & \longrightarrow & \bar{\mathcal{D}_0} \\ \downarrow & & \downarrow \\ \mathcal{C}_{\mathrm{det}} & \longrightarrow & \mathcal{C} \end{array}$$

We claim:

1. *$\bar{\mathcal{D}_0} \to \mathcal{C}$ is a Markov prefibration.*
2. *There is a bijection $\bar{\mathcal{D}_0}(\bar{A}, -) \cong \mathsf{T}(\mathcal{D}_0(\bar{A}, -))$. When the left-hand side is equipped with the canonical stochastic module structure coming from Proposition 3.3.6, and the right is equipped with the free one, this is moreover a homomorphism (hence isomorphism) of stochastic modules.*
3. *$\bar{\mathcal{D}_0} \to \mathcal{C}$ is initial among Markov prefibrations receiving a map from $\mathcal{D}_0$. In other words, this construction gives a left adjoint to the pullback functor* $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$

Proof. First observe that, by Lemma 3.3.7, the pullback $\bar{\mathcal{D}_0} \times_{\mathcal{C}} \mathcal{C}_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$ is indeed a fibration, with the image of the Cartesian lifts under the functor $\mathcal{D}_0 \to \bar{\mathcal{D}_0}$ being Cartesian again. (This also establishes that $\bar{\mathcal{D}_0}$ really does receive a map of fibrations from $\mathcal{D}_0$)

Given Cartesian $f : \bar{A} \to \bar{B} \leftarrow \bar{C}$, and a stochastic lift $A \to C (= p\bar{A} \to p\bar{C})$, consider the pullback $A \times_B C$, and the pullback of $\bar{A}$ to it. There is a unique lift of $A \to C$ to a section $A \to A \times_B C$, and this induces a unique lift $\bar{A} \to (f\pi_1)^*\bar{A}$ using the stochastic module structure. The composite of this with the projection to B is a lift of $A \bar{\to} \bar{B}$ over $A \to C$, as required by a Markov prefibration.

Analogously to the proof of Lemma 3.3.7, given some other lift $A \leftarrow N \to C, h : A \to N, h^*\bar{A} \to \bar{C}$, the fact this is a factorization implies the existence of some M with maps $M \to A, M \to N$ and a lift $A \to M$ of the section $A \to N$, so that the induced map between the pullbacks over M and N of $\bar{A}$ makes the triangle into $\bar{B}$ commute. But then since this is a triangle over *deterministic* bases, this implies the lifted triangle to $\bar{C}$ also commutes, hence this M lifts to another representative of the lift we started with. But then it's not hard to see that this M maps to $A \times_B C$ and exhibits an equation with the previously constructed "canonical" lift.

Hence $\bar{\mathcal{D}_0}$ is a Markov prefibration, and by the above, the induced stochastic module structure on the corepresentable presheaves $\bar{\mathcal{D}_0}(\bar{A}, -) = \mathsf{T}(\mathcal{D}_0(\bar{A}, -))$ is exactly the one given by T (in other words this equation is not merely a bijection of sets, but an isomorphism of stochastic modules).

Let $\phi : \mathcal{D}_0 \to \mathcal{D}'$ be a functor over $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ into some other Markov prefibration which preserves Cartesian maps. Using the algebra structure on $\mathcal{D}'(\phi\bar{B}, -)$, we see there is a unique extension of ϕ to $\bar{\mathcal{D}_0}(\bar{X}, -)$ which respects the stochastic module structure. By chasing the diagram around it's easy to see that this is functorial, and hence gives a map of Markov prefibrations—conversely, any such map extending ϕ must be a stochastic module homomorphism. Thus there is a unique functor, proving initiality. ◻

This characterization of the left adjoint makes it fairly easy to understand the induced monad on $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$.

Remark 3.3.9. We will sometimes refer to the morphisms of $\overline{\mathcal{D}_0}$ as *precharts*. Taking the fibration $\mathcal{C}^{\rightarrow}|_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$ as an example, it is not too hard to see that the precharts between $X \otimes A \to X$ and $B \otimes Y \to Y$ are representatives of co-optics $\binom{A}{X} \rightrightarrows \binom{B}{Y}$. (To see this, note that any prechart is equivalent to one where the apex of the span has the form $M \otimes Y$ and the right leg is the projection to Y. Then the rest of the data is a map $X \to M \otimes Y$ and a map $M \otimes B \to A$, since the X-coordinate of the latter map is determined by the span).

In fact their equivalence relation is given by sliding equivalence for *deterministic* maps (i.e morphisms in $\mathsf{Optic}_{\mathcal{C}_{\mathrm{det}}}(\mathcal{C}_{\mathrm{det}}, \mathcal{C})$). The precharts in $\mathcal{D}_0^{\mathrm{fop}}$ will be called *prelenses*. We will speak of the tuple

$$(M, p : M \to X, p' : M \to Y, s : X \to M, \phi : p^*\bar{X} \to p^*\bar{Y})$$

representing a prechart just as a "decorated span (representing ...)". When part of the structure is understood, or can just be left abstracted, we will denote such a decorated span simply by (M, s, ϕ), or even just (M, ϕ). It will be clear from context which part of the structure is being specified.

Proposition 3.3.10. *Let $\mathcal{D}_0$ be a fibration. Then the underlying fibration of the free Markov prefibration, $\overline{\mathcal{D}_0}|_{\mathrm{det}}$, has fiber over $X \in \mathcal{C}$ given by*

1. *Objects are simply objects of $\mathcal{D}_{0,X}$.*
2. *A morphism $\bar{X} \to \bar{X}'$ consists of a deterministic $f : M \to X$, a stochastic section $s : X \to M$, and a map $\phi : f^*\bar{X} \to f^*\bar{X}'$, up to the equivalence relation generated by, whenever $g : N \to M$ is deterministic and $s' : X \to N$ is a factorization of s, identifying (M, f, s, ϕ) with $(N, fg, s', g^*(\phi))$.*
3. *Given two such morphisms $(M, f, s, \phi), (N, f', s', \psi)$, their composite is represented by $M \times_X N \to X$ equipped with the section formed as the composite of $X \to M$ and the lift of $X \to N$ to the pullback, and the composite $\pi_M^*(\phi)\pi_N^*(\psi) \in \mathcal{D}_{0,M\times_X N}$.*
4. *Given deterministic $f : X \to Y$, the pullback is given on such a map by taking the pullback $M \times_Y X \to X$, the induced section, and the pullback of the map ϕ along the projection $M \times_Y X \to M$.*

It is immediately obvious that we have:

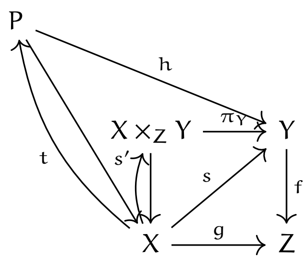
Corollary 3.3.11. *The free Markov prefibration monad $\overline{(-)}|_{\mathrm{det}}$ commutes with fiberwise opposites. In particular, algebra structures on $\mathcal{D}_0$ are in bijection with algebra structures on $\mathcal{D}_0^{\mathrm{fop}}$, and fiberwise opposites lifts to an involution of $\mathsf{Alg}(\overline{(-)}|_{\mathrm{det}})$.*

Definition 3.3.12 (Stochastic Module Fibration)**.** Let $\mathcal{C}$ be a pullback-positive Markov category. The adjunction $\overline{(-)} \dashv (-)|_{\mathrm{det}}$ induces a monad on $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$. A module for this monad is called a *stochastic module over* $\mathcal{C}$ (or, to distinguish it from the copresheaves of Definition 3.3.4, a stochastic module fibration). The category of stochastic module fibrations is denoted $\mathsf{SFib}(\mathcal{C})$

Let us try to understand the structure of a stochastic module fibration. It is easiest to understand in the case of a projection map $P \times X \to X$. Suppose we have two objects A, B over X. Then we think of a map $f : \pi_X^* A \to \pi_X^* B$ over $P \times X$ as a map $P \times A \to B$, that is a map parameterized by P (this is literally the case for a codomain fibration). Given a stochastic section s of π_X, which amounts to a stochastic map $X \to P$, the stochastic module stucture picks out a new map $s_* f$, which is given over each point $x \in X$ by choosing the parameter according to s, then applying f.

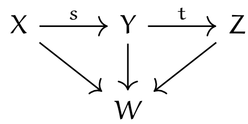
The definition of the composite in Proposition 3.3.10, in these terms, tells us that given maps $f : P \times A \to B, g : Q \times B \to C$, and maps $s : X \to P, t : X \to Q$, the composite of $s_*(f)$ and $t_*(q)$ is equal to the map obtained by forming the parameterized composite $Q \times P \times A \to C$ and applying the independent pairing $\langle t, s\rangle : X \to Q \times P$. This is of course how composition is supposed to work in a Markov category.

Lemma 3.3.13. *Let $\mathcal{D}$ be a stochastic module over $\mathcal{C}$, and let*



be given, so that every map except s, s', t is deterministic. Suppose the deterministic part of the diagram commutes, $fs = g$, s' is the induced section, and t is a section. Let $\bar{A}, \bar{B}$ be two objects over Z. Suppose given a map $\phi : f^*\bar{A} \to f^*\bar{B}$. Then $t^*h^*(P) = (s')^*\pi_Y^*(P) : g^*\bar{A} \to g^*\bar{B}$.

In particular, this operation depends only on s. Moreover, it is functorial, in the sense that given a diagram



with the downwards maps deterministic, $s^*t^* = (ts)^*$.

Remark 3.3.14. The operations α^* associated to stochastic lifts are "functorial" in the sense that $(\alpha\beta)^* = \beta^*\alpha^*$. However they are *not* functorial in the sense that $\alpha^*(fg) = \alpha^*(f)\alpha^*(g)$.

To make sense of this, consider a simple case of a map $m : I \to X$ in $\mathrm{Kl}(\Delta)$. Given two objects over $*$ (in $\mathrm{Kl}(\Delta)^{\to}$), a map $\bar{A}_X \to \bar{B}_X$ is equivalent to a parametrized map $X \times \bar{A} \to \bar{B}$. The operation

$$m^* : (A \times X \to B) \to (A \to B)$$

consists in sampling this parameter according to the distribution m—but since composition in the fiber over X is defined by *copying* the parameter, but composition in the fiber over $*$ (i.e just $\mathrm{Kl}(\Delta)$) is defined by composing the kernels under conditional independence, these only agree if the distribution m is assumed to be deterministic.

Note that if either f or g is pulled back from a map $\bar{A} \to \bar{B}$ (i.e, if they do not depend on the parameter X), the composition *is* preserved.

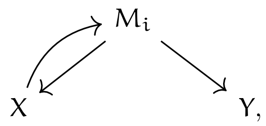
By construction, two morphisms in $\overline{\mathcal{D}}_0$ represented by spans with apex M, M', are identified if there exists a zig-zag $M \to K_0 \leftarrow K_1 \to \cdots \leftarrow M'$ of spans (decorated with sections from the domain X and morphisms in the fiber, satisfying equations, etc). We will now prove a lemma that allows us to cut this down to a smaller set in many conditions. We will need the following hypothesis:

Definition 3.3.15 (Weak conditionals)**.** Let $p : \mathcal{D} \to \mathcal{C}$ be a Markov prefibration. We say p (or, abusing notation, $\mathcal{D}$) *admits weak conditionals* if, given a Cartesian map $\bar{Y} \to \bar{Z}$ and any map $\bar{X} \to \bar{Z}$, the existence part of the Cartesian condition holds—that is, every factorization $p(\bar{X}) \to p(\bar{Y})$ admits a lift, although not necessarily a unique one.

We say a Markov category $\mathcal{C}$ admits weak conditionals if its codomain functor $\mathcal{C}^{\to} \to \mathcal{C}$ is a prefibration which admits weak conditionals—this is equivalent to requiring that it is pullback-positive, and that *all* deterministic pullbacks are carried to weak pullbacks by the inclusion $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ (in other words, that they satisfy the existence part of the universal property even for pairs of nondeterministic maps).

Observe that, if $\mathcal{C}$ admits conditionals, it certainly admits weak conditionals: given a pullback $X \times_Z Y$, and maps $P \to X, Y$, form a Bayesian inverse of $Y \to Z$ with respect to the given measure, and use that to build a lifting $X \to Y$, which gives $X \to X \times_Z Y$—then a diagram chase verifies that this map has the desired properties.

Lemma 3.3.16. *Suppose $\mathcal{C}$ admits weak conditionals, and let $\mathcal{D} \to \mathcal{C}_{\mathrm{det}}$ be a fibration. Then two morphisms $f_0, f_1 : \bar{X} \to \bar{Y}$ in $\overline{\mathcal{D}}$, represented by commutative diagrams*



as well as $\phi_i : \bar{X}_{M_i} \to \bar{Y}_{M_i}$, for $i = 0, 1$, are equal if and only if there exists a span $M_0 \leftarrow K \to M_1$ over X, Y, with a stochastic section $X \to K$ lifting both the sections to M_0, M_1, so that the pullbacks of ϕ_0, ϕ_1 to K agree.

Proof. The relation here described clearly implies identity, and contains all the generating identities, so it suffices to show it is an equivalence relation. Reflexivity and symmetry are clear, so transitivity is the only issue. It suffices to show that, given a span $X \leftarrow S \to Y$ and maps $M_0 \to S \leftarrow M_1$ so that the triangles commute, and so that the two induced sections $X \to S$ agree, and a map $\bar{X}_S \to \bar{Y}_S$ which pulls back to ϕ_0, ϕ_1, we can find K as above.

To do this, take $K = M_0 \times_S M_1$. Clearly the maps to M_0, M_1 are over X, Y, and by the existence of weak conditionals there exists a common lift of the sections to $X \to K$. By functoriality of pullbacks, the pullbacks of ϕ_0, ϕ_1 to K agree. □

3.4 Markov Fibrations

Recall that, given a functor $R : \mathcal{C} \to \mathcal{D}$ with left adjoint L, there is a "standard resolution" of any object $X \in \mathcal{C}$, given by the "cofork" $LRLRX \rightrightarrows LRX \to X$, where the two parallel maps are the two possible applications of the adjunction counit. The adjunction is monadic if and only if this is always a coequalizer, in which case the RL-algebra corresponding to X is $RLRX \to RX$—conversely, given an algebra $\alpha : RLA \to A$, there are two parallel maps $LRLA \rightrightarrows LA$ (given by $L(\alpha)$ and the counit), and the object in $\mathcal{C}$ corresponding to this algebra is given by this coequalizer.

Example 3.4.1. Consider the adjunction $|-| : \mathsf{Mon} \leftrightarrows \mathsf{Set} : (-)^*$ between the category of monoids and the category of sets. Given a monoid M, $|M|^*$ consists of lists of elements in M, and $||M|^*|^*$ consists of lists of such lists. The two maps $||M|^*|^* \to |M|^*$ consist in either concatenating the lists, or replacing each list with its product. Clearly these two maps are coequalized by the product map $|M|^* \to M$. Moreover it's clear that two lists have the same product if and only if they are identified in this coequalizer (simply consider a singleton list-of-lists, which identifies any given list with the singleton corresponding to its product).

As a generalization of this, if this coequalizer exists for every algebra, they form a left adjoint to the canonical functor $\mathcal{C} \to \mathsf{Alg}_{\mathcal{D}}(RL)$. Since we have seen that the monad of free Markov prefibrations commutes with taking fiberwise opposites, we may hope that such a left adjoint exists—a simple argument shows that, if it does, it is fully faithful, and we may say that those prefibrations in the image are the "fibrations" and define their fiberwise opposite as the fiberwise opposite applied to their underlying algebras. Although it turns out to not be quite so simple, we will take this idea as our starting point.

The following example illustrates this construction for a simpler adjunction:

Example 3.4.2. Let GrpTop be the category of topological groups, and let $R : \mathsf{GrpTop} \to \mathsf{Set}$ forget both the group structure and the topology. Clearly this is right adjoint to the free group in the discrete topology, and the monad of this adjunction is the free group monad, which we write RL for now. The canonical comparison functor $\mathsf{GrpTop} \to \mathsf{Grp}$ just forgets the topology. Given a (non-topological) group, described by a map $RLG \to G$, we can form the diagram of topological groups $LRLG \rightrightarrows LG$. Here LG is the free, discrete group on G and $LRLG$ is the free discrete group on the underlying set of LG. Their coequalizer is simply G equipped with the discrete topology, which is indeed the left adjoint to $\mathsf{GrpTop} \to \mathsf{Grp}$

In what follows, we will denote the monad $\overline{(-)}|_{\det}$ simply by Free to avoid too many complicated nestings of overlines and parentheses.

In this section, we will study the coequalizers of $\overline{\mathrm{Free}(\mathcal{D}_0)} \rightrightarrows \overline{\mathcal{D}_0}$ when $\mathcal{D}_0$ is a stochastic module. This corresponds to freely adding certain lifts to extend $\mathcal{D}_0$ into a prefibration (the free prefibration we already constructed), then identifying the formal operations with their actual values in the stochastic module. This yields a category over $\mathcal{C}$ which we denote $\mathsf{SChart}(\mathcal{D}_0)$. When applied to the fiberwise opposite of $\mathcal{D}_0$, this will be our notion of "stochastic lenses"

Definition 3.4.3 (Fibred supports)**.** Let $\mathcal{D}_0 \to \mathcal{C}_{\det}$ be a stochastic module fibration.

Let $\alpha : X \to M$ be a stochastic section, let $\bar{X} \in \mathcal{D}_{0,X}$ be an object, and let $\phi : \bar{X}_M \to \bar{X}_M \in \mathcal{D}_{0,M}$ be an endomorphism of its pullback. Observe that if there exists $f : N \to M$ so that $f^*(\phi) = 1$ and α factors over f, then $\alpha_*(\phi) = 1$. We say $\mathcal{D}_0$ has *fibred supports* if this implication is an equivalence.

We will make use of the following lemma:

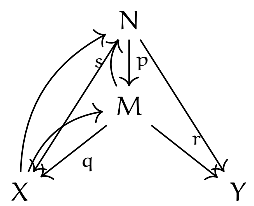
Lemma 3.4.4. *Let* $F, G : \mathcal{D} \rightrightarrows \mathcal{D}'$ *be a parallel pair in* $\mathsf{Cat}_{/\mathcal{C}}$*, and suppose both are identity on objects. Suppose moreover this is a reflexive pair, i.e there is some* $S : \mathcal{D}' \to \mathcal{D}$ *so that* $FS = GS = 1_{\mathcal{D}'}$*. Then the coequalizer in* $\mathsf{Cat}_{/\mathcal{C}}$ *is again identity on objects, and is given on hom-sets simply by the coequalizer of the parallel pair* $\mathcal{D}(x, y) \rightrightarrows \mathcal{D}'(x, y)$.

Proof. The only nontrivial part is to verify that composition is well-defined on the equivalence classes in $\mathcal{D}'(x,y)/\sim$. It suffices to see that post- and precomposition with a fixed morphism both preserve this equivalence relation. Take some $f : x \to y \in \mathcal{C}$, and $h : y \to z \in \mathcal{C}'$. We must show that $hF(f) = hG(f)$. But simply write

$$hF(f) = F(S(h)f) \sim G(S(h)f) = hG(f),$$

and we are done. Clearly the other side follows by duality, finishing the proof. □

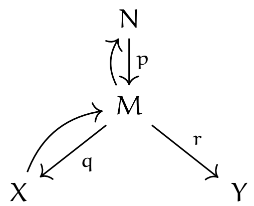
Proposition 3.4.5 (Construction of $\mathsf{SChart}(\mathcal{D}_0)$)**.** *Let $\mathcal{D}_0$ be a fibration equipped with a stochastic module structure. Consider the equivalence relation on $\overline{\mathcal{D}_0}(\bar{X},\bar{Y})$ which identifies two precharts $(M,\phi),(N,\phi')$ if there exists a map $f : N \to M$ over X, Y and a stochastic section s of f which preserves the section from X, so that $s^*\phi' = \phi$ (note that this makes sense because pullbacks compose).*



Then:

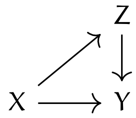
1. *This equivalence relation respects composition, and so defines a category which we denote $\mathsf{SChart}(\mathcal{D}_0)$*
2. *In $\mathsf{Cat}_{/\mathcal{C}}$, $\overline{\mathrm{Free}(\mathcal{D}_0)} \rightrightarrows \overline{\mathcal{D}_0} \to \mathsf{SChart}(\mathcal{D}_0)$ is a coequalizer diagram. In particular, $\mathcal{D}_0$ presents a Markov fibration if and only if $\mathsf{SChart}(\mathcal{D}_0)$ is a Markov prefibration (in which case $\mathsf{SChart}(\mathcal{D}_0)$ is the fibration it presents)*
3. *If $\mathcal{D}_0$ has fibred supports, $\mathsf{SChart}(\mathcal{D}_0)$ is again a prefibration*

Proof. Recall that the fibration $\mathrm{Free}(\mathcal{D}_0)$ has fibers whose morphisms $\bar{X} \to \bar{X}'$ are given by tuples $s : X \to M : p, p^*\bar{X} \to p^*\bar{X}'$ (up to a certain equivalence relation). Forming the free Markov prefibration $\overline{\mathrm{Free}(\mathcal{D}_0)}$ on this fibration, we find that the morphisms are given by diagrams



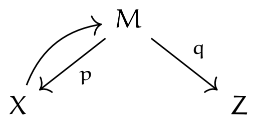
equipped with a map $\phi : p^*q^*\bar{X} \to p^*r^*\bar{Y}$. The two maps $\overline{\mathrm{Free}(\mathcal{D}_0)} \to \overline{\mathcal{D}_0}$ carry such a thing to first, the map resulting from forgetting M and just composing q, r with p to get a span (and composing the sections to get a new section), and secondly, the map with apex M obtained by using the stochastic module structure to push ϕ down into a map over M. It is clear that this is equivalently the equation described in the theorem. This establishes points 1. and 2., since by Lemma 3.4.4 we can compute such coequalizers hom-set by hom-set.

Now we wish to prove that $\mathsf{SChart}(\mathcal{D}_0)$ is a Markov prefibration given fibred supports. Since by the coequalizer presentation, its deterministic part is isomorphic to $\mathcal{D}_0$, the fibration property is automatic. It remains to verify that, given a triangle

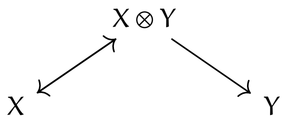


in $\mathcal{C}$ with the vertical and horizontal maps deterministic, an object $\bar{A} \in \mathcal{D}_{0,Y}$ and Cartesian maps $\bar{A}_X \to \bar{A} \leftarrow \bar{A}_Z$, there exists a unique lift $\bar{A}_X \to \bar{A}_Z$ in $\mathsf{SChart}(\mathcal{D}_0)$.

Such a lift is given by a diagram



equipped with $p^*\bar{A}_X \to q^*\bar{A}_Z$. By taking the equalizer of the two maps $M \to Y$ (the section factors over this), we may assume these two are equal, which implies that the pulled-back objects are equal—denote this object $\bar{A}_M$. Now the hypothesis is that after postcomposing with the Cartesian map $\bar{A}_Z \to \bar{A}$, this gives the map $\bar{A}_X \to \bar{A}$. This postcomposition is given simply by postcomposing the leg $M \to Z$ with the map $Z \to Y$ (and observing that, by functoriality of pullbacks, this does not alter the pulled-back object). Then the claim is that integrating this map $\bar{A}_M \to \bar{A}_M$ down into a map $\bar{A}_X \to \bar{A}_X$, it gives the identity. But by assumption this means we can pull back to some object $M' \to M$ (lifting the section from X) where the two maps are already equal to the identity. But this pull-back can be applied to the original map $\bar{A}_X \to \bar{A}_Z$ as well. But this implies every such map is equal to the one represented by the diagram



and the identity on $\bar{A}_{X\otimes Y}$, with the map $M' \to X \otimes Y$ giving the witness, since identities pull back. This map only depends on the underlying $X \to Y$, hence $\mathsf{SChart}(\mathcal{D}_0)$ is indeed a prefibration. □

Definition 3.4.6 (Stochastic charts and lenses)**.** We refer to $\mathsf{SChart}(\mathcal{D}_0)$ as the category of *stochastic charts* in $\mathcal{D}_0$. We refer to $\mathsf{SChart}(\mathcal{D}_0^{\mathrm{fop}})$ as *stochastic lenses* and denote it also $\mathsf{SLens}(\mathcal{D}_0)$

Lemma 3.4.7. *Let* (M, ϕ) *and* (N, ψ) *be two representatives of charts. Given some possibly stochastic map* $f : M \to N$ *over* X *and* Y*, recall (Lemma 3.3.13) that we can define* $f^*\psi$*, regardless of whether* f *is deterministic or the section of a deterministic map. If there exists such a map* f*, we can always factor it over the pullback* $M \times_{X\times Y} N$ *as a section followed by a deterministic map. Hence the equivalence relation defining*

SChart *is equivalent to the relation identifying two representatives whenever there exists such an* f *with* $f^*\psi = \phi$

Definition 3.4.8 (Markov Fibration)**.** Let $\mathcal{D}$ be a Markov prefibration. If the canonical map $\mathsf{SChart}(\mathcal{D}|_{\mathrm{det}}) \to \mathcal{D}$ is an isomorphism, we call $\mathcal{D}$ a *Markov fibration*. If $\mathcal{D}_0$ is a stochastic module so that $\mathsf{SChart}(\mathcal{D}_0)$ is a prefibration, we say $\mathcal{D}_0$ *presents a Markov fibration*.

The terminology "presents a Markov fibration" is justified by the following proposition.

Proposition 3.4.9. *Let* $\alpha : \mathrm{Free}(\mathcal{D}_0) \to \mathcal{D}_0$ *be stochastic module which presents a Markov fibration. Then the underlying stochastic module of* $\mathsf{SChart}(\mathcal{D}_0)$ *is isomorphic to* $\mathcal{D}_0$*, and in particular this prefibration* is *a Markov fibration.*

This correspondence determines an equivalence of categories between the full subcategory $\mathsf{MarkFib}(\mathcal{C})$ *of* $\mathsf{MarkPreFib}(\mathcal{C})$ *spanned by the Markov fibrations, and the full subcategory* $\mathsf{SFib}(\mathcal{C})^{\mathrm{p}} \subseteq \mathsf{SFib}(\mathcal{C})$ *spanned by those algebras which present a Markov fibration.*

Proof. Let α be an algebra as assumed, and let $\mathcal{D}_0^\alpha \in \mathsf{MarkPreFib}$ denote the coequalizer given. By general nonsense there is an induced functor $\mathcal{D}_0 \to \mathcal{D}_0^\alpha|_{\mathrm{det}}$ which is moreover an algebra homomorphism—the claim is that this is an isomorphism. By Lemma 3.4.4, pullback to the deterministic part preserves these coequalizers, so this amounts to the claim that the diagram $\mathrm{Free}^2(\mathcal{D}_0) \rightrightarrows \mathrm{Free}(\mathcal{D}_0) \to \mathcal{D}_0$ is a coequalizer. But this is true for any algebra of any monad (in fact, the unit gives a splitting of this coequalizer).

By general nonsense the fibration associated to an algebra which presents a fibration forms a partial left adjoint to $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{SFib}(\mathcal{C})$. This left adjoint, by the above, has its image inside $\mathsf{MarkFib}$, and hence there is an adjunction $\mathsf{MarkFib}(\mathcal{C}) \leftrightarrows \mathsf{SFib}(\mathcal{C})^{\mathrm{p}}$. The preceding furthermore proves that the unit of this adjunction is the identity, which implies that the left adjoint is fully faithful—but by definition it is essentially surjective, finishing the argument. □

Let us briefly summarize the relationship between Markov prefibrations, Markov fibrations, and stochastic module fibrations at this stage.

- A stochastic module fibration is a (Grothendieck) fibration $\mathcal{D}$ over $\mathcal{C}_{\mathrm{det}}$, equipped with some extra structure involving the whole category $\mathcal{C}$. Given a deterministic map $f : A \to B$ two objects $X, Y \in \mathcal{D}_B$, and a map $\phi : f^*X \to f^*Y$, we can think of this as a map parameterized by the fibers A_b. Given a stochastic section $s : B \to A$, the stochastic module structure picks out a map $X \to Y \in \mathcal{D}_B$ corresponding to choosing this parameter randomly according to s.

- A Markov prefibration is a category $\mathcal{D}$ over $\mathcal{C}$ with a particular unique lifting property. In the above situation, it gives a unique lift $X \to f^*X$ of s, corresponding to choosing $a \in A_b$ according to s and leaving the $x \in X$-coordinate unchanged. By composing this lift with ϕ, then with the Cartesian $f^*Y \to Y$, we get a stochastic module structure on the part of $\mathcal{D}$ lying over deterministic maps (which is also a Grothendieck fibration).

- Given a stochastic module structure, there is a way of generating a category over $\mathcal{C}$, by freely adding the lifts corresponding to a Markov prefibration, then quotienting by the relations implied by the stochastic module structure. This does not necessarily yield a Markov prefibration.

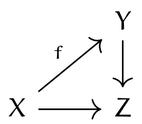
- A Markov prefibration is called a Markov fibration if it is presented by its underlying stochastic module in the above sense.

It is not apparent whether fibred supports are necessary for $\mathsf{SChart}(\mathcal{D}_0)$ to be a prefibration. We have not found any counterexample, but in general the equivalence relation on charts is fairly complicated, so it is not apparent how to prove the necessity. We will generally not be too bothered about assuming fibred supports instead of the more nebulous assumption that $\mathcal{D}_0$ presents a Markov fibration.

Remark 3.4.10. By construction, given a morphism of stochastic modules $F : \mathcal{D}_0 \to \mathcal{D}'_0$, there is an induced functor $\mathsf{SChart}(\mathcal{D}) \to \mathsf{SChart}(\mathcal{D}')$ over $\mathcal{C}$. This restricts to F on the deterministic part and in particular preserves Cartesian morphisms.

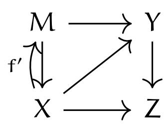
Proposition 3.4.11. *Let $\mathcal{D}_0$ be a stochastic module fibration. Then $\mathcal{D}_0$ presents a Markov fibration if and only if $\mathcal{D}_0^{\mathrm{fop}}$ does.*

Proof. Consider a triangle of this form in $\mathcal{C}$:



where the maps to Z are deterministic. Suppose given Cartesian lifts $\bar{X} \to \bar{Z}, \bar{Y} \to \bar{Z}$ of the cospan. These are the same in both cases, coming from Cartesian maps in $\mathcal{D}_0$ or $\mathcal{D}_0^{\mathrm{fop}}$ (which are the same). We must show that there is a unique lift of f to a chart in $\mathsf{SChart}(\mathcal{D}_0)$ if and only if there is a unique lift to a chart in $\mathsf{SChart}(\mathcal{D}_0^{\mathrm{fop}})$. Clearly it suffices to prove the "only if" implication, so suppose $\mathsf{SChart}(\mathcal{D}_0)$ is a prefibration.

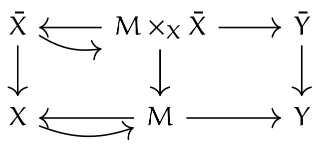
By passing to the equalizer as in the proof that $\overline{\mathcal{D}_0}$ is a prefibration, we may assume that any such lift is represented by a diagram



where the outer square and the triangle $X \to M \to Y$ commutes, and f' is a section.

Take such a diagram and let $\phi : \bar{Z}_M \to \bar{Z}_M$ be the map representing a chart. Then the claim is there exists some zig-zag of chart equivalences identifying (M, ϕ) with $(M, 1)$. But clearly this is invariant under passing to the fiberwise opposite, and so $\mathcal{D}^{\mathrm{fop}}$ is also a prefibration. □

Example 3.4.12. $\mathsf{Kl}(\Delta)^{\to}$ is a Markov fibration. We have already seen that it is a Markov prefibration, and that the map from the coreflection is full. So it suffices to prove faithfulness. Consider a map in $\mathsf{Kl}(\Delta)^{\to}|_{\mathrm{det}}$, given by a diagram



We can factor the section $X \to M$ as $X \to X \times Y \to M$, where the first map is just the pairing and the second is a conditional distribution. This induces a factorization of the lift $\bar{X} \to M \times_X \bar{X}$ over $\bar{X} \to \bar{X} \times Y$. By composing the map $M \times_X \bar{X} \to \bar{Y}$ with this factorization to build the map $\bar{X} \times Y \to \bar{Y}$, we have found a new representative for the same map.

Hence every map over $X \to Y$ has a representative where the residual is $X \times Y$. We would like to argue that, since the map $\bar{X} \times Y \to \bar{Y}$ is given by the conditional distribution of the composite map $\bar{X} \to \bar{Y}$, it is uniquely determined by it, and thus if two distinct maps in $\overline{\mathsf{Kl}(\Delta)^{\to}|_{\det}}$ have the same underlying map in $\mathsf{Kl}(\Delta)^{\to}$, they must have equal representatives of this form, and so be identified in the coreflection (which must therefore be isomorphic to $\mathsf{Kl}(\Delta)^{\to}$). But of course, the two maps may only be *almost certainly* equal.

In this case, there is a simple fix: instead of taking $X \times Y$ as the residual, take the subset S given by those pairs (x, y) where y has positive probability given x. The pairing factors over this, of course, and two maps $\bar{X} \times_X S \to \bar{Y}$ which give the same distribution $\bar{X} \to \bar{Y}$ really must have the same value on every point. This proves that $\mathsf{SChart}(\mathsf{Kl}(\Delta)^{\to}) \to (\mathsf{Kl}(\Delta))^{\to}$ is faithful and hence an isomorphism.

Example 3.4.13. $\mathsf{BorelStoch}^{\to}$, as we have noted, is a Markov prefibration, and hence induces a stochastic module structure on $\mathsf{BorelStoch}^{\to}|_{\det}$. This structure does not present a Markov fibration. To see this, note that in that case its fiberwise opposite would also present a Markov fibration. Then this fibration, $\mathsf{SLens}(\mathsf{BorelStoch}^{\to}|_{\det})$, would be a prefibration whose deterministic part was $\mathsf{BorelStoch}^{\to}|_{\det}^{\mathrm{fop}}$. But Example 3.2.11 shows that this is impossible.

Let us say that a Markov category has *enough points* if *deterministic* morphisms from the unit separate morphisms.

Proposition 3.4.14. *Let $\mathcal{C}$ be a Markov category with supports and enough points. Then the stochastic module induced by $\mathcal{C}^{\to} \to \mathcal{C}$ presents a Markov fibration. If $\mathcal{C}$ has conditionals, this Markov fibration is isomorphic to $\mathcal{C}^{\to}$.*

Proof. We will verify first that having enough points and supports implies the stochastic module has fibred supports. Given a section $s : X \to M : p$ and $\phi : A \times_X M \to A \times_X M$ so that $s^*(\phi) = 1_A$. Now consider the pullback $A \times_X S$ along the inclusion $S \hookrightarrow M$, S being the support of s. If the pulled back map $A \times_X S \to A$ is not equal to the projection, there must be some point $a : I \to A, \sigma : I \to S$ which witnesses this difference. But this point then proves that ϕ is not s-almost surely equal to the identity (since σ is in the support), which is equivalent to $s^*(\phi)$ not being equal to the identity by positivity.

Hence by Proposition 3.4.5, $\mathcal{C}^{\to}$ presents a Markov fibration.

There is an induced map $\mathsf{SChart}(\mathcal{C}^{\to}|_{\det}) \to \mathcal{C}^{\to}$, which we claim is an isomorphism. So consider a map in $\mathcal{C}^{\to}$:

$$\begin{array}{ccc} \bar{X} & \longrightarrow & \bar{Y} \\ \downarrow & & \downarrow \\ X & \longrightarrow & Y. \end{array}$$

This map is in the image of

$$\begin{array}{ccccc} \bar{X} & \curvearrowright & \bar{X} \times Y & \longrightarrow & \bar{Y} \\ \downarrow & & \downarrow & & \downarrow \\ X & \overset{\curvearrowright}{\longleftarrow} & X \times Y & \longrightarrow & Y, \end{array}$$

where the map $\bar{X} \times Y \to \bar{Y}$ is taken to be a conditional. Note that every map in $\mathsf{SChart}(\mathcal{C}^{\to}|_{\mathrm{det}})$ can be represented in this form, by taking a conditional of M given Y to build a section to $M \to X \otimes Y$. Since conditionals are almost-surely equal, by restricting to the support of $X \to X \otimes Y$, we can find a representative which only depends on the overall map $\bar{X} \to \bar{Y}$, which proves that the map from SChart is faithful, concluding the proof.

□

We note that the definition of Markov fibration over $\mathcal{C}$ makes sense as soon as the pullback functor $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$ admits a left adjoint. In general we will not make a study of this outside of the case where $\mathcal{C}$ is pullback-positive, but we briefly note the following trivial case, continuing from Example 3.2.5.

Proposition 3.4.15. *If $\mathcal{C}$ is Cartesian,* $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{Fib}(\mathcal{C})$ *admits a left adjoint (being an identity), the Markov fibrations are exactly the Grothendieck fibrations, and their fiberwise opposites are exactly the fiberwise opposites in the ordinary sense.*

Proof. This is trivial because $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{Fib}(\mathcal{C}_{\mathrm{det}} = \mathcal{C})$ is simply the identity functor, hence it is monadic (with the identity monad,) hence every fibration/prefibration presents a Markov fibration, namely itself, and is in particular a Markov fibration. The fiberwise opposite is simply given by applying the identity (taking the stochastic module on the deterministic part), taking the fiberwise opposite, then applying the identity again (passing to the presented Markov fibration). □

It is worth noting that, even in the case where $\mathsf{SChart}(\mathcal{D}_0^{\mathrm{fop}})$ is not a prefibration, it may still deserve the name "stochastic lenses". For example the stochastic lenses in $\mathsf{BorelStoch}$ can be seen to contain $\mathsf{Optic}(\mathsf{BorelStoch})$ as a full subcategory, even though it does not form a Markov fibration (see Theorem 3.7.2 below).

Part of the motivation for the theory of dependent optics is to identify a category of stochastic optics which admits all coproducts. If $\mathcal{C}$ is distributive, $\mathsf{Optic}(\mathcal{C})$ satisfies $\binom{A}{X} + \binom{A}{Y} = \binom{A}{X+Y}$, but this coproduct fails to exist in general if the two secondary objects are distinct. The idea is that this coproduct $\binom{A}{X} + \binom{A'}{Y}$ should exist as a family indexed by $X + Y$, where $E_x = A$ for $x \in X$, and $E_y = A'$ for $y \in Y$. Our theory accommodates this example under the mild additional hypothesis of *extensiveness*

Definition 3.4.16. A Markov category is said to be an *extensive Markov category* if it admits finite coproducts, whose injections are deterministic, and which satisfy the following equivalent conditions:

1. If we let $\mathcal{C}^{\mathrm{det}}_{/A}$ refer to the full subcategory of the slice spanned by the deterministic morphisms $X \to A$, we have an equivalence of categories $\mathcal{C}^{\mathrm{det}}_{/A} \times \mathcal{C}^{\mathrm{det}}_{/B} \cong \mathcal{C}^{\mathrm{det}}_{/A+B}$, given by taking coproducts

2. $\mathcal{C}_{\mathrm{det}}$ is an extensive category in the usual sense and the inclusion $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ preserves pullbacks along coproduct inclusions.

Proposition 3.4.17. *Let $\mathcal{C}$ be an extensive Markov category, let $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}}$ be a fibration which satisfies $\mathcal{D}_{0,X+Y} = \mathcal{D}_{0,X} \times \mathcal{D}_{0,Y}$. Note that this implies $\mathcal{D}_0$ admits finite coproducts, and they're given exactly by this pairing. Suppose $\mathcal{D}_0$ is equipped with a stochastic module structure. Then $\mathcal{D}_0 \hookrightarrow \mathsf{SChart}(\mathcal{D}_0)$ preserves the finite coproducts. In particular, $\mathcal{D}_0^{\mathrm{fop}}$ has the same coproducts as $\mathcal{D}_0$, and $\mathcal{D}_0^{\mathrm{fop}} \to \mathsf{SLens}(\mathcal{D}_0)$ preserves them as well.*

Proof. This is straightforward to check—the residual $M \to X_1 + X_2$ splits into $M_1 + M_2$ by extensivity of $\mathcal{C}$, which also implies the section must split as the copairing of a section $s_1 : X_1 \to M_1, s_2 : X_2 \to M_2$. By the condition on the fibration, the map in the fiber over M splits into a map over M_1 and a map over M_2. Using the extensivity again, it is straightforward to see that this decomposition respects the equivalence relation. □

In particular, $\mathsf{SChart}(\mathcal{C}^{\to}|_{\mathrm{det}}), \mathsf{SLens}(\mathcal{C}^{\to}|_{\mathrm{det}})$ both admit coproducts given simply as coproducts in $\mathcal{C}^{\to}$, if $\mathcal{C}$ is extensive.

Proposition 3.4.18. *Suppose $\mathcal{D}$ is a Markov fibration so that each pullback functor $f^* : \mathcal{D}_Y \to \mathcal{D}_X$ for $f : X \to Y \in \mathcal{C}_{\mathrm{det}}$ can be taken to be bijective on objects, and that these can furthermore be chosen strictly functorial (so that $(fg)^* = g^* f^*$). Then, writing objects $\bar{X} \in \mathcal{D}_X$ as $\binom{A \in \mathcal{D}_*}{X \in \mathcal{C}}$, where A is the unique object in $\mathcal{D}_*$ which pulls back to $\bar{X}$ under the deletion $X \to *$, (note that this means $f^*\binom{A}{Y} = \binom{A}{X}$) we may characterize the fiberwise dual as having hom-sets*

$$\mathcal{D}^{\mathrm{fop}}(\begin{pmatrix} A \\ X \end{pmatrix}, \begin{pmatrix} B \\ Y \end{pmatrix}) = \mathcal{D}(\begin{pmatrix} B \\ X \end{pmatrix}, \begin{pmatrix} A \\ Y \end{pmatrix})$$

Proof. The correspondence in both directions is obvious by just formally reversing the direction of the map $f : \binom{A}{M} \to \binom{B}{M}$ in a representing tuple $X \leftarrow M \to Y, a : X \to M, f$—the fact that this assignment respects the equivalence relation follows from the fact that the monad preserves fiberwise opposites. □

3.5 Limits of Stochastic Modules and Markov Fibrations

The class of fibrations is stable under pullback. This turns $\mathsf{Fib}(-)$ into an indexed category, which represents a fibration over Cat—this is the "global" category of fibrations Fib, whose objects are fibrations $\mathcal{D} \to \mathcal{C}$, and whose morphisms are commutative squares

$$\begin{array}{ccc} \mathcal{D} & \longrightarrow & \mathcal{D}' \\ \downarrow & & \downarrow \\ \mathcal{C} & \longrightarrow & \mathcal{C}' \end{array}$$

where the top map preserves Cartesian morphisms.

By considering universal constructions like limits and colimits in Fib, additional fibrations can be constructed. It would similarly be useful to study limits in the category of Markov fibrations. Moreover, the products in Fib allow one to express notions of internal pseudomonoid—these turn out to be *monoidal fibrations*, and this is a key part of Moeller and Vasilakopoulou's treatment of the monoidal Grothendieck construction, [MV20]. Since we want to study monoidal Markov fibrations, we should study their limits.

Lemma 3.5.1. *Let $\mathcal{D} \to \mathcal{C}$ be a Markov prefibration, and let $\mathsf{F} : \mathcal{C}' \to \mathcal{C}$ be any functor from another Markov category which preserves deterministic maps. Then the pullback $\mathcal{D}' = \mathcal{D} \times_{\mathcal{C}} \mathcal{C}' \to \mathcal{C}'$ is again a Markov prefibration, and the functor $\mathcal{D}' \to \mathcal{D}$ preserves Cartesian maps.*

Proof. Since pullbacks compose, $\mathcal{D}'|_{\mathrm{det}} \to \mathcal{C}'_{\mathrm{det}}$ is the pullback of $\mathcal{D}|_{\mathrm{det}}$ along $\mathcal{C}'_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$. Since fibrations are stable under pullback, this is a fibration.

Now consider a triangle in $\mathcal{C}'$:

$$\begin{array}{ccc} & & A \\ & \overset{f}{\nearrow} & \downarrow \\ X & \longrightarrow & B, \end{array}$$

with $X, A \to B$ deterministic, and let $\bar{B}_X \to \bar{B} \leftarrow \bar{B}_A$ be Cartesian maps lying over these. We must show there is a unique lift $\bar{f} : \bar{B}_X \to \bar{B}_A$ rendering the lifted triangle commutative. By definition, to give such a morphism is to given one over $F(f)$, and the triangle commutes if and only if its image in $\mathcal{D}$ commutes.

The triangle in $\mathcal{C}'$ goes to a triangle of the same class in $\mathcal{C}$, and the Cartesian maps go to Cartesian maps of the same type. Hence there is a unique lift of $F(f)$ of the given type, which is exactly what we needed to show. □

Remark 3.5.2. Given any functor $F : \mathcal{C} \to \mathcal{C}'$ between Markov categories, we may attempt to define an oplax monoidal structure $F(X \otimes Y) \to F(X) \otimes F(Y)$ by pairing the projections.

This is not necessarily a natural transformation. However, if it is, it automatically equips F with the structure of an oplax monoidal functor. Recall that oplax monoidal functors carry comonoids to comonoids. An oplax monoidal functor between Markov categories preserves the given comonoids if and only if it is induced like this.

Hence, there is at most one way to equip a functor between Markov categories with such a structure—it is a property, not extra structure. Call such a functor an *oplax Markov functor*. Note that oplax Markov functors preserve deterministic morphisms.

(Fritz [Fri20] defines a *Markov functor* to be a *strong* monoidal functor which preserves the comonoids. Clearly this is a proper subset of our oplax Markov functors.)

Definition 3.5.3. We will denote by $\mathsf{MarkPreFib}$ the category whose objects are Markov prefibrations $\mathcal{D} \to \mathcal{C}$, and whose functors are commutative squares

$$\begin{array}{ccc} \mathcal{D} & \xrightarrow{\bar{F}} & \mathcal{D}' \\ \downarrow & & \downarrow \\ \mathcal{C} & \xrightarrow{F} & \mathcal{C}' \end{array}$$

where $\bar{F}$ preserves Cartesian maps, and F is an oplax Markov functor.

We will let $\mathsf{Markov}^{\mathrm{oplax}}$ denote the category of Markov categories and oplax Markov functors. Note that by Lemma 3.5.1, the forgetful functor $\mathsf{MarkPreFib} \to \mathsf{Markov}^{\mathrm{oplax}}$ is a fibration.

There is an obvious functor $\mathsf{MarkPreFib} \to \mathsf{Fib} \times_{\mathsf{Cat}} \mathsf{Markov}^{\mathrm{oplax}}$, which carries a prefibration to the pair of its Markov category and its underlying fibration onto the deterministic part. On each fiber, this admits a left adjoint, as constructed in section 3.3.

By abstract nonsense these left adjoints commute *laxly* with the pullbacks—that is, given an oplax Markov functor $f : \mathcal{C} \to \mathcal{C}'$ and a map of fibrations $\mathcal{D} \to \mathcal{D}'$ over the deterministic part, there is an induced functor $\overline{\mathcal{D}} \to \overline{\mathcal{D}'}$ over f, although this assignment does not preserve Cartesian squares.

However this does give a functor $\mathsf{Fib} \times_{\mathsf{Cat}} \mathsf{Markov}^{\mathrm{oplax}} \to \mathsf{MarkPreFib}$, left adjoint to the restriction. The category of algebras over this monad is fibred over $\mathsf{Markov}^{\mathrm{oplax}}$, with each fiber being the category of stochastic modules over that Markov category, and we get a global functor from $\mathsf{MarkPreFib}$. In the same way, we get a global functor $\mathsf{SChart}(-)$ to $\mathsf{Cat}^{\to}$ which carries each stochastic module $\mathcal{D}_0 \to \mathcal{C}_{\mathrm{det}} \to \mathcal{C}$ to the functor $\mathsf{SChart}(\mathcal{D}_0) \to \mathcal{C}$

We would like to study the limits in here. At this point, we are forced to consider for a moment a bit of higher category theory. Structures on a category defined "up to isomorphism" generally don't play well together with limits in the category Cat, since they are defined "up to equality". For example, we cannot infer from the fact that $\mathcal{C}, \mathcal{D}$ have products and the functors $F, G : \mathcal{C} \rightrightarrows \mathcal{D}$ preserve them that the equalizer of F, G has products, since given A, B in the equalizer, we have $F(A \times B) \cong G(A \times B)$, but not necessarily equality!

For the moment we will restrict ourselves to limits of *strict* (and in particular, strong) monoidal Markov functors, since these always exist. In general one should probably consider some form of homotopy limit, but we will not go into that now. We clearly have:

Proposition 3.5.4. *Let* $\mathsf{Markov}_s \subseteq \mathsf{Markov}^{\mathrm{oplax}}$ *denote the subcategory of* strictly monoidal *Markov functors (i.e those where the oplaxator* $F(X \otimes Y) \to F(X) \otimes F(Y)$ *is the* identity*).*

1. Markov_s *and* $\mathsf{Markov}^{\mathrm{oplax}}$ *admit all products, computed simply as products in* Cat.
2. Markov_s *admits all finite limits, and these are preserved by the inclusion into* $\mathsf{Markov}^{\mathrm{oplax}}$

Proposition 3.5.5. *The category* $\mathsf{MarkPreFib}(\mathcal{C})$ *admits all products, and pullbacks along isofibrations. These are simply computed as limits in* $\mathsf{Cat}_{/\mathcal{C}}$

Proof. It is immediately apparent that products (that is, pullbacks over $\mathcal{C}$) of prefibrations are again prefibrations, since the Cartesian lifts can simply be computed coordinatewise, and the uniqueness property checked coordinatewise.

Let $\mathcal{D} \times_{\mathcal{E}} \mathcal{D}'$ be a pullback of Markov prefibrations, with $\mathcal{D} \to \mathcal{E}$ an isofibration. (Note that their pullback in $\mathsf{Cat}_{/\mathcal{C}}$ is simply their pullback in Cat). First note that since limits commute, the deterministic part is given by $\mathcal{D}_{\mathrm{det}} \times_{\mathcal{E}_{\mathrm{det}}} \mathcal{D}'_{\mathrm{det}}$. Thus to prove this is a fibration, it suffices to note that fibrations are stable under pullback along isofibrations. Given $f : X \to Y \in \mathcal{C}$ and two lifts $\bar{Y}, \bar{Y}'$ which are identified in $\bar{E}$, we get two Cartesian lifts $f^*\bar{Y} \to \bar{Y}, f^*\bar{Y}' \to \bar{Y}'$. These go to two Cartesian lifts in $\mathcal{E}$, and are therefore identified *up to isomorphism,* but we can lift this isomorphism to $\mathcal{D}$ and obtain a pair of lifts in the strict pullback—this is a Cartesian lift.

Since Cartesian lifts are given by pointwise Cartesian lifts, given a triangle $X \to Y \leftarrow Z$ and a stochastic lift $X \to Y$, we have a unique lift in both $\mathcal{D}, \mathcal{D}'$. These both go to lifts in $\mathcal{E}$—since such a lift is also unique, they are identified. Hence there is a unique lift in the pullback. □

Proposition 3.5.6. *The functor* $\overline{(-)} : \mathsf{Fib} \times_{\mathsf{Cat}} \mathsf{Markov}^{\mathrm{oplax}} \to \mathsf{MarkPreFib}$ *preserves products.*

For each Markov category $\mathcal{C}$ *with weak conditionals,* $\overline{(-)} : \mathsf{Fib}(\mathcal{C}_{\mathrm{det}}) \to \mathsf{MarkPreFib}(\mathcal{C})$ *preserves the terminal object, and pullbacks along isofibrations.*

Proof. The product-preservation is clear from the description of $\overline{(-)}$. Let $\mathcal{C}$ admit weak conditionals. It suffices to show that $\overline{(-)}$ preserves the terminal object and pullbacks in $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$

The terminal fibration is $\mathcal{C}_{\mathrm{det}} \to \mathcal{C}_{\mathrm{det}}$. Clearly the terminal Markov prefibration is $\mathcal{C} \to \mathcal{C}$, so we must show that $\overline{\mathcal{C}_{\mathrm{det}}} \to \mathcal{C}$ is an isomorphism.

Its morphisms are simply spans $X \leftarrow M \to Y$ with the left leg equipped with a stochastic section $s : X \to M$, which goes to the composite $X \to Y$. As noted before, this is clearly *full,* by taking $M = X \otimes Y$, and faithful because the pairing $M \to X \otimes Y$ exhibits the equality of this canonical representative with any other.

Now let $\mathcal{D} \to \mathcal{E} \leftarrow \mathcal{D}'$ be a cospan of fibrations over $\mathcal{C}_{\mathrm{det}}$, with $\mathcal{D} \to \mathcal{E}$ an isofibration, and consider the pullback $\mathcal{D} \times_{\mathcal{E}} \mathcal{D}'$. There is a natural transformation

$$\overline{\mathcal{D} \times_{\mathcal{E}} \mathcal{D}'} \to \overline{\mathcal{D}} \times_{\overline{\mathcal{E}}} \overline{\mathcal{D}'}$$

which we must show to be an isomorphism.

Maps on the left-hand side are given by a span $X \leftarrow M \to Y$, a stochastic section $X \to M$, and a map in the pullback of the fibers $\mathcal{D}_M \times_{\mathcal{E}_M} \mathcal{D}'_M$. A map on the right-hand side is given by two spans each equipped with a map, so that they become identified in $\overline{\mathcal{E}}$. Let the apexes of the two spans be M, M'. It suffices to consider the case of a span $M \leftarrow K \to M'$ with a common lifting $X \to K$, so that the pullbacks of the two maps to $\mathcal{E}_K$ agree. But then the original maps may also be pulled back to have K as the underlying span, and thus are in the completion of the pullback.

Similarly, given two maps which become identified in the image, we can again use the identifying maps in $\mathcal{C}$ to identify the original maps, proving faithfulness. This finishes the proof. ◻

Proposition 3.5.7. *For any (pullback-positive)* $\mathcal{C}$*, the monad* $\overline{(-)}|_{\mathrm{det}}$ *on* $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$ *preserves products.*

Proof. There is a canonical map $\overline{\mathcal{D} \times_{\mathcal{C}} \mathcal{D}'}|_{\mathrm{det}} \to \bar{\mathcal{D}}|_{\mathrm{det}} \times_{\mathcal{C}} \bar{\mathcal{D}}|_{\mathrm{det}}$. Clearly the deterministic parts are both isomorphic to $\mathcal{D} \times_{\mathcal{C}} \mathcal{D}'$, and so on this part it is an isomorphism—in particular, bijective on objects. To see it is full, consider an morphism in the codomain, given by a pair of maps $M', M \to X$, sections $s : X \to M, s' : X \to M'$, and maps ϕ, ϕ' in $\mathcal{D}_M, \mathcal{D}'_{M'}$. Then this pair is equivalent to $M \times_X M'$ equipped with the pairing $\langle s, s' \rangle : X \to M \times_X M'$ and the pullbacks of ϕ, ϕ', which is in the image. Given two maps $M \to N, M' \to N'$ witnessing equations with another pair of maps, it's easy to see that this lifts to a map $M \times_X M' \to N \times_X N'$ witnessing the identity between these, so it's faithful. This concludes the proof. ◻

Lemma 3.5.8. *The property of having fibred supports is stable under equalizers in stochastic module fibrations over* $\mathcal{C}$*. If* $\mathcal{C}$ *has weak conditionals, it is also stable under finite products (hence all finite limits).*

Proof. (Note that stochastic modules themselves do not admit all finite limits, requiring some sort of isofibration property—we merely claim here that if the limit exists, it again admits fibred supports)

It is clear that the terminal object $1_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$ has fibred supports (regardless of $\mathcal{C}$). Given an equalizer $\mathcal{E} \hookrightarrow \mathcal{D} \rightrightarrows \mathcal{D}'$, if $M \to X$ is a deterministic map with a stochastic section and $\phi : \bar{X}_M \to \bar{X}_M$ is a map in $\mathcal{E}_M$ which goes to the identity in $\mathcal{E}_X$, find a factorization $X \to N \to M$ so that the image in $\mathcal{D}$ pulls back to the identity over N. Then clearly the same is true for ϕ itself.

Now consider a product $\mathcal{D} \times_{\mathcal{C}} \mathcal{D}'$. The point is that given a pair of maps that go to the identity, we can find $N_0 \to M, N_1 \to M$ where the pullbacks are the identity. We form the pullback $N_0 \times_M N_1$, and use the weak conditionals to find a common lift of the two given sections to this. This gives the required fibred supports. □

3.6 Monoidal Stochastic Modules and Markov Fibrations

Recall that if $\mathcal{C}$ is symmetric monoidal, $\mathsf{Optic}(\mathcal{C})$ inherits a symmetric monoidal structure. At the same time, if $\mathcal{D} \to \mathcal{C}$ is a monoidal fibration, the fiberwise opposite retains a monoidal structure. Since these monoidal structures play an important role both in compositional game theory (where it would not be much of an exaggeration to say the entire point is to use string diagrammatic syntax to work with games) and in categorical systems theory, it is clearly important to understand the monoidal structure on Markov fibrations. Luckily, as we will see in this section, there are essentially no difficulties in accounting for the monoidal structure.

The theory of monoidal fibrations has been developed by Moeller and Vasilakopoulou, [MV20], and Shulman [Shu09]. We briefly sketch it here for convenience. There are essentially two available notions of monoidal fibration:

1. For any category $\mathcal{C}$, the 2-category $\mathsf{Fib}(\mathcal{C})$ admits products, and we can ask for an internal pseudomonoid in this 2-category. This is equivalent to asking for a functor $\mathcal{C}^{op} \to \mathsf{MonCat}$—in other words, for a monoidal structure on each fiber so that the base-change functors become (strong) monoidal.

2. The global category of fibrations Fib admits products, and we may ask for an internal pseudomonoid here. This is what Shulman calls a monoidal fibration: a fibration where $\mathcal{D}, \mathcal{C}$ both come equipped with monoidal structures, the fibration is a *strict* monoidal functor, and Cartesian maps are stable under monoidal product.

By a result of Moeller and Vasilakopoulou, these notions coincide in the case where $\mathcal{C}$ is Cartesian monoidal. Since we are only interested in ordinary fibrations over $\mathcal{C}_{\mathrm{det}}$, which is indeed Cartesian, we may apply this result. However, our notion of monoidal Markov fibration will be a modified version of the latter.

Definition 3.6.1 (Monoidal Markov prefibration)**.** A *monoidal Markov prefibration* is a Markov prefibration $p : \mathcal{D} \to \mathcal{C}$ equipped with a monoidal category structure on $\mathcal{D}$ so that p is *strict* monoidal and so that the underlying fibration is a monoidal fibration (i.e so that p preserves Cartesian lifts).

A *braided* or *symmetric* monoidal Markov prefibration is a monoidal prefibration equipped with a braiding or symmetry on $\mathcal{D}$ so that p is moreover a braided monoidal functor.

Note that a monoidal Markov prefibration is the same thing as an internal pseudomonoid in the global category of prefibrations. We will not delve further into this point, however.

We will need this lemma:

Lemma 3.6.2. *Let* $\mathsf{F}, \mathsf{G} : \mathcal{C}' \rightrightarrows \mathcal{C}$*,* $\mathsf{S} : \mathcal{C} \to \mathcal{C}'$ *be a reflexive pair of identity-on-objects, strict monoidal functors. Let* $\mathsf{E} : \mathcal{C} \to \mathcal{D}$ *be the coequalizer in* Cat*. Then* $\mathcal{D}$ *inherits a monoidal structure making* E *strict monoidal. Moreover, if* $\mathcal{C}$ *is symmetric or braided,* $\mathcal{D}$ *inherits this structure making* E *a braided functor.*

Proof. The only thing to check is that the equivalence relation on morphisms is stable under tensoring. But this is clear: let $f : X \to Y \in \mathcal{C}'$, $g : A \to B \in \mathcal{C}$. Then $f \otimes S(g)$ witnesses the identification of $F(f) \otimes g$ and $G(f) \otimes g$. Tensoring on the right is analogous. This finishes the proof for the monoidal structure.

In the braided or symmetric case, it is clear that the image of the braiding of $\mathcal{C}$ in $\mathcal{D}$ becomes a braiding on $\mathcal{D}$: it clearly satisfies the coherence equations (being a quotient), and naturality follows by simply choosing representatives and noting that the tensor in $\mathcal{D}$ is defined by tensoring representatives in $\mathcal{C}$. Finally if $\mathcal{C}$ is symmetric clearly the equation $\sigma_{A,B}\sigma_{B,A} = 1$ passes to $\mathcal{D}$. □

Theorem 3.6.3. *1. The free prefibration monad on* $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$ *has a canonical lifting to* $\mathsf{MonFib}(\mathcal{C}_{\mathrm{det}})$

2. *Given a monoidal prefibration, its underlying stochastic module acquires the structure of an algebra of this lifted monad.*

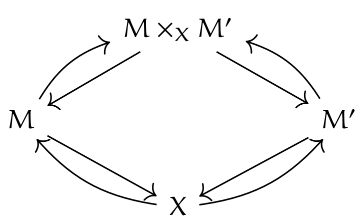
3. *Given an algebra for the lifted monad* $\mathcal{D}_0$*,* $\mathsf{SChart}(\mathcal{D}_0)$ *acquires a monoidal structure so that* $\mathsf{SChart}(\mathcal{D}_0) \to \mathcal{C}$ *is a strict monoidal functor.*

4. *If* $\mathcal{D}_0$ *moreover has fibred supports, this forgetful functor is a monoidal prefibration.*

   Every statement holds also for braided or symmetric fibrations.

Proof. By [MV20], $\mathsf{MonFib}(\mathcal{C}_{\mathrm{det}})$ is equivalent to the category of pseudomonoids in $\mathsf{Fib}(\mathcal{C}_{\mathrm{det}})$. Since the monad preserves products, it must preserve pseudomonoids, which is all we need.

Now suppose $\mathcal{D} \to \mathcal{C}$ is a monoidal prefibration. Then its underlying fibration is a monoidal fibration, hence an object of $\mathsf{MonFib}(\mathcal{C}_{\mathrm{det}})$. The claim is that the functor $\overline{\mathcal{D}|_{\mathrm{det}}}|_{\mathrm{det}} \to \mathcal{D}|_{\mathrm{det}}$ is monoidal. The induced monoidal structure is given on objects by $\otimes_{\mathcal{D}}$ and takes a pair of morphisms in the fiber represented by sections $(s : X \to M, \phi : \bar{X}_{0M}) \to \bar{X}_{1M}$ and $s' : X \to M', \phi' : \bar{X}'_{0M} \to \bar{X}'_{1M}$ to $\langle s, s' \rangle : X \to M \times_X M', \pi^*_M \phi \otimes_{\mathcal{D}} \pi^*_{M'} \phi'$. Recalling that the algebra structure is defined by taking (s, ϕ) to the composite $\bar{X}_0 \to \bar{X}_{0M} \xrightarrow{\phi} \bar{X}_{1M} \to \bar{X}_1$, and chasing the below diagram around, it is apparent that the algebra structure preserves the monoidal structure.



Now let $\mathcal{D}_0$ be a monoidal stochastic module in this sense. It suffices to show that the free prefibration $\overline{\mathcal{D}_0}$ is a monoidal prefibration, by Lemma 3.6.2, and the induced functor $\mathsf{SChart}(\mathcal{D}_0) \to \mathcal{C}$ will clearly be strict monoidal if $\overline{\mathcal{D}_0} \to \mathcal{C}$ is.

To construct this monoidal structure on $\mathcal{D}_0$, simply note that since $\overline{(-)}$ preserves global limits as well, there is an induced monoidal structure on $\overline{\mathcal{D}_0}$ so that the forgetful functor is strict monoidal. Recalling that the Cartesian lifts of $f : X \to Y \in \mathcal{C}$ to $\overline{\mathcal{D}_0}$ are given by the span $X = X \to Y$ and the morphism $1_{f^*\bar{Y}}$, it is easy to see by unwinding the definition that these are stable under tensor.

If $\mathcal{D}_0$ has fibred supports, we have already proven that $\mathsf{SChart}(\mathcal{D}_0) \to \mathcal{C}$ is a strict monoidal functor, and fibred supports are equivalent to the claim that it is a prefibration. Since the Cartesian lifts are just the equivalence classes of the Cartesian lifts in $\overline{\mathcal{D}_0}$, the preceding claim that they are stable under tensor implies the same for $\mathsf{SChart}(\mathcal{D}_0)$, finishing the proof.

The last point is mostly trivial. The product preservation still establishes the lifting to $\mathsf{BrMonFib}(\mathcal{C}_{\mathrm{det}})$ and $\mathsf{SymMonFib}(\mathcal{C}_{\mathrm{det}})$. Given a braided or symmetric monoidal prefibration, the braidings are Cartesian and in the deterministic part, so the underlying fibration is braided/symmetric and they are preserves by the stochastic module structure. The braiding/symmetry on SChart follows again from Lemma 3.6.2, and there is nothing to show for the last point. □

Definition 3.6.4 (Monoidal Markov fibration)**.** A *monoidal Markov fibration* is a monoidal prefibration $\mathcal{D} \to \mathcal{C}$ which is a Markov fibration. It is *braided* or *symmetric* if it is braided or symmetric as a prefibration

Remark 3.6.5. Since $\overline{(-)}$ preserves both global products and fiberwise ones, it induces both a fiberwise monoidal structure and a "global" monoidal structure on $\overline{\mathcal{D}_0}$. Here we only use the global one. If $\mathcal{C}$ were Cartesian, the global one would be induced from the local one by, given maps $\bar{X}_1 \to \bar{X}_2, \bar{Y}_1 \to \bar{Y}_2$, pulling each of them back along the square

$$\begin{array}{ccc} X_1 \otimes Y_1 & \longrightarrow & X_2 \otimes Y_2 \\ \downarrow & & \downarrow \\ X_1 & \longrightarrow & X_2 \end{array}$$

(and the analogous one for Y) and tensoring them over $X_1 \otimes Y_1$. In a Markov prefibration, of course, these pullbacks are not unique unless $X_1 \to X_2$ is deterministic, but there are "canonical" lifts given by tensoring globally with the (fiberwise) unit map over $Y_1 \to Y_2$, and the global tensor is indeed given by the tensor of these canonical lifts (this doesn't provide a noncircular definition of the global tensor, of course). This provides a consistency relation between the two tensor products. Again, we will not dwell on this point.

Proposition 3.6.6 (Markov structure on stochastic charts)**.** *Let $\mathcal{D}$ be a monoidal stochastic module fibration over $\mathcal{C}$ and suppose each fiber $\mathcal{D}_X$ is a Markov category, and this structure is preserved by the pullback functors f^*. Then $\mathsf{SChart}(\mathcal{D})$ carries the structure of a Markov category, so that $\mathsf{SChart}(\mathcal{D}) \to \mathcal{C}$ is a Markov functor.*

If $\mathcal{D} \to \mathcal{C}$ is a Markov prefibration which is also a strict Markov functor, the induced monoidal structure on $\mathcal{D}|_{\mathrm{det}}$ acquires a fiberwise Markov structure.

Proof. Every equation in the definition of Markov category involves only deterministic maps, so this can be verified entirely over $\mathcal{C}_{\text{det}}$. Thus this reduces to the claim: given a monoidal fibration over a Cartesian base, if each fiber has a Markov structure, the global monoidal structure has one as well. Given an object $\bar{X} = \binom{\bar{X}}{X}$, a map $\bar{X} \to \bar{X} \otimes \bar{X}$ is by definition a map $f : X \to X \otimes X$ plus a map $\bar{X} \to f^*(\bar{X} \otimes \bar{X})$. Taking f to be the copy map, the codomain there is by definition the monoidal product in the fiber $\mathcal{D}_X$, and so we simply use the copying map of the fiberwise monoidal structure.

Given a Markov structure on the total category $\mathcal{D}$, we simply apply this idea in reverse and take the map $\bar{X} \to \text{copy}_X^*(\overline{(X)} \otimes \bar{X}) =: \bar{X} \otimes_X \bar{X}$ to be the copy map.

The deletion maps can be handled in an analogous way. □

Corollary 3.6.7. *Let $\mathcal{D}$ be a stochastic module fibration, and suppose each fiber has* coproducts, *and these are preserved by the pullback functors. Then* $\mathsf{SLens}(\mathcal{D})$ *is a Markov category, with monoidal structure given by*

$$\binom{\bar{X}}{X} \& \binom{\bar{Y}}{Y} = \binom{\pi_X^* \bar{X} + \pi_Y^* \bar{Y}}{X \otimes Y}$$

The Markov structure of Corollary 3.6.7 is a generalization of the fact that, if $\mathcal{C}$ has products and coproducts, and the products distribute over the coproducts, then $\mathsf{Lens}(\mathcal{C})$ has products given by $\binom{A}{X} \times \binom{B}{Y} \cong \binom{A \coprod B}{X \times Y}$. See eg [Hed17, Section 8] for more on this.

3.7 Examples

We have already noted several examples throughout. We'll gather a few more here, and also collect a few scattered throughout to make the picture more clear.

First, let us make explicit the example of optics, as strongly as it can be stated

Proposition 3.7.1. *Let $\mathcal{C}$, a Markov category, act on $\mathcal{D}$. Then* $\mathsf{Optic}_{\mathcal{C}}(\mathcal{D}) :=$ $\mathsf{Optic}_{\mathcal{C}}(\mathcal{C}, \mathcal{D})$ *has a functor to $\mathcal{C}$. The deterministic part* $\mathsf{Optic}_{\mathcal{C}}(\mathcal{D})|_{\text{det}}$ *admits the structure of a stochastic module fibration. There is an isomorphism* $\mathsf{Optic}_{\mathcal{C}}(\mathcal{D}) \to$ $\mathsf{SChart}(\mathsf{Optic}_{\mathcal{C}}(\mathcal{D})|_{\text{det}})$.

If $\mathcal{D}$ is itself symmetric monoidal and the action is symmetric (meaning it is given by $M \cdot A = F(M) \otimes A$ for some symmetric monoidal functor $F : \mathcal{C} \to \mathcal{D}$, see [CG23] 5.4.3 and 5.5.12), this stochastic module is symmetric monoidal and the isomorphism is an isomorphism of symmetric monoidal categories.

Proof. We have essentially already seen that the deterministic part $\mathsf{Optic}_{\mathcal{C}}(\mathcal{D})|_{\text{det}} \to$ $\mathcal{C}_{\text{det}}$ is a fibration, with maps $\binom{A}{X} \leftrightarrows \binom{B}{X}$ over X given by $X \cdot B \to A$, and with the pullback functors acting by reparametrization. Since a map $X \to M \otimes Y$ with deterministic marginal on Y is always equal to the independent pairing of $X \to M, X \to Y$, we can slide the former through and identify each optic over a given $X \to Y$ with a map $X \times B \to A$, and note that this map is conversely an invariant of an optic, since it is obtained by postcomposing with $\binom{B}{Y} \to \binom{B}{*}$.

Given $M \to X$, objects $\binom{A}{X}, \binom{B}{X}$, and a map over M classified by $M \cdot B \to A$, a section $s : X \to M$ act by reparametrization. It is clear that this gives the structure of a stochastic module.

The functor from optics takes $f : X \to M \otimes Y, g : M \cdot B \to A$ to the span $X \leftarrow M \otimes X \otimes Y \to Y$ equipped with the obvious map $(M \otimes X \otimes Y) \cdot B \to A$ that simply forgets X, Y. Note that every chart is equivalent to one of this form (given $X \leftarrow M' \to Y$

and $\phi : M' \cdot B \to A$, the map $M' \to M' \otimes X \otimes Y$ exhibits the required equivalence), hence the functor is full.

Moreover, note that each chart is associated with a well-defined optic, given by the maps $X \to M \to M \otimes Y, M \cdot B \to A$. It is straightforward to see both of these maps are preserved by chart equivalence. This gives an inverse to the functor, proving it is faithful. Since it is identity on objects, this finishes the argument.

In the symmetric monoidal case, it is immediately clear that the fibration on $\mathcal{C}_{\mathrm{det}}$ is symmetric monoidal. Since the action is symmetric monoidal, given maps $X \to M, X \to M'$ and $M \cdot B \to A$, $M' \cdot B' \to A'$, it is clear that composing to get maps $X \cdot B \to A, X \cdot B' \to A'$, then tensoring and composing with the diagonal to get $X \cdot (B \otimes B') \to A \otimes A'$, gives the same map as tensoring, then using the map $X \to M \otimes M'$. Hence we have a symmetric monoidal module. It's straightforward to see the functor is symmetric monoidal, and that finishes the argument. □

Slightly orthogonally, we have the following comparison between $\mathsf{SLens}(\mathcal{C}^{\to})$ and $\mathsf{Optic}(\mathcal{C})$:

Theorem 3.7.2. *Let $\mathcal{C}$ be any pullback-positive Markov category. Then $\mathcal{C}^{\to} \to \mathcal{C}$ is a Markov prefibration which thus induces a stochastic module structure on $\mathcal{C}^{\to}|_{\mathrm{det}}$. Writing simply* $\mathsf{SChart}(\mathcal{C}), \mathsf{SLens}(\mathcal{C})$ *for* $\mathsf{SChart}(\mathcal{C}^{\to}|_{\mathrm{det}}), \mathsf{SLens}(\mathcal{C}^{\to}|_{\mathrm{det}})$*, we have:*

1. *There is a functor* $\mathsf{Optic}(\mathcal{C}) \to \mathsf{SLens}(\mathcal{C})$*, which is fully faithful. Dually there is a functor* $\mathsf{coOptic}(\mathcal{C}) \to \mathsf{SChart}(\mathcal{C})$ *which is fully faithful.*

2. $\mathsf{SLens}(\mathcal{C})$ *and* $\mathsf{SChart}(\mathcal{C})$ *both admit symmetric monoidal structures, which make the functors* $\mathsf{SChart}(\mathcal{C}), \mathsf{SLens}(\mathcal{C}) \to \mathcal{C}$ *strict symmetric monoidal, as well as the functors* $\mathsf{Optic}(\mathcal{C}) \to \mathsf{SLens}(\mathcal{C}), \mathsf{coOptic}(\mathcal{C}) \to \mathsf{SChart}(\mathcal{C})$ *strong symmetric monoidal.*

3. *If $\mathcal{C}$ is extensive, this functor preserves the coproducts $\binom{A}{X} + \binom{A}{Y} = \binom{A}{X+Y}$, and* $\mathsf{SChart}(\mathcal{C}), \mathsf{SLens}(\mathcal{C})$ *both admit all finite coproducts.*

4. *If $\mathcal{C}$ moreover has conditionals and supports,* $\mathsf{SChart}(\mathcal{C}) = \mathcal{C}^{\to}$

Thus we have our previous claim that $\mathsf{SLens}(\mathsf{BorelStoch}^{\to}|_{\mathrm{det}})$ contains $\mathsf{Optic}(\mathsf{BorelStoch})$.

We also have some examples of an analytic flavor:

Example 3.7.3. Given a compact Hausdorff space X, a *Banach space bundle* is a space over X, $V \to X$, equipped with a fiberwise (complex) vector space structure $+ : V \times_X V \to V, \cdot : \mathbb{C} \times V \to V$, so that there exists a cover $\{U_i\}$ of X so that for each U_i, there exists a Banach space V_i and a homeomorphism $V \times_X U_i =: V_{U_i} \cong V_i \times U_i$ over U_i, which is moreover linear in each fiber, where V_i is equipped with the norm topology.

Note that this determines a *local* norm on each V_{U_i} (and in particular each V_x) up to equivalence (but no stricter than that). In particular each V_x is a Banach space.

A morphism of Banach space bundles is a continuous map $f : V \to W$ over X which is linear on each bundle. Note that this implies that on a suitable cover U_i, the maps $f : V_{U_i} \to W_{U_i}$ obey $\|f(v)\| \leqslant C_i \|v\|$ for some $C_i \in \mathbb{R}$, for any local norms inducing the topologies, and hence by compactness there exists some C so that $\|f(v)\| \leqslant C \|v\|$ for each v.

If $X \to Y$ is a continuous map, there is a pullback functor $\mathsf{Ban}_Y \to \mathsf{Ban}_X$. The Grothendieck construction of this gives a fibration $\mathsf{BanBun} \to \mathsf{CHaus}$.

Note that Tychonoff spaces include all compact Hausdorff spaces. Therefore consider the full subcategory $\mathsf{CHausStoch} \hookrightarrow \mathsf{TychStoch}$ spanned by these. The fibration BanBun admits the structure of a stochastic module: given $M \to X$, $s : X \to M$ a kernel, and a linear continuous map $f : V \times_X M \to W \times_X M$, given $v \in V_x$, there is an induced function $M_x \to W_x$ given by $f(v, -)$. Since this is bounded (being continuous on a compact space) it is (Bochner) integrable, define $s^* f(v)$ to be this integral.

This example is analogous to optics for the action of categories of Markov kernels on categories of vector spaces (as in Proposition 3.7.1). Note that we do *not* expect this type of example to present a Markov fibration. The reason is simply that, given a parametrized linear map $M \times \mathbb{R}^n \to \mathbb{R}^n$ and a measure on M, the fact that the expectation map $\mathbb{R}^n \to \mathbb{R}^n$ is the identity by no means implies that the original map is almost surely the identity or anything like that. If $(m, e_0) \mapsto e_1$ and m has positive probability, this can be canceled out by $(m', e_0) \mapsto -e_1$. This is impossible for probability kernels.

If we add an assumption of positivity, it seems plausible that examples of this type will present Markov fibrations—but of course, that brings us quite close to categories of Markov kernels in any case.

Note that, as in this example, we do not generally expect $\mathsf{Optic}_{\mathcal{C}}(\mathcal{C}, \mathcal{D})$ to yield a Markov fibration if $\mathcal{D}$ is not another Markov category (and not even then in general, as the case of $\mathsf{BorelStoch}$ shows)—for these, we expect to need a sort of positivity in the fiber as well, which restricts us to things that look like probability kernels.

Proposition 3.7.4 (Stochastic module of P-algebras)**.** *Let $\mathcal{C}$ be a representable, positive Markov category so that $\mathcal{C}_{\mathrm{det}}$ admits intersections and the probability monad P preserves them. Then each slice $(\mathcal{C}_{\mathrm{det}})_{/X}$ inherits a monad structure given by $P_X(B \to X) = P(B) \times_{P(X)} X$. This is pseudofunctorial in X. Moreover, the stochastic module structure on $\mathcal{C}^{\to}|_{\mathrm{det}}$ extends to a stochastic module structure on the fibration representing the pseudofunctor $X \mapsto \mathsf{Alg}(P_X)$.*

Proof. The monad is induced by the adjunction $\mathcal{C}_{\mathrm{det}}/X \leftrightarrows \mathsf{Alg}(P)_{PX}$. Let $f : X \to Y$ (deterministic). For abstract reasons there is a natural transformation $P_Y f^* \to f^* P_X$. Writing this out, we find

$$P(A \times_Y X) \times_{PX} X \to P(A) \times_{PY} X$$

By representability, the unit $X \to PX$ is a monomorphism. Hence a map into $P(A \times_Y X) \times_{PX} X$ is precisely a map in $\mathcal{C}$ into $A \times_Y X$ so that the marginal on X is deterministic. But by pullback-positivity this is precisely a map (in $\mathcal{C}_{\mathrm{det}}$) into $P(A) \times_{PY} X$.

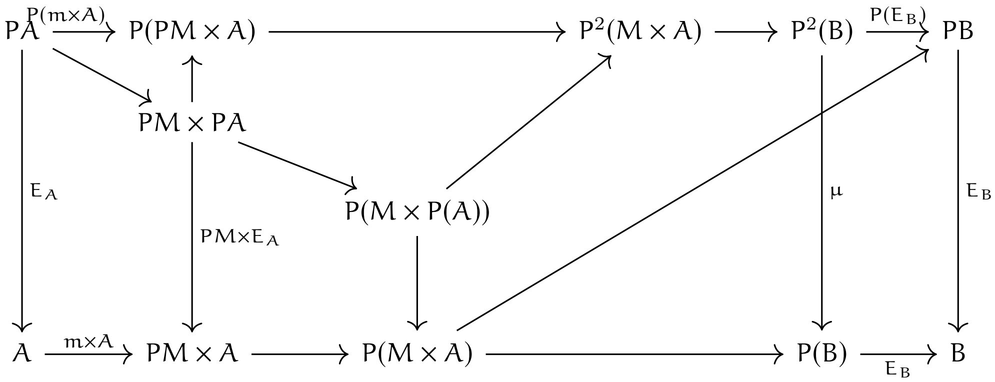
Let $M \to X$, two P_X-algebras A, B, and a map $M \times_X A \to M \times_X B$ which is a homomorphism for the induced P_M-algebras, and a map $X \to PM$ be given. Note that by the above, $P_M(M \times_X A) \cong M \times_X P_X A$. Then the induced map $A \to B$ is given by

$$A \to PM \times_{PX} A \cong P_X(M) \times_X A \hookrightarrow P_X(M \times_X A) \to P_X(B) \to B$$

We must show this is a P_X-homomorphism. Let us simplify by working internally to $(\mathcal{C}_{\mathrm{det}})_{/X}$—thus we have a Cartesian category equipped with a strong commutative monad P, a map $* \to PM$, and a map $M \times A \to B$ which is a parametrized algebra homomorphism, in the sense that the diagram

$$\begin{array}{ccccc} M \times PA & \longrightarrow & P(M \times A) & \longrightarrow & PB \\ \downarrow & & & & \downarrow \\ M \times A & & \longrightarrow & & B \end{array}$$

commutes. Now we must show the map $A \to PM \times A \to P(M \times A) \to PB \to B$ is a P-homomorphism. Write E_A, E_B for the structure maps of the two algebras. Consider this diagram:



The triangle at the top left commutes because P is strong. The square to the right does not commute in general—however, since P is commutative, the composite maps $PM \times PA \to P^2(M \times A) \xrightarrow{\mu} P(X \times A)$ agree. Since the map $P^2(M \times A) \to B$ factors over this, we may replace one edge of this square with another. The square to the right of that is simply $P(-)$ applied to the previous diagram, and so commutes by assumption. The "triangle" under that is just two copies of the same maps, so commutes. The square on the left of the diagram commutes by functoriality of product. The square to the right of that commutes again because P is strong. Hence the outer square commutes, which is precisely the homomorphism property we wanted.

It is apparent that, if $A, B = P_X A', P'_B$ are free algebras, this restricts to the stochastic module structure of $\mathcal{C}^{\to}|_{\mathrm{det}}$ (viewing $\mathcal{C}$ as the Kleisli category of P). But since every algebra is a coequalizer of free algebras, it follows that the action on general algebras is determined uniquely by this. This implies the equations of a stochastic module. □

As in Example 3.7.3, this stochastic module cannot be expected to come from a Markov prefibration in general.

The vast majority of examples seem to occur as subcategories of stochastic modules of the form given by Proposition 3.7.4 (of course, $\mathcal{C}^{\to}$ is just the subcategory spanned fiberwise by the free algebras). In fact, since a stochastic module necessitates in some sense an action of P on the morphisms of the fiber, it seems they do all have this form in a generalized way, although we have not found a better way to make this precise than the existing definition of stochastic module.

Here at the end, let us try to summarize again the relation between the various concepts, to organize the examples:

1. A Markov prefibration is a certain type of category over $\mathcal{C}$. The main examples are $\mathcal{C}^{\to}$ when $\mathcal{C}$ is pullback-positive (most cases of interest), or subcategories of this.

2. A stochastic module fibration is a structure on a fibration over $\mathcal{C}_{\mathrm{det}}$, which allows us to generate a category over $\mathcal{C}$ (and also a fiberwise opposite version of this). Optics are an example of such a presented category.

3. A prefibration gives rise to a stochastic module. Thus we can generate SLens, SChart in this case, which generalizes optics in the case $\mathcal{C}^{\to}$ above.

4. Finally if a prefibration generates *itself* in this case, we call it a Markov fibration. It requires quite strong assumptions on $\mathcal{C}$ for $\mathsf{SChart}(\mathcal{C})$ to be isomorphic to $\mathcal{C}$ in this way, but under weaker assumptions $\mathsf{SChart}(\mathcal{C})$ is itself a Markov fibration, which sits as a subcategory of $\mathcal{C}^{\rightarrow}$.

Chapter 4

The Para construction in generic 2-categories

4.1 Introduction

Myers' theory of categorical systems theory (see section 2.6) gives a rich categorical structure to a wide variety of types of dynamical system. The central idea can be summarized by saying that there are two different, but tightly related, notions of morphism at play in dynamical systems. Open dynamical systems themselves involve a bidirectional information flow, captured by the notion of lens $\binom{X'}{X} \leftrightarrows \binom{Y'}{Y}$, and composition of such lenses describes the composition of subsystems into systems. But morphisms *between* systems are unidirectional, captured by the notion of chart $\binom{X'}{X} \rightrightarrows \binom{Y'}{Y}$. Algebraically, the relationship between these two notions is that they assemble into a *double category*, which indexes the category of systems.

There is another double category involving lenses which has been considered in the categorical study of systems. That is the double category of *parametrized morphisms,* applied to the monoidal category of lenses. These have been studied as an abstraction for gradient descent in several papers, see eg [Cru+21; Gav24], (the idea goes back to [FST19]). Given any action of a monoidal category $\mathcal{M}$ on another category $\mathcal{C}$ (most simply, if $\mathcal{C}$ is monoidal it acts on itself) we obtain a category of parametrized morphisms $f : X \cdot P \to Y$ (where $P \in \mathcal{M}, X, Y \in \mathcal{C}$) and these turn out to be extremely useful. We will mention two applications:

1. A parametrized lens $\binom{P}{P} \otimes \binom{X}{X} \to \binom{Y}{Y}$ is essentially what is called a *learner* in [FST19]—it contains the information necessary to compute a new parameter value p' given an existing $p \in P$ and a sample pair $x \in X, y \in Y$, as well as the additional information required to compose such things. Thus the functoriality of backpropagation can be derived from two facts: the reverse derivative defines a monoidal functor $\mathsf{Euc} \to \mathsf{Lens}(\mathsf{Euc})$, and the construction $\mathsf{Para}(-)$, taking a category to its category of parametrized maps, is itself functorial. This viewpoint has been significantly developed in [Cru+21] and other papers.

2. An *open game*, in the sense of Hedges [Gha+18a], is almost the same thing as a parametrized lens $\binom{\Sigma'}{\Sigma} \otimes \binom{S}{X} \leftrightarrows \binom{R}{Y}$, equipped with a subset $E \subset \Sigma \times \mathsf{Set}(\Sigma, \Sigma')$. Given a context for the game—that is, a state $x \in X$ (describing the state of information when the decision is made) and a continuation $Y \to R$ (describing

how decisions $y \in Y$ map to outcomes $r \in R$,) we obtain a function $k : \Sigma \to \Sigma'$, and we say $\sigma \in \Sigma$ is an equilibrium strategy if $(\sigma, k) \in E$. Since $\mathsf{Set}(\Sigma, \Sigma') = \mathsf{Lens}(\mathsf{Set})(\binom{\Sigma'}{\Sigma}, I)$, and $\Sigma = \mathsf{Lens}(\mathsf{Set})(I, \binom{\Sigma'}{\Sigma})$, this neatly captures the extra data of an open game in terms of the category of lenses. The potential of this idea as a generalized approach to "cybernetic systems" is explored in [Cap+22].

In this chapter, we will develop the theory of the category Para of parametrized maps. We will begin by reviewing the existing literature briefly. We will describe a double categorical version of this category—this does not seem to have appeared in the literature yet, although it has been folklore for at least a few years (and there is nothing complicated about this construction, certainly). The remainder of this chapter will be dedicated to lifting this construction to a generic 2-category $\mathbb{C}$ (with the above being the specialization to $\mathbb{C} = \mathsf{Cat}$). We will derive this lifting using the machinery of 2-category theory. We will see how this generalization accounts for much structure which can be seen to exist on Para, such as its symmetric monoidal structure (assuming $\mathcal{M}, \mathcal{C}$ are symmetric monoidal). But the true application of this will be in chapter 6, where we use this to construct a *triple category* of open dynamical systems.

The definition of the double category $\mathbb{P}\mathsf{ara}$ involves two types of categorical structure with which the reader may be unfamiliar—*actegories*, which are the input to the construction, and pseudo double categories, which are the output. Since we will shortly introduce the abstract internal versions of these, internal pseudomonoid actions and internal pseudocategories, we will not give a separate introduction here. The reader who is unfamiliar with these should refer to [CG23] for actegories, and [Shu09] for (pseudo) double categories. A reader who simply needs a definition may look at Definition 4.5.1 and Example 4.4.2.

4.2 The Para Construction as a double category

In many different situations, we want to understand some morphism as *parametrized* by some data. For example, in machine learning one tries to find a function $f : X \to Y$ with some desirable behaviour by choosing a parametrization $F : P \times X \to Y$ and searching for some $p \in P$ so that $F(p, -)$ has this behaviour (for example by gradient descent on p).

In situations where we want to understand F as being built up as a composite of multiple functions (for example, the layers of a neural network), it is convenient to introduce a category of parametrized morphisms, where the composition combines the parameter spaces of each composite map. We can do this in a general setting with the following definition:

Definition 4.2.1. Let $\mathcal{M}$ be a monoidal category, and let $\bullet : \mathcal{M} \times \mathcal{C} \to \mathcal{C}$ be an action of $\mathcal{M}$ on another category $\mathcal{C}$. Then the Para *construction* is the category $\mathsf{Para}_{\mathcal{M}}(\mathcal{C})$ where

1. Objects are objects of $\mathcal{C}$
2. Morphisms $A \to B$ are tuples $P \in \mathcal{M}, P \bullet A \to B$ up to the natural notion of isomorphism.
3. Composition is by tensoring the parameter objects and composing.

The fact that the parametrization object may live in a different category than the domain and codomain objects, which initially seems like a superfluous generalization, is in fact highly useful. For example, we will often want to consider parametrized morphisms where the parameter space is decorated with some additional data, for example a probability distribution. In most such cases, the category of spaces decorated with such data will act on the category of spaces without such data (simply by forgetting the data and tensoring), and hence we can realize such parametrized morphisms as examples of this Para construction. We may also want only Euclidean parameter space (so that we can run gradient descent simply), but allow the parametrized morphism to go between more general manifolds.

Para, and generalizations of it, have been introduced many times. One of the earliest occurrences is by Hermida and Tennent, [HT12], in the special case of a symmetric monoidal category acting on another such via a functor $i : \mathcal{C} \to \mathcal{D}$ (by a result of Capucci and Gavranović, [CG23], every "structure-preserving" action of a symmetric monoidal category on another has this form). Hermida and Tennent actually give a very intuitive universal property of $\mathsf{Para}_{\mathcal{C}}(\mathcal{D})$: it is freely generated by adding a morphism $I \to i(C)$ for each $C \in \mathcal{C}$, subject to the equations that this must be a monoidal natural transformation. This type of idea actually goes all the way back to Pavlović in [Pav97].

The notation Para was introduced in [FST19], where the special case of parametrized morphisms of Euclidean spaces was used to study gradient descent. In [Cap+22], a bicategorical variant (replacing the quotient by isomorphism in the above variant in the obvious way) is introduced.

Example 4.2.2. Let $\mathcal{C}$ be a monoidal category, and let $S : \mathcal{C} \to \mathsf{Set}$ be any functor. Recall that $\int S$ is the category whose objects are pairs $(X \in \mathcal{C}, s \in S(X))$ and whose morphisms $(X, s) \to (Y, s')$ are $f : X \to Y$ so that $S(f)(s) = s'$. If S is lax monoidal (for $(\mathsf{Set}, \times)$,) $\int S$ acquires a lax monoidal structure making the forgetful functor $\int S \to \mathcal{C}$ strong (even strict) monoidal.

With the action induced by this functor, the bicategory $\mathsf{Para}_{\int S}(\mathcal{C})$ has morphisms given by maps $M \otimes X \to Y$, $m \in S(M)$, and maps given by reparametrizations which preserve the decoration m.

For example, let $\mathcal{C}$ be a Markov category and let $S(X) = \mathcal{C}(I, X)$. Then morphisms are parametrized maps equipped with a measure on the parameter space.

Example 4.2.3. If $\mathbb{C} = \mathsf{Set}$ regarded as a discrete 2-category, as noted, a pseudomonoid action is just a monoid action in the ordinary sense—that is, a monoid M, a set X and a function $m \cdot x$ so that $m \cdot (n \cdot x) = (mn) \cdot x$. The para construction is then the *action category,* whose objects are the points of X and whose morphisms $x \to y$ are elements m so that $m \cdot x = y$ (with multiplication as composition).

Example 4.2.4. Let Ab-Cat be the 2-category of Ab-enriched categories, functors and natural transformations. Recall that a ring R is the same as a one-object category enriched in Ab, and it admits a monoidal structure if and only if it is commutative (this is not particular to Ab-enriched categories). An enriched R-action on an Ab-category $\mathcal{C}$ is then an R-module structure on each hom-set so that composition is R-linear in each variable.

If a category has coproducts, then Set acts on it via $S \cdot X = \coprod_{s \in S} X$. (This is the Set-enriched case of what in enriched categories is called a *copower* or *tensor*, the dual of the power objects from Example 4.3.2). The morphisms of $\mathsf{Para}_{\mathsf{Set}}(\mathcal{C})$ are pairs $(I \in \mathsf{Set}, (f_i : X \to Y \in \mathcal{C})_{i \in I})$, which compose in the obvious way.

Note that this definition clearly makes sense even if $\mathcal{C}$ does not actually have coproducts. This is an example of another construction which has been called Para, which takes a monoidal category $\mathcal{V}$ and a $\mathcal{V}$-enriched category $\mathcal{C}$ and constructs a double category where the morphisms are pairs $(J \in \mathcal{V}, J \to \mathcal{C}(X, Y))$. It is not hard to see that this also extends to a double category in the same way, but we do not presently know the correct definition of *internal enriched object* that would replace pseudomonoid actions to replicate our general theory for this case. Note that the literature contains a notion of *internal enriched category*, [Ghi20], but these are categories enriched over *internal monoidal categories*—that is, the ambient category $\mathcal{E}$ is a 1-category and the categorical structure of the base of enrichment $\mathcal{V}$ is formulated on top of this, not as part of the structure of the objects of the category $\mathcal{E}$

There is a very natural *double* categorical version of Para, where the vertical morphisms are just the unparametrized morphisms of $\mathcal{C}$, and the 2-cells with this boundary:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow & \Downarrow g & \downarrow \\ X' & \longrightarrow & Y' \end{array}$$

are morphisms $M \to N$ (if f is parametrized by M, g by N) so that the square

$$\begin{array}{ccc} M \cdot X & \longrightarrow & Y \\ \downarrow & & \downarrow \\ M \cdot X' & \longrightarrow & Y' \end{array}$$

commutes.

This notion is not original—we learned of it from Matteo Capucci—but it seems to have remained somewhat on the level of folklore. Our main goal for this section will be to construct this (pseudo) double category, and show that it is functorial—that is, given a homomorphism of actions, there is an induced functor between double categories. In fact, we will show that this construction works in any 2-category. Later we will apply it to a pseudomonoid action in (strict) double categories to construct a pseudocategory internal to double categories—our triple category of bimachines.

Theorem 4.2.5. *Let $\mathcal{M}$ be a monoidal category, and let $\mathcal{C}$ be a category equipped with an $\mathcal{M}$-action $\cdot : \mathcal{M} \times \mathcal{C} \to \mathcal{C}$.*

Then there is a pseudo double category $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})$, with (using the notation of internal categories)

1. $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})_0 = \mathcal{C}$
2. $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})_1 = (\cdot : \mathcal{M} \times \mathcal{C} \to \mathcal{C}) \downarrow \mathcal{C}$*, with domain and codomain given by the two projections to* $\mathcal{C}$
3. *The identities map* $\mathcal{C} \to \mathcal{M} \times \mathcal{C} \downarrow \mathcal{C}$ *is given by* $(I, 1_{\mathcal{C}}, 1_{\mathcal{C}}, \lambda)$*, where* $\lambda : I \cdot - \to -$ *is the left unitor of* $\mathcal{M}$*.*
4. *The horizontal composition map is the composition in* Para *: given* $M \cdot X \to Y$, $N \cdot Y \to Z$*, their composite is given by* $(N \otimes M) \cdot X \cong N \cdot (M \cdot X) \to N \cdot Y \to Z$*. The horizontal composition of 2-cells is defined analogously.*

Moreover, if $\mathcal{C} \to \mathcal{D}$ is strict *homomorphism of $\mathcal{M}$-modules, there is an induced pseudofunctor $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C}) \to \mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{D})$. If $\mathcal{N} \to \mathcal{M}$ is a strict monoidal functor, then regarding $\mathcal{C}$ as an $\mathcal{N}$-module along this map, there is an induced pseudofunctor $\mathbb{P}\mathsf{ara}_{\mathcal{N}}(\mathcal{C}) \to \mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})$. These combine into a 2-functor $\mathsf{Act}_s \to \mathsf{PsDbl}_s$ between the 2-category of actions and strictly linear functors and the category of pseudo double categories and strict double functors. This functor preserves (strict) limits.*

Proof. Note that every actegory $\mathcal{M} \curvearrowright \mathcal{C}$ is equivalent to a strict action of a strict monoidal category $\mathcal{M}_s \curvearrowright \mathcal{C}_s$ (in the sense that there exists $\mathcal{M}_s \xrightarrow{\sim} \mathcal{M}$ strong monoidal equivalence and $\mathcal{C}_s \xrightarrow{\sim} \mathcal{C}$ an equivalence such that these maps together form a map of actegories). It follows that it suffices to show unitality and associativity for our composition in the strict case (since these are plainly preserved by equivalence).

Let $\mathcal{M} \curvearrowright \mathcal{C}, \mathcal{M}' \curvearrowright \mathcal{C}'$, $F : \mathcal{M} \to \mathcal{M}'$ be a strict monoidal functor, and let $G : \mathcal{C} \to \mathcal{C}'$ be a linear functor "over F", that is a (strict) $\mathcal{M}$-linear functor when $\mathcal{C}'$ is viewed as a $\mathcal{M}$-actegory along F. Then the induced functor $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C}) \to \mathbb{P}\mathsf{ara}_{\mathcal{M}'}(\mathcal{C}')$ is given by G on the vertical category, and on the horizontal category by

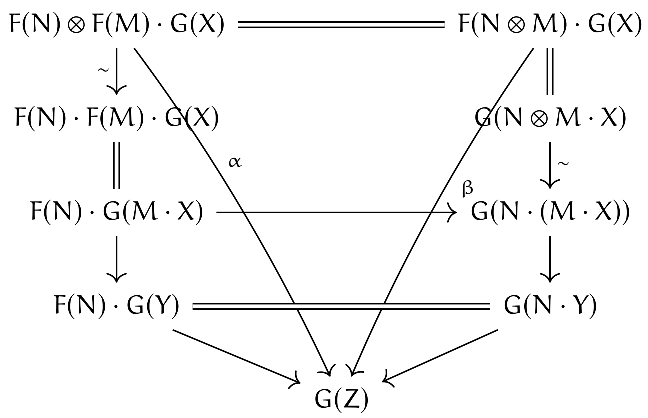
$$(M, X, X', \phi : M{\cdot}X \to X') \mapsto (F(M), G(X), G(X'), F(M){\cdot}G(X) \simeq G(M{\cdot}X) \to G(X'))$$

where the isomorphism is the linearity—noting that the M-action on $\mathcal{C}'$ is precisely acting by $F(M)$.

By naturality of the lineator it is straightforward to see that this is a functor, and it clearly preserves domain and codomain. It preserves units (strictly!) by the compatibility between the unitor and lineator. It remains to see that this preserves composition. For brevity, denote the induced functor H from here. It suffices to verify this for strict actions. So the composite of $M \cdot X \to Y, N \cdot Y \to Z$ is given by the composite

$$(N \otimes M) \cdot X = N \cdot M \cdot X \to N \cdot Y \to Z.$$

Given two such composable maps, call them f, g, the two objects $H(fg), H(f)H(g)$ are given by $(F(M \otimes N), G(X), G(Z), \beta), (F(M) \otimes F(N), G(X), G(Z), \alpha)$,where α, β are respectively the maps across the top and bottom of this diagram:



This proves that H preserves the composition strictly, as desired. (The rightmost triangle commutes by definition, and the leftmost rectangle is a lineator coherence)

Since the strict limits in PsDbl_s and Act_s are computed levelwise, it suffices to observe that the slice category also preserves limits, being a weighted limit itself. □

Remark 4.2.6. For $\mathcal{M}$ a monoidal category, there is a pseudo double category $\mathbb{B}\mathcal{M}$, with trivial vertical category and horizontal category given by $\mathcal{M}$ (see Example 4.4.2). Then $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})$ comes equipped with a strict functor to $\mathbb{B}\mathcal{M}$. In fact, one can recover the action $\cdot : \mathcal{M} \times \mathcal{C} \to \mathcal{C}$ from this data: the object $M \cdot X$ is characterized up to isomorphism by its universal property: parametrized maps $M \cdot X \twoheadrightarrow Y$ with parameter N are in bijection with maps $X \to Y$ parametrized by $N \otimes M$.

4.3 2-Limit sketches

We now give a review of the theory of limit sketches as specialized to 2-categories (that is, categories enriched in categories). Nothing here is new, see [Kel82].

Definition 4.3.1 (Weighted limit)**.** Let $\mathbb{C}$ be a 2-category, let $D : \mathbb{I} \to \mathbb{C}$ be a 2-diagram in it, and let $W : \mathbb{I} \to \mathsf{Cat}$ be another 2-functor. A *limit of* D *weighted by* W is an object $\lim^W D \in \mathbb{C}$ equipped with a natural isomorphism of categories $\mathbb{C}(X, \lim^W D) \cong [\mathbb{I}, \mathbb{C}](W, \mathbb{C}(X, D(-)))$ (natural in $X \in \mathbb{C}$).

Example 4.3.2 (Particular weighted limits)**.** If $W(-) = *$ is constant at the point, and $\mathbb{I}$ is an 1-category, then a weighted limit is simply a $L \to D(i)$ in the underlying category $\mathbb{C}_0$ so that there is additionally an isomorphism of categories $\mathbb{C}(X, L) \xrightarrow{\sim} \lim_i \mathbb{C}(X, D(I))$. Note that this *is* stronger than just a limit in $\mathbb{C}_0$. We will refer to limits of this form by their ordinary names—speaking for example of pullbacks, products, and so on.

If $I = *$, $D(*) = D \in \mathbb{C}$, and $W(*) = C \in \mathsf{Cat}$, the universal property of the weighted limit is that $\mathbb{C}(X, \lim^W D) = \mathsf{Cat}(C, \mathbb{C}(X, D))$. In this case we write $C \pitchfork D$ for this limit if it exists, and call it a *power* of D by C.

If $\mathbb{I} = \{A \to B \leftarrow C\}$ and the weighting is given by $W(B) = \{1 \to 2\}, W(A) = \{1\}, W(B) = \{2\}$ (with the obvious inclusions), then the weighted limit is called the *comma object* and will be denoted $D(A) \downarrow_{D(B)} D(C)$. Note that if $\mathbb{C} = \mathsf{Cat}$ this is precisely the ordinary comma category.

The analogue for a cone on an object C in the setting of weighted limits is called a *cylinder:* it is a natural transformation $W(-) \to \mathbb{C}(C, D(-))$. A cylinder on C induces a natural transformation $\mathbb{C}(X, C) \to [\mathbb{I}, \mathbb{C}](W(-), \mathbb{C}(X, D(-)))$, and we say it's a limit cylinder if this is an isomorphism.

Definition 4.3.3 (2-limit sketch)**.** A *2-limit sketch* is a small 2-category $\mathcal{T}$ equipped with a (small) set of cylinders Θ. A *model* of the sketch in $\mathbb{C}$ is a 2-functor $\mathcal{T} \to \mathbb{C}$ which carries each cylinder in Θ to a limit cylinder. We write the category of models and natural transformations $\mathsf{Mod}(\mathcal{T}, \mathbb{C})$ (leaving the set of cylinders implicit). If $\mathbb{C} = \mathsf{Cat}$, we write simply $\mathsf{Mod}(\mathcal{T})$.

Lemma 4.3.4. *Let* $\mathcal{T}$ *be a 2-limit sketch. Then the functor* $\mathsf{Mod}(\mathcal{T}, \mathbb{C}) \to [\mathbb{C}^{op}, \mathsf{Mod}(\mathcal{T})]$ *given by* $A \in \mathsf{Mod}(\mathcal{T}, \mathbb{C}) \mapsto (C \mapsto (S \in \mathcal{T} \mapsto \mathbb{C}(C, A(S))))$ *is fully faithful, and its essential image consists of those functors* F *so that each* $F(-)(S) : \mathbb{C}^{op} \to \mathsf{Cat}, S \in \mathcal{T}$ *is representable.*

Proof. First note that the codomain can be identified with the subcategory of $[\mathbb{C}^{op} \times \mathcal{T}, \mathsf{Cat}]$ spanned by those F where each $F(C, -)$ is a model. Since the Yoneda embedding preserves limits, the functor $A \mapsto \mathrm{Hom}(-, A(=))$ from $\mathsf{Mod}(\mathcal{T})$ clearly lands inside here. Since $\mathsf{Mod}(\mathcal{T}, \mathbb{C})$ is itself a full subcategory of the functor category $[\mathcal{T}, \mathbb{C}]$, this is fully faithful.

Clearly for each model A and for each $S \in \mathcal{T}$, the functor $\mathbb{C}(-, A(S))$ is representable, by $A(S)$. Conversely, if $F(C, S)$ is such that each $F(-, S)$ is representable, then the currying of $F\,\mathcal{T} \to [\mathbb{C}^{op}, \mathsf{Cat}]$ factors over $\mathbb{C}$, and since the Yoneda embedding preserves those limits that exist, this factorization must be a model as well. □

Proposition 4.3.5. *Let* $\mathcal{T}, \mathcal{T}'$ *be limit sketches and suppose given a functor* $F : \mathsf{Mod}(\mathcal{T}) \to \mathsf{Mod}(\mathcal{T}')$ *which preserves limits and is* accessible, *that is it preserved* κ*-filtered colimits for some* κ*. Then* F *admits a left adjoint* L.

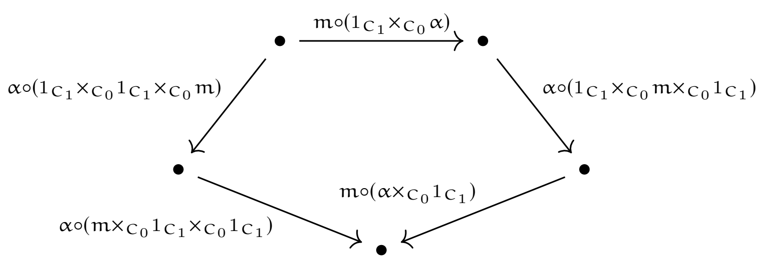
In particular, $F(A)(S) = \mathsf{Mod}(\mathcal{T})(L(y(S)), A)$.

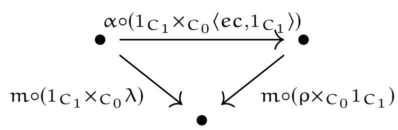
4.4 Pseudocategories

Given a 2-category $\mathcal{C}$, there is a natural loosening of the notion of "internal category in $\mathcal{C}$", given by requiring that associativity and unitality hold only up to chosen coherent isomorphism 2-cells. This is the notion of *pseudocategory*. We refer to [Fer06] for a thorough description of this concept (as well as a review of the earlier literature). However, we will review the basic facts here for convenience.

Definition 4.4.1 (Internal pseudocategory)**.** Let $\mathcal{C}$ be a 2-category with pullbacks. Then an *internal pseudocategory in* $\mathcal{C}$ (or simply a pseudocategory in $\mathcal{C}$) consists of the following data:

1. Two objects $C_0, C_1 \in \mathcal{C}$
2. Morphisms $d, c : C_1 \to C_0, e : C_0 \to C_1$ so that $de = ce = 1_{A_0}$.
3. A morphism $m : C_1 \times_{C_0} C_1 \to C_1$, so that $dm = d\pi_2, cm = c\pi_1$
4. Isomorphism 2-cells $\alpha : m(1_{C_1} \times_{C_0} m) \to m(m \times_{C_0} 1_{C_1})$, $\lambda : m\langle ec, 1_{C_1}\rangle \to 1_C A_1$, and $\rho : m\langle 1_{A_1}, ed\rangle \to 1_{A_1}$
5. Satisfying the following equations:
    (a) $d \circ \lambda = 1_d = d \circ \rho$
    (b) $c \circ \lambda = 1_c = c \circ \rho$
    (c) $d \circ \alpha = 1_{d\pi_3}, c \circ \alpha = 1_{c\pi_1}$
    (d) $\lambda \circ e = \rho \circ e$
6. And so that the following diagrams commute:





A *homomorphism* or *strict functor* of pseudocategories $A \to B$ is a pair of morphisms $F_1 : A_1 \to B_1, F_0 : A_0 \to B_0$ which commute with all the structure—that is, $dF_1 = F_0 d$, $F_1 \alpha = \alpha(F_1 \times_{F_0} F_1 \times_{F_0} F_1)$, and so on. There is a clear notion of natural transformation of homomorphisms. We write $\mathsf{PsCat}_s(\mathbb{C})$ for the 2-category of pseudocategories, homomorphisms and natural transformations in $\mathbb{C}$.

It is important to note that, although this is a weakening of the definition of internal category, pseudocategories are themselves a strict concept—they are defined in terms of equations that must hold up to *strict* equality. Thus for example there is an enriched monad (a 2-monad) on Cat whose *strict* algebras are the pseudo double categories.

Example 4.4.2.

1. An internal pseudocategory in Cat is, as mentioned above, a pseudo double category.
2. An internal pseudocategory A with A_0 terminal is the same thing as an internal pseudomonoid. If we fix a specific terminal object $*$ and require $A_0 = *$, this is an *isomorphism* of 2-categories (both for the categories of pseudomorphisms and homomorphisms).
3. An internal pseudocategory with α, λ, ρ identities is the same thing as a (strict) internal category. In particular, internal pseudocategories in discrete 2-categories are merely internal categories.

Example 4.4.3. Let MonCat_s be the 2-category of monoidal categories, *strict* monoidal functors and monoidal natural transformations. Note that this has pullbacks. Then the objects of $\mathsf{PsCat}(\mathsf{MonCat})$ are a stricter version of *monoidal pseudo double categories*: they are pseudo double categories C_1, C_0 where both C_1, C_0 carry a monoidal structure so that d, c, e, m are strict monoidal functors and ρ, λ, α are monoidal natural transformations.

Here we are already beginning to feel the limitations of this internalization a bit—really we should ask for m, at least, to be only a strong monoidal functor. But we move on for now with this problematic definition.

Of course, there is a very crucial notion of *pseudofunctor* between pseudo double categories. This has an internal formulation in terms of *pseudomorphisms:*

Definition 4.4.4 (Pseudomorphism of pseudocategories)**.** Let

$$C = (C_0, C_1, d, c, e, m, \alpha, \lambda, \rho)$$

$$D = (D_0, D_1, d', c', e', m', \alpha', \lambda', \rho')$$

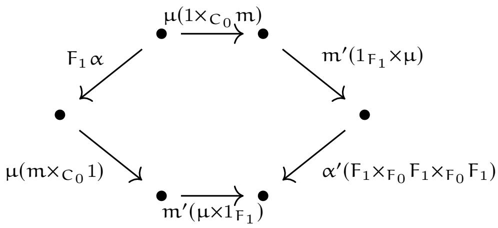
be internal pseudocategories in a 2-category $\mathbb{C}$. A *pseudomorphism* $F : C \to D$ consists of the following data:

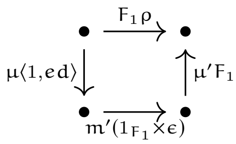
1. Morphisms $F_0 : C_0 \to D_0, F_1 : C_1 \to D_1$ in $\mathbb{C}$
2. Isomorphism 2-cells $\mu : F_1 m \to m(F_1 \times_{F_0} F_1)$ and $\epsilon : F_1 e \to e F_0$

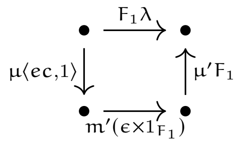
So that the following equations hold:

1. $d' F_1 = F_0 d, c' F_1 = F_0 c$
2. $d' \circ \mu = 1_{F_0} d\pi_2, c' \circ \mu = 1_{F_0 c \pi_1}$
3. $d' \circ \epsilon = 1_{F_0}, c' \circ \epsilon = 1_{F_0}$

And so that the following diagrams commute:







One should not spend too much time looking at Definition 4.4.4. This is clearly a horrible concept, and we will prefer to work around it rather than referring to it explicitly. It may be a useful exercise to convince oneself that, when $C_0 = *$ and the pseudocategory is a pseudomonoid—eg a monoidal category—this coincides with the notion of strong monoidal functor.

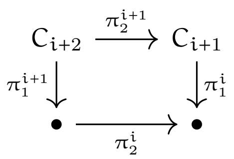
One can also define lax and oplax functors between pseudo double categories, but we will not need that concept here.

Proposition 4.4.5. *There exists a* Cat*-limit sketch* $(\mathcal{T}_{\mathsf{PsCat}}, \Theta_{\mathsf{PsCat}})$ *whose strict models are internal pseudocategories, and whose strict natural transformations are homomorphisms.*

Proof. There is nothing surprising about this to those familiar with limit sketch presentations of other algebraic gadgets, but for completeness we give an explicit construction here.

Take $\mathcal{T}_{\mathsf{PsMon}}$ to have objects C_0, C_1, C_2, C_3, C_4. The 2-category is freely generated by these objects and the following data:

1. Morphisms $d, c : C_1 \to C_0$
2. Morphisms $\pi_1^i, \pi_2^i : C_{i+1} \to C_i$ for $i = 1, 2, 3$, so that

   commutes.

3. Morphisms $m : C_2 \to C_1$ and $e : C_0 \to C_1$ satisfying $dm = d\pi_1^1, cm = c\pi_2^1, de = ce = 1_{C_0}$

4. Further morphisms, and invertible 2-cells α, λ, ρ, as in the following diagrams:

$$\begin{array}{ccc} C_3 & \xrightarrow{\langle m, 1_{C_1}\rangle} & C_2 \\ {\scriptstyle \langle 1_{C_1}, m\rangle}\downarrow & \overset{\alpha}{\swarrow\!\!\!\swarrow} & \downarrow{\scriptstyle m} \\ C_2 & \xrightarrow[m]{} & C_1 \end{array}$$

$$\begin{array}{ccccc} C_1 & \xrightarrow{\langle e, 1_{C_1}\rangle} & C_2 & \xleftarrow{\langle 1_{C_1}, e\rangle} & C_1 \\ & {\scriptstyle 1_{C_1}}\searrow & \overset{\lambda}{\Rightarrow} \downarrow{\scriptstyle m} \overset{\rho}{\Leftarrow} & \swarrow{\scriptstyle 1_{C_1}} & \\ & & C_1 & & \end{array}$$

   and all the other induced morphisms appearing in Definition 4.4.1, subject to those equations holding.

5. Equations asserting that the morphisms and 2-cells with codomain $C_i, i > 1$ behave as expected under postcomposition with the pullback projections. □

The prescribed pullback cones Θ are the diagrams involving π_j^i for $i = 0, 1, 2$.

An explicit description of this sketch seems not to have appeared in the literature until [ABK24, Example 5.13]. Note that they construct a richer object, what they term an $\mathcal{F}$*-sketch,* which includes the information that the domain and codomain are "tight" and must be preserved strictly by pseudofunctors, but composition and identities are "loose" and may be preserved only up to natural isomorphism. This idea will be highly useful later, although we will not go into a full treatment of their theory.

4.5 Actegories

Definition 4.5.1 (Actegory)**.** Let $\mathcal{M}$ be a monoidal category. A $\mathcal{M}$-*actegory* (also $\mathcal{M}$-module, $\mathcal{M}$-action) is a category $\mathcal{C}$ equipped with a functor $\cdot : \mathcal{M} \times \mathcal{C} \to \mathcal{C}$ called the *action* (written infix, $M \cdot C$), and natural isomorphisms $\mu : (M \otimes M') \cdot C \to M \cdot (M' \cdot C), \eta : I \cdot C \to C$ satisfying the following coherence equations:

$$\begin{array}{ccccc} M \cdot (N \cdot (P \cdot C)) & & \xrightarrow{\mu_{M,N,P \cdot C}} & & (M \otimes N) \cdot (P \cdot C) \\ {\scriptstyle M \cdot \mu_{N,P,C}}\downarrow & & & & \downarrow{\scriptstyle \mu_{M \otimes N, P, C}} \\ M \cdot ((N \otimes P) \cdot C) & & & & ((M \otimes N) \otimes P) \cdot C \\ & {\scriptstyle \mu_{M, N \otimes P, C}}\searrow & & \swarrow{\scriptstyle \alpha_{M,N,P} \cdot C} & \\ & & (M \otimes (N \otimes P)) \cdot C & & \end{array}$$

$$\begin{array}{ccc} I \cdot (M \cdot C) & \xrightarrow{\mu_{I,M,C}} & (I \otimes M) \cdot C \\ & {\scriptstyle \eta^{-1}_{M \cdot C}}\searrow & \downarrow{\scriptstyle \lambda_M \cdot C} \\ & & M \cdot C \end{array}$$

$$\begin{array}{ccc} M \cdot (I \cdot C) & \xrightarrow{\mu_{I,M,C}} & (M \otimes I) \cdot C \\ & \searrow^{M \cdot \eta_C^{-1}} & \downarrow{\scriptstyle \rho \cdot C} \\ & & M \cdot C \end{array}$$

Actegories are a very natural concept, and as one would expect their history goes back a long way. The concept has been considered by Benabou all the way back in [Ben67], and used many times since them. We will not delve too deeply into their theory here—see [CG23] for a thorough treatment.

Just as pseudo double categories, actegories really make sense in every 2-category, as a weakening of monoid actions.

Definition 4.5.2 (Pseudomonoid action)**.** Let $\mathbb{C}$ be a 2-category. A pseudomonoid action internal to $\mathbb{C}$ consists of the following data: An internal pseudomonoid M, an object C, a morphism $\cdot : M \times C \to C$, natural isomorphisms $\mu : \cdot(\otimes \times 1_C) \to \cdot(1_M \times \cdot)$ and $\eta : \cdot\langle e, 1_C \rangle \to 1_C$, satisfying the coherence equations from Definition 4.5.1.

A *strict homomorphism of actions* is a pair $F_m : M \to M', F_c : C \to C'$ so that F_m is a strictly monoidal functor and F_c preserves the action strictly, i.e $F_m(M) \cdot F_c(C) = F_c(M \cdot C)$, $F_c(\eta) = \eta'$, etc. Note that this makes sense even if the monoidal category or action is not itself strict.

Proposition 4.5.3. *There is a limit sketch* $\mathcal{T}_{\mathsf{Act}}, \Theta_{\mathsf{Act}}$ *whose models are tuples* $(M, C, \cdot)$ *of a pseudomonoid* $M = (M, \otimes, I)$*, an object* C *and an action* $\cdot$ *of* M *on* C*. The strict natural transformations between models are strictly linear functors.*

Proof. As above, there is nothing difficult about this. Let $\mathcal{T}_{\mathsf{Act}}$ contain as a subcategory the sketch of pseudomonoids $\mathcal{T}_{\mathsf{PsMon}}$, along with three additional objects C, MC, MMC, projections $\pi^1_M, \pi^1_C : MC \to M, C$, $\pi^2_{M_1}, \pi^2_{M_2}, \pi^2_C : MMC \to M, M, C$, and limit cones expressing these as the products $M \times C$ and $M \times M \times C$ respectively, a morphism $\cdot : MC \to C$, and isomorphism 2-cells $\eta : 1_C \Rightarrow \cdot\langle I, 1_C \rangle$ and $\mu : \cdot\langle \pi^2_{M_1}, \cdot\langle \pi^2_{M_2}, \pi^2_C \rangle\rangle \to \cdot\langle \otimes\langle \pi^2_{M_1}, \pi^2_{M_2} \rangle, \pi^2_C \rangle$, subject to the coherence equations in Definition 4.5.1. □

Example 4.5.4 (Discrete pseudomonoid)**.** If $\mathbb{C} = \mathcal{C}$ is an ordinary category with finite products (viewed as a discrete 2-category), a pseudomonoid is simply an internal monoid, and an action is just a monoid action in the ordinary sense.

Example 4.5.5 (Monoidal and symmetric monoidal actegories)**.** In [CG23], the authors study actegories with extra monoidal structure. Although they do not introduce pseudomonoid actions in a general 2-category, they study actegories internal to monoidal categories, which they call *monoidal actegories*. It is straightforward to check that pseudomonoid actions in MonCat agree with their notion: such an action consists of a braided monoidal category $\mathcal{M}$, an ordinary monoidal category $\mathcal{C}$, a monoidal functor $\mathcal{M} \times \mathcal{C} \to \mathcal{C}$, and monoidal natural transformations μ, η satisfying the coherence equations for an actegory.

(Note that since the functor $\otimes$ appears on one side of μ, to speak of μ being a monoidal natural transformation we must have a monoidal structure on $\otimes$—this makes $\mathcal{M}$ into a braided monoidal category).

For the case of symmetric monoidal actegories, since SymMonCat is cocartesian, PsMon(SymMonCat) $\simeq$ SymMonCat. Given two symmetric monoidal categories $\mathcal{M}, \mathcal{C}$, [CG23] show that an action is simply given by a strong symmetric monoidal functor $F : \mathcal{M} \to \mathcal{C}$ (and in this case the action is $M \cdot C = F(M) \otimes C$). Unwinding this construction we see that $\mathsf{Act}_s(\mathsf{SymMonCat})$ is bi-equivalent to a category which has as objects strong (but not strict) symmetric monoidal functors $C \to C'$, and morphisms squares

$$\begin{array}{ccc} C & \longrightarrow & D \\ \downarrow & & \downarrow \\ C' & \rightsquigarrow & D' \end{array}$$

of symmetric monoidal functors, where the top map is strict, and which commute strictly.

Again, as expected, the notion of strict linear morphism is usually too strict. One can expect an equation like $F_m(M) \cdot F_c(C) = F_c(M \cdot C)$ to hold only up to coherent isomorphism. Let us now make this clear:

Definition 4.5.6 (Pseudolinear map)**.** Let $\mathbb{C}$ be a 2-category, and let $(M, C, \cdot), (N, D, \star)$ be internal pseudomonoid actions. A *pseudolinear map* (or just *linear*) is a pair of maps $F_m : M \to N, F_c : C \to D$, equipped with a pseudomonoid pseudohomomorphism structure ϕ, τ on F_m and a natural isomorphism $l : F_m(-) \cdot F_c(=) \to F_c(- \cdot =)$ satisfying the following coherence conditions (note that writing functors using element-notation like this is well-defined)

$$\begin{array}{ccc} F_m(M) \star (F_m(N) \star F_c(C)) & \xrightarrow{F_m(M)\star l_{N,C}} & F_m(M) \star F_c(N \cdot C) \\ \downarrow \mu^{\star}_{F_m(M),F_m(N),F_c(C)} & & \\ (F_m(M) \otimes F_m(N)) \star F_c(C) & & \downarrow l_{M,N\cdot C} \\ \downarrow \phi_{M,N} \star F_c(C) & & \\ F_m(M \otimes N) \star F_c(C) & & F_c(M \cdot (N \cdot C)) \\ & \searrow^{l_{M\otimes N,C}} \quad {}^{F_c(\mu^{\cdot}_{M,N,C})}\swarrow & \\ & F_c((M \otimes N) \cdot C) & \end{array}$$

$$\begin{array}{ccc} F_m(I) \star F_c(C) & \xrightarrow{l_{I,C}} & F(I \cdot C) \\ \tau \uparrow & & \uparrow F_c(\eta_C) \\ I' \star F_c(C) & \xrightarrow[\eta^{\star}_{F_c(C)}]{} & F_c(C) \end{array}$$

In [CG23], these are defined in two steps: first linear functors for the same $\mathcal{M}$ (those with $F_m = 1_{\mathcal{M}}$) are considered. Then, given a strong monoidal $\mathcal{M} \to \mathcal{N}$, a functorial assignment of an $\mathcal{M}$-action to every $\mathcal{N}$-action is constructed, and the full category of actions is defined as the Grothendieck construction of this. There is nothing preventing this from working in the setting of a general 2-category, and it is straightforward to verify that our notion of linear morphism of actions agrees with theirs.

4.6 $\mathbb{P}$ara for a general 2-category

We are now almost ready to prove:

Theorem 4.6.1. *Let $\mathbb{C}$ be a 2-category which admits pullbacks, products and comma objects. Then there is a functor $\mathbb{P}\mathsf{ara} : \mathsf{Act}(\mathbb{C}) \to \mathsf{PsCat}(\mathbb{C})$ from the 2-category of pseudomonoid actions and pseudolinear maps to the 2-category of pseudocategories and pseudofunctors, with $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})_0 = \mathcal{C}$ and $\mathbb{P}\mathsf{ara}_{\mathcal{M}}(\mathcal{C})_1 = \mathcal{M} \times \mathcal{C} \downarrow \mathcal{C}$.*

This functor preserves strict maps, limits, and filtered colimits.

The main ingredient missing is a characterization of the respective notions of pseudomorphism in terms of the limit sketches. This we do now:

Proposition 4.6.2. *Let $\mathbb{C}$ be a 2-category and let $C, D : \mathcal{T}_{\mathsf{PsCat}} \to \mathbb{C}$ be pseudocategories, represented as models of the theory. Then a pseudonatural transformation $F : C \to D$ which is strict on c, d, e and the limit projections π^i_j is equivalently a pseudomorphism between the pseudocategories.*

Now let $(M, C), (N, D) : \mathcal{T}_{\mathsf{Act}} \to \mathbb{C}$ be pseudomonoid actions. Then a pseudonatural transformation F_m, F_c which preserves the product projections strictly is equivalently a pair of a pseudomorphism $F_m : M \to N$ and a functor $F_c : C \to D$ which is pseudolinear with respect to the action $F_m(-) \cdot =$ of M on D.

Proof. Both of these are essentially a matter of unwinding the definitions. Let us start with pseudocategories. The requirement that F is strict on the product projections amounts to requiring that $F_{C_2} : C_2 \to D_2$ is given *strictly* as the pullback $F_1 \times_{F_0} F_1$, and not merely up to 2-isomorphism. With this, μ is just a naturality transformation for m, and ϵ is a naturality transformation for e. Note that since every map in $\mathcal{T}_{\mathsf{PsCat}}$ is induced from m, e, c, d and the pullback structure, this means the other naturality transformations are determined by μ, ϵ.

The composite $m'(1_{F_1} \times \mu)\mu(1 \times_{C_0} m)$ is the naturality transformation for the map $m(1 \times m)$, and analogously for the composite $m'(\mu \times 1_{F_1})\mu(m \times_{C_0} 1)$. Hence the hexagon is a pseudonaturality coherence for the 2-cell α. Analogously the two squares can be obtained as coherence 2-cells. Conversely, these suffice to make F a pseudonatural transformation, again since all the other 2-cells are generated by α, ρ, λ.

Now let us look at pseudomonoid actions. As above, such a pseudonatural transformation is determined by the functors $F_m : M \to N$ and $F_c : C \to D$ plus some 2-cells involving these and their products. As a special case of the above when $C_0, D_0 = *$ we find that the restriction F_m of the pseudonatural transformation to the pseudomonoid part is equivalently a strong monoidal functor. Now the required lineator $l : \cdot_D(F_m \times F_c) \to F_m(\cdot)$ is a naturality square for the multiplication map $\cdot \in \mathcal{T}_{\mathsf{Act}}$. As above this determines all the other naturality transformations because all the other maps are generated from $\cdot$ (and the pseudomonoid structure) using the product structure. Also analogously to the above argument, the coherence pentagon is a pseudonaturality coherence for the natural isomorphism μ, and the coherence triangle for η. ◻

Remark 4.6.3. This characterization of the notions of pseudomorphism cannot help but seem a little ad hoc. The basic idea here seems to go back to Power in [Pow99]. After constructing an enriched Lawvere theory (which is just a special kind of limit sketch) corresponding to any finitary enriched monad, he notes that in the Cat-enriched case the pseudomorphisms of monad algebras correspond to exactly those pseudonatural transformations which preserve the products strictly.

An elegant theory which generalizes this basic idea has recently been developed by Bourke, Ko, and Varkor in [ABK24] (as we mentioned briefly before). Their theory involves decorating the sketch with a set of *tight* morphisms, which the transformations must be strictly natural with respect to. Beyond developing a general theory of this (not specialized to a few examples as above), they also show how to construct the categories of *(op)lax* homomorphisms, and develop a commutativity of internalization principle for their models. This could be profitably applied in our case to understand better, for example, the implied notion of monoidal triple category, but we will not pursue this here.

Proof of Theorem 4.6.1. Recall that $\mathsf{Act}(\mathbb{C}) = \mathsf{Mod}(\mathcal{T}_{\mathsf{Act}})$ and $\mathsf{PsCat}(\mathbb{C}) = \mathsf{Mod}(\mathcal{T}_{\mathsf{PsCat}})$ (by definition). Also recall that $\mathsf{Mod}(\mathcal{T}, \mathbb{C})$ can be identified with the subcategory of $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}(\mathcal{T})]$ spanned by the levelwise representable presheaves. Note that for $\mathcal{T} = \mathcal{T}_{\mathsf{PsCat}}$, it suffices to verify representability for $C_0, C_1 \in \mathcal{T}_{\mathsf{PsCat}}$, since the rest are pullbacks of these, and pullbacks of representable presheaves are again representable since $\mathbb{C}$ admits pullbacks.

Postcomposition with the previously-constructed functor $\mathsf{Mod}(\mathcal{T}_{\mathsf{Act}}) = \mathsf{Act}(\mathsf{Cat}) \to \mathsf{PsCat}(\mathsf{Act}) = \mathsf{Mod}(\mathcal{T}_{\mathsf{PsCat}})$ gives a functor $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}(\mathcal{T}_{\mathsf{Act}})] \to [\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}(\mathcal{T}_{\mathsf{PsCat}})]$. Clearly this functor preserves representability (again, since limits of representable functors are representable). This gives the desired functor $\mathsf{Mod}_s(\mathcal{T}_{\mathsf{Act}}, \mathbb{C}) \to \mathsf{Mod}_s(\mathcal{T}_{\mathsf{PsCat}}, \mathbb{C})$.

Under the identification of $\mathsf{Mod}(\mathcal{T}, \mathbb{C})$ with a subcategory of $[\mathbb{C}^{\mathrm{op}} \times \mathcal{T}, \mathsf{Cat}]$, it is clear that the pseudonatural transformations of models correspond to those pseudonatural transformations which are strict on morphisms in $\mathbb{C}^{\mathrm{op}}$. This implies the full category $\mathsf{Mod}_p(\mathbb{C})$ of models and pseudonatural transformations can be identified with the subcategory of $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}_p(\mathcal{T}, \mathsf{Cat})]$ spanned by those 2-functors which come from models—that is, $A : \mathbb{C}^{\mathrm{op}} \to \mathsf{Mod}_p(\mathcal{T}) = \mathsf{Mod}_p(\mathcal{T}, \mathsf{Cat})$ must factor over $\mathsf{Mod}(\mathcal{T}) \hookrightarrow \mathsf{Mod}_p(\mathcal{T})$.

(To be clear: the morphisms in $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}_p(\mathcal{T})]$ are strictly natural transformations between functors $\mathbb{C}^{\mathrm{op}} \to \mathsf{Mod}_p(\mathcal{T})$, but each component $A(C) \to B(C)$ of such a natural transformation is a pseudonatural transformation of models).

Under *this* identification, it is clear that pseudofunctors correspond to those natural transformations in $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}_p(\mathcal{T}_{\mathsf{PsCat}})]$ which are valued in pseudofunctors, and analogously for pseudolinear morphisms and maps $[\mathbb{C}^{\mathrm{op}}, \mathsf{Mod}_p(\mathcal{T}_{\mathsf{Act}})]$. Hence it suffices to observe that the Cat-valued version $\mathsf{Mod}(\mathcal{T}_{\mathsf{Act}}) \to \mathsf{Mod}(\mathcal{T}_{\mathsf{PsCat}})$ carries pseudolinear maps to pseudofunctors.

Note that $\mathbb{P}\mathsf{ara} : \mathsf{Mod}(\mathcal{T}_{\mathsf{Act}}) \to \mathsf{Mod}(\mathcal{T}_{\mathsf{PsCat}})$ is given objectwise as a finite limit. Moreover, since the limit sketch of pseudocategories only involved finite limits, the class of pseudocategories is stable under filtered colimits in $[\mathcal{T}_{\mathsf{PsCat}}, \mathsf{Cat}]$. Hence $\mathbb{P}\mathsf{ara}$ commutes with filtered colimits, and is in particular accessible. Hence it admits a left adjoint $L\ (M, C) \mapsto \mathrm{Hom}_{\mathsf{Mod}(\mathcal{T}_{\mathsf{Act}})(L(-),(M,C))}$ where $L : \mathcal{T}_{\mathsf{PsCat}}^{\mathrm{op}} \to [\mathcal{T}_{\mathsf{Act}}, \mathsf{Cat}]$ is the restriction of the left adjoint of $\mathbb{P}\mathsf{ara}$. Both forming the hom-category $\mathrm{Hom}(-, (M, C))$ and precomposing with L are 2-functors, and as such preserve pseudonatural transformations. Thus it only remains to observe that $\mathbb{P}\mathsf{ara}$ also preserves the partial strictness property, but this is clear. □

We have not yet given any thought to functoriality in $\mathbb{C}$, but passing through the constructions, it is apparent that we have:

Proposition 4.6.4. *If $\mathbb{C} \to \mathbb{D}$ is a 2-functor that preserves finite 2-limits, the square*

$$\begin{array}{ccc} \mathsf{Act}(\mathbb{C}) & \longrightarrow & \mathsf{Act}(\mathbb{D}) \\ \downarrow & & \downarrow \\ \mathsf{PsCat}(\mathbb{C}) & \longrightarrow & \mathsf{PsCat}(\mathbb{D}) \end{array}$$

commutes up to strict natural isomorphism. (In particular, the square involving the subcategories of strict homomorphisms also commutes up to natural isomorphism).

(It may seem wrong that, after working strictly all this time, this square commutes only up to natural isomorphism, but in fact that is the strict notion—the weak version of this statement would be that it commuted up to natural *equivalence*. Note that eg. the comma objects $\mathsf{M} \times \mathsf{C} \downarrow \mathsf{C}$ are only characterized up to isomorphism, so this is really the best we can hope for.)

The main problem with Theorem 4.6.1 is that for many 2-categorical notions of interest, requiring (for example) the action $\mathcal{M} \times \mathcal{C} \to \mathcal{C}$ to be a strict homomorphism of whatever structure under consideration is too strict to work, while working with the full category of pseudomorphisms prevents the pullbacks required for pseudocategories from existing. Thus for example, an internal pseudomonoid in MonCat_s is a *commutative* monoidal category, which is far too strict for most purposes—generally speaking, symmetric monoidal categories can *not* be strictified into commutative ones.

In chapter 6, we will want to construct a symmetric monoidal "triple category"—that is, a symmetric monoidal pseudocategory in strict double categories. We will manage this via the preceding by noting that since $\mathsf{Act}(\mathbb{C}) \to \mathsf{PsCat}(\mathbb{C})$ preserves products, it carries (symmetric) pseudomonoids to pseudomonoids—in other words, we can apply the internalization the other way around, taking pseudomonoids in actions rather than actions in pseudomonoids. This works because strict *products* still exist in the category of pseudohomomorphisms, but for a more general categorical structure, we would be in trouble.

Chapter 5

Open Games with external choice in Markov Fibrations

5.1 Introduction

Game theory is a field of economics which studies mathematical models of human decision-making. Classically, game theory is particularly interested in the behavior resulting from individual agents optimizing simple objectives (such as the expected value of some real-valued function of each players' decision) when multiple such players interact. Although the analysis of games, obviously, has a long prehistory, the field was put on its modern theoretical footing by Von Neumann in 1928 ([V N28], see also [VM07] for a more thorough treatment from this era). The basic problem studied here is this: given two finite sets Σ_1, Σ_2 of *strategies*—that is, the choices available to the two players—and a function $f : \Sigma_1 \times \Sigma_2 \to \mathbb{R}$ which assigns to a pair of such choices a score, which player 1 seeks to maximize and player 2 seeks to minimize, what can be said about the rational decision-making of each player?

Suppose each player has access to some source of randomness, so that they may each choose a distribution on their respective strategy set to play. This is called a *mixed strategy*. Suppose player 2 has the opportunity to move *knowing* the distribution chosen by player 1 (but not the value drawn from it), and suppose he prefers to minimize the expected outcome of the game. Writing $f(\mu_1, \mu_2)$ for $\mathbb{E}_{x\sim\mu_1, y\sim\mu_2} f(x, y)$ for brevity, if player 1 selects the distribution μ_1, clearly player 2 must select the distribution $\mathrm{argmin}_{\mu_2} f(\mu_1, \mu_2)$. Knowing this, player 1 will select the distribution

$$\mathrm{argmax}_{\mu_1} \min_{\mu_2} f(\mu_1, \mu_2),$$

and the expected score of the game will be

$$\max_{\mu_1} \min_{\mu_2} f(\mu_1, \mu_2).$$

If the selection happens in the other order, of course, the result will be

$$\min_{\mu_2} \max_{\mu_1} f(\mu_1, \mu_2).$$

Clearly choosing *knowing* your opponent's (mixed) strategy can not be a disadvantage compared to choosing with no information, and so we have

$$\min_{\mu_2} \max_{\mu_1} f(\mu_1, \mu_2) \leqslant \max_{\mu_1} \min_{\mu_2} f(\mu_1, \mu_2).$$

Von Neumann's great *minimax* theorem is that these values agree, and this implies that by choosing as if your opponent would know your (mixed) strategy, you obtain a result as good as you could have obtained if you knew your opponents' strategy—hence neither player *could* improve their outcome, even if they had the advantage of greater information.

This is a strong argument for the optimality of such decisions. The key assumption is that the players are perfectly opposed, that is, player 2 seeks to minimize precisely the value that player 1 seeks to maximize. Nash in [Nas51] extended the theory to the more general class of games where players simply each have their own utility function, although it should be noted that Nash merely proved the *existence* of equilibria—that is, strategy sets where no player can improve their situation by switching strategies. It is computationally intractable [DGP09] to identify Nash equilibria in an arbitrary game, and so to apply the theory we're forced to analyze each game in an ad hoc way.

We will not go into a serious review of game theory here, we mention the above details mainly to contextualize our future definitions. For a modern reference on game theory see eg [Mas+20], or [Osb04] for a textbook treatment. Although we won't go into a formal comparison, the traditional counterpart to the external choice considered here is *extensive form games*, see eg [Kuh53; OR94]. For more on the algorithmics of game theory see [Nis+07; SL09].

In his thesis [Hed16], Hedges introduced a novel approach to game theory which he named *compositional game theory,* studying objects called *open games*. The idea of open games is to describe a type of "partial" game which has an interface to the world—some part of the payoff function being dependent on an undetermined environment, into which hole can be inserted another game. This is the sense in which they are *open*. Based on these, one can build a complex game up out of simpler subparts, as well as leverage a string diagrammatical syntax to analyze games.

The original definition of open game cannot help but seem somewhat ad hoc. It was quickly realized that a large part of the definition can be understood to say that a game is a function $\Sigma \to \mathsf{Lens}(\binom{S}{X}, \binom{R}{Y})$ from a strategy set into a set of lenses, and this part of the game composes by lens composition.

The additional data of an open game is a so-called *equilibrium relation*, which determines which strategies are equilibria in a given situation (each strategy really represents a set of strategies, one for each player). In [Cap+22], the author, Capucci, Gavranovic and Hedges demonstrated that this can be further simplified by realizing an open game as a parametrized map in lenses—a morphism in $\mathsf{Para}(\mathsf{Lens}(\mathcal{C}))$ (in fact, one often wants to consider $\mathsf{Optic}(\mathcal{C})$ for, for example, Markov categories $\mathcal{C}$, to account for mixed strategies) along with a notion of equilibrium relation defined on the parameter object $\binom{\Omega}{\Sigma}$. We will begin this chapter by recapping this idea.

On a conceptual level, there is a natural operation on games, called *external choice,* which assigns to two open games $\binom{A_1}{X_1} \to \binom{B_1}{Y_1}, \binom{A_2}{X_2} \to \binom{B_2}{Y_2}$ a new game, whose environment begins by making a choice between one of these games, lets that one happen, then provides some payoff for it at the end. The natural type for the interface of this operation is $\binom{A_1}{X_1} + \binom{A_2}{X_2} \to \binom{B_1}{Y_1} + \binom{B_2}{Y_2}$, but unfortunately the category of optics doesn't have coproducts. The category of *lenses* can be extended with coproducts (into dependent lenses), but the analysis of mixed strategies makes probability an absolute necessity for a useful theory of games.

The introduction of Markov fibrations provides a solution to this problem, and in the second half of this chapter, we provide such an external choice operation on open games. Although the approach based on Markov fibrations is novel, the approach to defining the external choice operator is otherwise very similar to one appearing in presently unpublished work by the author, Braithwaite, Hedges and Videla (this paper used a more specialized approach to adding coproducts to Optic).

5.2 Open games in Monoidal categories

Definition 5.2.1. Let $\mathcal{C}$ be a monoidal category. A *selection relation* on an object $A \in \mathcal{C}$ is a relation $\epsilon \subseteq \mathcal{C}(I, A) \times \mathcal{C}(A, I)$. We write $\epsilon(a, k)$ for the statement $(a, k) \in \epsilon$.

Selection relations are ordered by inclusion, and so form a (posetal) category, which we denote $\mathbb{S}_{\mathcal{C}}(A)$. Given $f : A \to B$, we define the pushforward on selection relations by $f_*\epsilon = \{(fa, k) \mid (a, kf) \in \epsilon\}$. In other words, $f_*\epsilon(b, k)$ if and only if there exists $a : I \to A$ so that $fa = b$ and $\epsilon(a, kf)$.

It is clear that pushforward is monotone, so that this defines a functor $\mathbb{S}_{\mathcal{C}} : \mathcal{C} \to \mathsf{Cat}$

Remark 5.2.2. Of course, there is an equally good "pull-back" operation on selection relations given by $\{(x, kf) \mid (fx, k) \in \epsilon\}$ (this is the pushforward in $\mathbb{S}_{\mathcal{C}^{op}}$). These are adjoint representatives of the same profunctor, which should arguably be regarded as the primary object of interest—that is, we could work with a functor $\mathcal{C} \to \mathbb{C}\mathsf{at}$, the category of categories and profunctors, where we say two selection relations $\epsilon \in \mathbb{S}_X$, $\epsilon' \in \mathbb{S}_Y$ are related by $f : X \to Y$ if $\epsilon(x, kf) \Rightarrow \epsilon'(fx, k)$.

However, we will stick with the pushforward definition for now, since it is conceptually simpler and good enough for our purposes.

Example 5.2.3.

1. Let $\mathcal{C} = \mathsf{Optic}(\mathsf{Set})$. There is a selection relation $\mathrm{argmax}_X \in \mathbb{S}(\binom{\mathbb{R}}{X})$ defined by $\mathrm{argmax}(x, f)$ if and only if x is a maximum of f. (Identifying maps $I \to \binom{\mathbb{R}}{X}$ with points $x \in X$, and maps $\binom{\mathbb{R}}{X} \to I$ with functions $X \to \mathbb{R}$)
2. In the same category, for each $r \in \mathbb{R}$, there is a selection relation given by $\epsilon(x, f) \Leftrightarrow f(x) \geqslant r$. This corresponds to *satisficing* at the value r [Sim56]—that is, selecting any strategy which achieves this value or greater.
3. For $\mathcal{C} = \mathsf{Optic}(\mathcal{C}')$, with $\mathcal{C}'$ semiCartesian, there is a selection relation $\epsilon(x, k) \Leftrightarrow kx = x$ (using the same identification as above). The agents with this selection relation are called *predicting agents* in [BHW18]. These agents attempt to predict the value the environment will return to them.
4. For $\mathcal{C} = \mathsf{Bun}^{fop}$, the fiberwise opposite of manifolds and smooth bundles—that is, the category of lenses between smooth manifolds—there is a selection relation on the tangent bundle $\mathsf{T}X$ given by $\epsilon(x, k)$ if and only if $k(x) = 0$—that is, if x is a fixpoint of the dynamical system identified by $k : X \to \mathsf{T}X$.
5. Let $\mathcal{A} \to \mathcal{C}$ be a dynamical systems theory and work in the category of lenses. Suppose the monoidal structure on $\mathcal{A}$ is Cartesian, so that a lens $I \to \binom{A}{X}$ is the same as a map $* \to X$. Note that $I = \mathsf{T}(*)$ has a distinguished section given by the identity. Then there is a selection relation on each object $\mathsf{T}X$ where a lens $I \to \mathsf{T}X$, given by $x : * \to X$, is in equilibrium with respect to a lens $\mathsf{T}X \leftrightarrows I$ if x is a trajectory between those systems—that is, if it is an equilibrium state of the smooth dynamical system $\mathsf{T}X \leftrightarrows I$. This subsumes the two previous examples.

Definition 5.2.4 (The Nash Product)**.** Let $\epsilon \in \mathbb{S}(X)$, $\epsilon' \in \mathbb{S}(Y)$ be selection relations. Their *Nash product* $\epsilon \boxtimes \epsilon' \in \mathbb{S}(X \otimes Y)$ is given by $\{(x \otimes y, k) \mid \epsilon(x, k(1_X \otimes y)), \epsilon'(y, k(x \otimes 1_Y))\}$

Let us unpack this very dense definition. Given a context $k : X \otimes Y \to I$, a strategy $I \to X \otimes Y$ is in equilibrium if:

1. It decomposes as a tensor product of $x : I \to X, y : I \to Y$.
2. Composing k with y, we obtain a map $X \to I$. This is the context of the first player assuming the second player plays y. x must be an equilibrium strategy for this map. Simultaneously, y must be an equilibrium for the analogous composite of k and x.

Example 5.2.5. Consider the selection relation $\mathrm{argmax}_X \boxtimes \mathrm{argmax}_Y$ on $\binom{\mathbb{R}\times\mathbb{R}}{X\times Y}$. A point (x, y) is an equilibrium for a function $k : X \times Y \to \mathbb{R}^2$ if and only if x maximizes $k_1(-, y)$ and y maximizes $k_2(x, -)$. In other words, if x, y is a Nash equilibrium [Nas51] in the usual sense for the game with payoff matrix k.

Proposition 5.2.6. *The Nash product $\mathbb{S}(X) \times \mathbb{S}(Y) \to \mathbb{S}(X \otimes Y)$, along with the map $* \to \mathbb{S}(I)$ given by the full set $S(I) \times S(I)$ equips $\mathbb{S}$ with a lax monoidal structure.*

Proof. It is trivial to verify associativity and unitality of the monoidal structure. The only hard part is to verify that $\boxtimes$ is actually a natural transformation. To that end, let $f : X \to Y$ and $g : X' \to Y'$ be morphisms of $\mathcal{C}$. Let $\epsilon \in \mathbb{S}(X), \epsilon' \in \mathbb{S}(Y)$.

First consider the statement $(f \otimes g)_*(\epsilon \boxtimes \epsilon')(s, k)$. This holds if and only if $s : I \to X \otimes Y$ factors as $(f \otimes g)s'$ so that $(\epsilon \boxtimes \epsilon')(s', k(f \otimes g))$. This in turn means $s' = x \otimes x'$ and $\epsilon(x, k(f \otimes f')(1_X \otimes x'))$ and analogously for ϵ'.

Now consider the statement $(f_*\epsilon \boxtimes g_*\epsilon')(s, k)$. This means that s factors as $y \otimes y'$ so that $f_*\epsilon(y, k(1_Y \otimes y'))$ and analogously for y', which in turn means that y factors as fx so that $\epsilon(x, k(1_Y \otimes y')f)$. Using the axioms of a monoidal category, it is straightforward to see that these two requirements on x, x' are equivalent, hence we have our naturality. □

Corollary 5.2.7. *Since $\mathbb{S}$ is lax monoidal, its Grothendieck construction $\int \mathbb{S}_{\mathcal{C}}$ acquires a monoidal structure (see [MV20]). We write this category $\mathcal{C}_{\mathbb{S}}$—explicitly, it is given as follows:*

1. *The objects are pairs (X, ϵ) where $X \in \mathcal{C}$ and $\epsilon \in \mathbb{S}(X)$ is a selection relation on it*
2. *The morphisms are morphisms $f : X \to Y$ so that, for every $x : I \to X, k : Y \to I$, $\epsilon(x, kf) \Rightarrow \epsilon(fx, k)$*
3. *The monoidal structure is given by $(X, \epsilon) \otimes (Y, \epsilon') = (X \otimes Y, \epsilon \boxtimes \epsilon')$*

Now we are ready to make the slick definition of open games.

Definition 5.2.8. Let $\mathcal{C}$ be a semiCartesian symmetric monoidal category. The symmetric monoidal double category of *open games in* $\mathcal{C}$ is $\widetilde{\mathsf{Game}}(\mathcal{C}) = \mathbb{P}\mathsf{ara}_{\mathsf{Optic}(\mathcal{C})_{\mathbb{S}}}(\mathsf{Optic}(\mathcal{C}))$.

Note that in this case, $\mathsf{Optic}(\mathcal{C})(I, \binom{A}{X}) = \mathcal{C}(I, X)$, and $\mathsf{Optic}(\mathcal{C})(\binom{A}{X}, I) = \mathcal{C}(X, A)$. Thus a selection function decides, for each *payoff function* $X \to A$, which of the states $I \to X$ are suitable equilibria.

An open game $\binom{A}{X} \to \binom{B}{Y}$ involves an agent or set of agents choosing some element of the space Y—this is their move (or tuple of moves). This move may be contingent on an element of the space X, describing the information the player(s) see before they make their move. The space B, the elements of which Hedges called "utilities", is the type of feedback received by the players from their environment—the outcome they are optimizing over or otherwise trying to control. Finally the space A, the "coutilities", is the feedback that they provide to other players in their causal past, who they are acting as the environment of.

Before we proceed to the case of stochastic lenses, we will pause briefly to make a small modification to the preceding theory as presented in [Cap+22]. In a game with forwards play function $\Sigma \times X \to Y$, there are two ways to talk about the player's "choice"—we may say that the player chooses a strategy $\sigma \in \Sigma$, which then has some effect. Or we may say that the choice is really the $y \in Y$, and the strategy σ is the "precommitment" of choosing what to do given each possible $x \in X$.

Once we introduce randomness—working in $\mathrm{Kl}(\Delta)$, for example—we see that there are two *distinct* ways for a player to make a random choice: first, his strategy $I \to \Delta(\Sigma)$ may be stochastic—that is, he is choosing a random strategy. Or the morphism $\Sigma \times X \to \Delta(Y)$ may be stochastic—this means each strategy σ contains a specification of how to randomly choose y given each possible x.

In the case where $X = *$, the distinction is between taking $\Sigma = Y$ and letting the play function be the identity, and taking $\Sigma = \Delta(Y)$ and letting the play function be the *sampling map* which stochastically draws an element from a distribution.

We take the view that the latter is the proper presentation of this game—this goes along with the terminology in the classical game theory literature, which would certainly regard a distribution on the set of possible moves as a (mixed) strategy. Having represented our games like this, we may restrict ourselves to considering *deterministic* maps $I \to \Sigma$ as strategies. This also fixes the awkwardness in the definition of the Nash product, since now every strategy in $\Sigma_1 \otimes \Sigma_2$ decomposes uniquely as a pair of strategies.

We quickly modify the preceding definitions to make sense of this. Note that we also modify the reparametrization maps to be deterministic (in the base).

Definition 5.2.9 (Open games in a stochastic module)**.** Let $\mathcal{D}$ be a symmetric monoidal stochastic module fibration over the Markov category $\mathcal{C}$. Recall that $\mathsf{SLens}(\mathcal{D})$ acquires a symmetric monoidal structure. Denote as usual $\mathsf{SLens}(\mathcal{D})|_{\mathrm{det}} = \mathsf{SLens}(\mathcal{D}) \times_{\mathcal{C}} \mathcal{C}_{\mathrm{det}}$. Note that this is stable under the monoidal product, and acts on $\mathsf{SLens}(\mathcal{D})$ via the inclusion.

Then the category of *open games in* $\mathcal{D}$ is the category

$$\widetilde{\mathsf{Game}}(\mathcal{D}) = \mathsf{Para}_{(\mathsf{SLens}(\mathcal{D})|_{\mathrm{det}})_{\mathbb{S}}}(\mathsf{SLens}(\mathcal{D})).$$

When $\mathcal{D} \to \mathcal{C}$ is a Markov prefibration, we overload the notation by writing $\widetilde{\mathsf{Game}}(\mathcal{D}) = \mathsf{Game}(\mathcal{D}|_{\mathrm{det}})$.

We now introduce the notion of *strategic equivalence,* which identified two open games if they have the same equilibria for every costate $\overline{\Sigma} \to I$ which can actually occur as a result of pasting the game $\overline{\Sigma} \otimes \bar{X} \to \bar{Y}$ into some larger diagram.

Definition 5.2.10 (Context)**.** Let X, Y be objects of a symmetric monoidal category $\mathcal{C}$. A *context* for X, Y is a tuple $M \in \mathcal{C}, s : I \to X \otimes M, k : Y \otimes M \to I$. We denote the set of contexts $\mathrm{Ctx}(X, Y)$.

Given a morphism $f : P \otimes X \to Y$, and a context $c = (M, s, k)$, the costate $k(f \otimes 1_M)(1_P \otimes s)$ will be called the induced costate.

Remark 5.2.11. The preceding definition clearly works in a non-symmetric monoidal category as well. In this case, arguably, a context should be defined to consist of maps $I \to M_1 \otimes X \otimes M_2, M_1 \otimes Y \otimes M_2 \to I$. However, the problem with this from our point of view is that it doesn't allow the definition of a costate on P given a parametrized map (since one cannot commute the P past the M_1), which is what we're interested in.

It is also worth observing that contexts are essentially the same thing as optics $\binom{I}{I} \to \binom{Y}{X}$. We have not specified an equivalence relation on contexts (we do not need it,) but it is easy to see that for any parametrized map, two contexts that are sliding equivalent give the same induced costate.

Definition 5.2.12 (Strategic Equivalence)**.** Let $\mathcal{D}$ be a monoidal stochastic module fibration over a Markov category $\mathcal{C}$, and consider a 2-cell $\alpha : G_1 \to G_2 : \bar{X} \to \bar{Y} \in \widetilde{\mathsf{Game}}(\mathsf{SLens}(\mathcal{D}))$.

We say this is a *strategic equivalence* if the underlying map $\Sigma_1 \to \Sigma_2$ in $\mathcal{C}$ is an isomorphism, and for every context $c \in \mathrm{Ctx}(\bar{X}, \bar{Y})$, the induced costate k on $\overline{\Sigma_2}$ has the same equilibria (under this isomorphism) as the composite costate $k\alpha$ on $\overline{\Sigma_1}$

Note that a game up to strategic equivalence is determined by a relation between strategies $I \to \Sigma$ and contexts. This brings us closer to Hedges' original definition of open game from [Gha+18a]. The chief difference is that a game in our sense is prevented from "inspecting" the context $Y \to \bar{Y}$ for those $y \in Y$ which are not in the image of $X \times \Sigma \to Y$, in the sense that whether a given strategy is in equilibrium or not cannot depend on this (since we only see a certain costate on Σ).

Proposition 5.2.13. *Strategic equivalence is compatible with composition and tensor in* $\widetilde{\mathsf{Game}}(\mathsf{SLens}(\mathcal{D}))$.

Proof. Let $\alpha : (G_1 \to G_1') : \bar{X} \to \bar{Y}, \beta : (G_2 \to G_2') : \bar{Y} \to \bar{Z}$ be strategic equivalences. By 2-cell composition there is a map $G_2 G_1 \to G_2' G_1'$, which we must show is a strategic equivalence.

Let $c = s, k$ be a context in $\mathrm{Ctx}(\bar{X}, \bar{Z})$. Now a pair of strategies σ_1, σ_2 for G_1, G_2 are in Nash equilibrium for the costate induced by this context if and only if σ_1 is in equilibrium for the costate induced by the context $s : I \to \bar{X} \otimes \bar{M}, k(p_2(\sigma_2) \otimes 1_{\bar{M}})$, where p_2 is the play function of G_2, and the analogous condition holds for σ_2.

But if α, β are equivalences, this is clearly equivalent to asking that σ_1, σ_2 be in Nash equilibrium for the costate induced by c on $\overline{\Sigma_1'} \otimes \overline{\Sigma_2}'$. This concludes the proof. □

Definition 5.2.14. We denote by $\mathsf{Game}(\mathsf{SLens}(\mathcal{D}))$ the symmetric monoidal category of strategic equivalence classes of open games.

Note that this category of open games retains a 2-categorical structure, given by deterministic maps $\Sigma_1 \to \Sigma_2$ which preserve equilibria in every context. However, we will leave a deeper investigation of this structure for future work.

5.3 Open Games with External Choice

Let $\mathcal{C}$ be an extensive Markov category. Let $\mathcal{D}$ be a monoidal stochastic module fibration with Markov structure, which has coproducts which are preserved by the pullbacks. Recall that $\mathsf{SLens}(\mathcal{D})$ acquires two monoidal structures: one from dualizing the given monoidal structure on $\mathcal{D}$, which we simply denote $\otimes, I$, and one from taking the coCartesian monoidal structure in the fiber (which is Cartesian after taking the fiberwise dual, of course), which we denote $\&, \top$. Note that $(\mathsf{SLens}(\mathcal{D}), \&)$ is a Markov category. For the rest of this section, fix $\mathcal{D}, \mathcal{C}$ like this.

Lemma 5.3.1. *Let* $\bar{X}, \bar{Y}$ *be objects in* $\mathsf{SLens}(\mathcal{D})$*, and let* $I \to I + I$ *be a morphism in* $\mathcal{C}$*. Then there is a canonical map* $\bar{X} \& \bar{Y} \to \bar{X} + \bar{Y}$*, so that the underlying map is* $X \otimes Y \to (X \otimes Y) \otimes (I + I) \cong X \otimes Y + X \otimes Y \to X + Y$

Proof. The first map in the factorization has a deterministic retract (deleting the $I + I$ component,) and using the coproduct-preservation, the coproduct over $X + Y$ and $\bar{X}\&\bar{Y}$ pull back to the same object over $X \otimes Y \otimes (I + I)$. Composing the induced stochastic-Cartesian map and the Cartesian map gives the canonical map we wanted. □

With the interpretation that $\overline{X}\&\overline{Y}$ is the object $\overline{X}_x + \overline{Y}_y$ indexed over $X \otimes Y$, this map simply selects one branch randomly and marginalizes to that coordinate in $X \otimes Y$, then includes the returned value into the coproduct.

Definition 5.3.2. When $\mathcal{C}, \mathcal{D}$ as above, $\widetilde{\mathsf{Game}}(\mathsf{SLens}(\mathcal{D}))$ acquires a monoidal structure which we call *external choice,* and write $\oplus$, given on objects by the coproduct $+$ in $\mathsf{SLens}(\mathcal{D})$, and on morphisms by the following formula:

Given two open games $G_1 = (\overline{\Sigma_A} \otimes \overline{A_1} \to \overline{A_2}, \epsilon_A), G_2 = (\overline{\Sigma_B} \otimes \overline{B_1} \to \overline{B_2}, \epsilon_B)$, their external choice is has parameter $\Sigma_A\&\Sigma_B$. The play map is given by

$$(\Sigma_A\&\Sigma_B)\otimes(A_1+B_1) \cong (\Sigma_A\&\Sigma_B)\otimes A_1+(\Sigma_A\&\Sigma_B)\otimes B_1 \to \Sigma_A\otimes A_1+\Sigma_B\otimes B_1 \to A_2+B_2$$

The selection relation $\epsilon_A \oplus \epsilon_A$ is given (up to equivalence) as follows: Given a context $k : \overline{\Sigma_A}\&\overline{\Sigma_B} \to I$, and a deterministic state $I \to \overline{\Sigma_A}\&\overline{\Sigma_B}$, they are in equilibrium if

1. k factors over the canonical $\overline{\Sigma_A}\&\overline{\Sigma_B} \to \overline{\Sigma_A} + \overline{\Sigma_B}$ for some $c : I \to I + I$.
2. The factorization being given by $k_A, k_B : \overline{\Sigma_A}, \overline{\Sigma_B} \to I$,and $I \to \overline{\Sigma_A}\&\overline{\Sigma_B}$ being given by $\sigma_A, \sigma_B : I \to \overline{\Sigma_A}, \overline{\Sigma_B}$, we have $\epsilon_A(\sigma_A, k_A), \epsilon_B(\sigma_B, k_B)$

The idea behind the external choice operator is that, for all the contexts which can actually occur as a result of pasting a game $G \oplus G'$ into a larger string diagram, the map $X \times X' \to \overline{X} + \overline{X}'$ has the given form—that is, the probability of landing in each of the two fibers does not depend on the chosen x, x' and the conditional distributions on the $\overline{X}$ component of the fiber depend only on $x \in X$. Hence we need only concern ourselves with which states are equilibria for contexts of this form. The choice of the empty set of equilibria for other contexts is merely a convention.

Given a state $I \to A + B$ in a Markov category with coproducts, we say a pair $I \to A, I \to B$ form a pair of conditionals if the copairing $I + I \to A + B$ is a Bayesian inverse of the map $A + B \to I + I$.

Given a state $I \to \bar{X} + \bar{Y}$ in $\mathsf{SLens}(\mathcal{D})$, we say a pair of maps $I \to \bar{X}, I \to \bar{Y}$ form a pair of conditionals if the underlying maps do.

Theorem 5.3.3. *As defined above, $\oplus$ is a symmetric monoidal structure on* $\widetilde{\mathsf{Game}}(\mathsf{SLens}(\mathcal{D}))$.

Proof. The monoidal coherences come from the monoidal structure of $+$, and it is trivial to see that they preserve the selection relations. The only nontrivial part is proving that $\oplus$ is functorial. Hence let $G_1 : A_1 \to B_1, G_1' : B_1 \to C_1, G_2 : A_2 \to B_2, G_2' : B_2 \to C_2$ be games. The strategy set of $(G_1' \oplus G_2') \circ (G_1 \oplus G_2)$ is given by $(\Sigma_1\&\Sigma_2) \otimes (\Sigma_1' \otimes \Sigma_2')$ For $(G_1' \circ G_1) \oplus (G_2' \circ G_2)$, by $(\Sigma_1 \otimes \Sigma_1')\&(\Sigma_2 \otimes \Sigma_2')$.

In the base, these are the same object $\Sigma_1 \otimes \Sigma_2 \otimes \Sigma_1' \otimes \Sigma_2'$. In the fiber, they are given respectively by

$$(\overline{\Sigma_1} + \overline{\Sigma_2}) \otimes (\overline{\Sigma_1}' + \overline{\Sigma_2'})$$

and

$$(\overline{\Sigma_1} \otimes \overline{\Sigma_1'}) + (\overline{\Sigma_2} \otimes \overline{\Sigma_2}').$$

Note the coproducts here are the fiberwise ones. There is an obvious lens from the former to the latter (given by the identity map on the base, and the inclusion of two summands in a fourfold coproduct—note that lenses go backwards in the fiber). Letting $I \to (A_1 \oplus A_2) \otimes M, (C_1 \oplus C_2) \otimes M \to I$ be a context, and going through the definitions, it is clear that the resulting contexts for the former game factors as this lens followed by the context for the latter game. In other words, this lens is a reparametrization map between the two games. It suffices to verify it is an equivalence.

Unpacking the equivalence relation on $(G_1' \oplus G_2')(G_1 \oplus G_2)$, note that (in all the possible contexts,) the signal to G_1 does not depend on the action of G_2' and vice versa, and so for G_1' and G_2. Hence they are in equilibrium if and only if they are in equilibrium in $G_1' G_1$ for the given context (conditioned on that branch), and similarly the other two. This proves the desired equivalence. □

Example 5.3.4. Consider the (Grothendieck) fibration $\mathsf{Set}^{\to} \to \mathsf{Set}$, which can be viewed as a Markov fibration. Let $f : \binom{\mathbb{R}}{\Sigma} \otimes \binom{\bar{X}}{X} \leftrightarrows \binom{\bar{Y}}{Y}$ be a parameterized lens. Denote by argmax_f the open game $\binom{\bar{X}}{X} \to \binom{\bar{Y}}{Y}$ with parameters $\binom{\mathbb{R}}{\Sigma}$, underlying parameterized lens f, and equilibrium relation given by argmax.

Then if $g : \binom{\mathbb{R}}{\Sigma'} \otimes \binom{\overline{X'}}{X'} \leftrightarrows \binom{\overline{Y'}}{Y'}$ is another parameterized lens, we have

$$\mathrm{argmax}_f \oplus \mathrm{argmax}_g \cong \mathrm{argmax}_{f \oplus g},$$

where by an abuse of notation $f \oplus g$ denotes the paramterized lens

$$\begin{pmatrix} \mathbb{R} \\ \Sigma \times \Sigma' \end{pmatrix} \otimes (\begin{pmatrix} \overline{X} \\ X \end{pmatrix} \oplus \begin{pmatrix} \overline{X'} \\ X' \end{pmatrix}) \to \begin{pmatrix} \overline{Y} \\ Y \end{pmatrix} \oplus \begin{pmatrix} \overline{Y'} \\ Y' \end{pmatrix}$$

given by distributing into the coproduct, then projecting into the relevant factor of the product $\Sigma \times \Sigma'$ and applying either f or g

A context for either of these games consists of an element of $X + X'$ and a function $k : Y + Y' \to \mathbb{R}$. The external choice game can be seen as having two players, one who gets to play if the context chooses an element in X, who must output an element $y \in Y$ and optimize $k(y)$ (according to his private utility function $\Sigma \times X \times \bar{Y} \to \mathbb{R}$), the other playing when the input is in X' and who must choose an element in y'. The single argmax game can be seen as a single player, who is constrained to play inside the same "branch" of the game as the input (this constraint is encoded in the lens $f \oplus g$), and whose utility function is given by the first players' in the first branch, and the second players' in the second branch.x

Example 5.3.5. Consider the Markov fibration $\mathsf{Kl}(\Delta)^{\to} \to \mathsf{Kl}(\Delta)$. Let $G_1 : \binom{A_1}{X_1} \to \binom{B_1}{Y_1}, G_2 : \binom{A_2}{X_2} \to \binom{B_2}{Y_2}$ be two open games in $\widetilde{\mathsf{Game}}(\mathsf{SLens}(\mathsf{Kl}(\Delta)^{\to}))$, with underlying lenses parameterized by $\binom{\Sigma_1'}{\Sigma_1}, \binom{\Sigma_2'}{\Sigma_2}$. Observe that a context for $G_1 \oplus G_2$ consists of a residual set M, a distribution $x : I \to (X_1 \sqcup X_2) \times M$, and a section of $B_1 \sqcup B_2 \to Y_1 \sqcup Y_2$ parameterized by M. This latter can be identified with two sections $k_1 : M \times Y_1 \to B_1, k_2 : M \times Y_2 \to B_2$. The former can be written as two conditional distributions $x_1 : I \to X_1 \times M, x_2 : I \to X_2 \times M$, composed with the composite $I \to (X_1 \sqcup X_2) \times M \to I + I$.

The equilibrium relation for the external choice $G_1 \oplus G_2$ says that a pair of strategies (σ_1, σ_2) is in equilibrium in the given context if they are both in equilibrium of the given conditional context (that is, σ_1 is equilibrium for the context consisting of x_1, k_1, and analogously for σ_2).

Example 5.3.6. Let $G : \binom{*}{X} \longrightarrow \binom{R}{Y}$ be a game, representing an agent who is optimizing the return value $r \in R$ in some sense. Then $G \otimes (1_I \oplus 1_I = 1_{I+I})$ represents the same agent, whose payoff may now depend on an additional bit (a value in $I + I$), but whose decisions may not depend on that bit (his selection function may still depend on its distribution). On the other hand, $G \oplus G (\cong G \otimes 1_I \oplus G \otimes 1_I)$ represents the same situation, but where the player's strategy *may* depend on the bit—he provides two strategies σ_1, σ_2, one for each possibility. This proves that $\otimes$ does not distribute over $\oplus$

Chapter 6

Categories of stochastic dynamical systems

6.1 Introduction

Given a set of inputs A and a set of outputs B, there are essentially two ways to make sense of the informal description "finite-state automaton which reads inputs from A and produces outputs in B". These are the notions of *Mealy machine* and *Moore machine*. Simply put, if the set of states is S, a mealy machine is a function $A \times S \to B \times S$, whereas a Moore machine is a pair $A \times S \to S, S \to B$. In other words, in a Moore machine the output does not depend on the current input, but only on previous inputs (through their effect on the state), but in a Mealy machine, the input can be passed through directly.

Using the language of categorical systems theory, we can make the following definitions:

Definition 6.1.1 (Moore Machine)**.** In a dynamical systems theory, a *Moore machine* with state space S and interface A is a lens $TS \leftrightarrows A$. The category of Moore machines with interface A is the slice category of the functor T over A—that is, a morphism of Moore machines is a morphism $S \to S'$ so that the obvious triangle commutes.

Definition 6.1.2 (Mealy Machine)**.** In a dynamical systems theory, a *Mealy machine* with state space S and interface A consists of a costate lens $TS \otimes A \leftrightarrows I$. The category of Mealy machines with interface A is the comma category of the functor $T(-) \otimes A$ over I—that is, a morphism of Mealy machines is $S \to S'$ so that the obvious triangle commutes.

A Moore machine in the sense of Definition 6.1.1 is what Myers calls an (open) dynamical system, and they are the central object of study in [Mye23]. Arguably, both Moore machines and Mealy machines deserve the name of "open dynamical system"—the difference is how they interact with the external world.

Observe in particular that, if $A = I$, the categories of Mealy and Moore machines agree, both being equal to the slice of T over the unit, I. In other words, the two notions of *closed* dynamical system coincide.

It is clear that both Mealy and Moore machines, in this sense, are special kinds of *parametrized morphism* in lenses, namely those parametrized by an object of the form

TS. This leads naturally to the idea that there should be a triple category of morphisms of this type, charts, and lenses.

We call the parametrized lenses $\mathsf{TS} \otimes \bar{\mathsf{A}} \leftrightarrows \bar{\mathsf{B}}$ *bisystems* since they generalize the two types of machine, Moore and Mealy (but we choose to stick with "system" rather than "machine").

Example 6.1.3 (Machine Learning as a bisystem)**.** It is an observation which goes back at least to [FST19], and was more thoroughly developed in [Cru+21], that the process of training a machine learning algorithm by gradient descent can be abstracted in the following way:

- Wanting to learn a function $X \to Y$, we choose a parametrized function $f : X \times P \to Y$, where X, Y, P are, in the simplest case, Euclidean spaces $\mathbb{R}^k$.
- We take the backwards derivative of f, obtaining a *lens:*

$$(f, Df) : \begin{pmatrix} TX \\ X \end{pmatrix} \otimes \begin{pmatrix} TP \\ P \end{pmatrix} \leftrightarrows \begin{pmatrix} TY \\ Y \end{pmatrix}.$$

- For each datum (x_n, y_n), we compute the loss gradient $\nabla L(-, y_n) : Y \to TY$, and combining this with $x_n \in X$ and the current parameter p, we get a gradient on the parameter space which we can use to update.

Thus a machine learning algorithm is a sort of bisystem. Indeed our bisystems are essentially an abstracted version of the *learners* of Fong–Spivak–Tuyeras. The functoriality of this assignment is the main point of the above-mentioned papers.

In this chapter, we will put together the ingredients we have assembled so far and construct a triple category of stochastic dynamical systems. We will also give the construction of triple categories of systems in the ordinary case.

Note that the theory of Markov fibrations does not quite generalize the ordinary theory of fibrations—it only generalizes fibrations with a Cartesian base (Example 3.2.5 and Proposition 3.4.15). Since the pullbacks in $\mathcal{C}_{\mathrm{det}}$ play such a key role in the theory, it is unclear whether this can be ignored. Although most examples from categorical systems theory do involve a base cartegory which is Cartesian, this is not a requirement for the theory. Moreover, we will see that the construction of the double category of lenses and charts encounters certain problems for a general Markov fibration not seen for ordinary fibrations. Hence we will give a separate description of the triple categories in each of the two cases.

The notions of Mealy and Moore machine are both quite old, going back to [Mea55; Moo56]. While we do obtain finite-state automata of these types as special cases, our interest is primarily in the analysis of dynamical systems, which tends to ask rather different questions than automata theory. Thus, despite using the terminology, we will not be particularly interested in the actual theory of Mealy and Moore machines. We do mention one recent paper, [Boc+23], which has a category-theoretic approach similar in spirit to our own. Their category of Mealy machines can be obtained, not as our category of Mealy machines above, but by considering maps $\mathsf{TS} \otimes \mathsf{A} \to \mathsf{B}$ in the case where $\mathsf{A} = \binom{*}{A}, \mathsf{B} = \binom{*}{B}$ are trivial in the secondary component. In the discrete case, such a map is given by $A \times S \to B \times S$.

It should also be noted that the idea of embedding Mealy machines as the morphisms $\mathsf{TS} \otimes \mathsf{A} \to \mathrm{I}$ is not original, but was communicated to the author by Matteo Capucci. It seems not to have appeared in the literature so far. The idea that there "should" be a triple category of systems, like the one we will construct, has also circulated as folklore, although again an explicit construction has yet to appear. Very similar ideas appear in the work of Shapiro and Spivak, see for example [SS23].

We begin this chapter with a treatment of double categories of charts and lenses in the context of Markov fibrations (and stochastic modules). The construction does not work quite as well as in the classical case—the difficulty is essentially that the equivalence relation which defines stochastic charts has a directed nature, and given a 2-cell defined in an obvious way $\phi \to \psi$, and an equation $\psi \overset{\sim}{\leftarrow} \psi'$, there is not (apparently) in general a way to lift this to an arrow $\phi' \to \psi'$ (with $\phi' \simeq \phi$). However, we can construct a double category whose globular horizontal 2-category has *connected components* given by the stochastic charts (or lenses).

We proceed to give an account of categorical systems theory for these double categories. The chief problem posed by the above is that we may not have any good clock systems (Example 2.6.11). Two equivalent charts should represent the same system, but they may receive different sets of maps from the supposed clock system (and thus have different sets of trajectories). We resolve this by proving that, for lenses with a deterministic base, every equivalence class of lenses has an initial representative. Moreover, mapping out of this initial representative to a representative of some other lens, a 2-cell exists filling a given square if and only if it commutes as a map of lenses and charts in the classical case (recall that over deterministic bases, Markov fibrations become ordinary fibrations). In particular the trajectories of a system with respect to such a clock system depend only on their equivalence class.

We follow this up by constructing the above discussed triple categories of "bimachines", that is systems which combine Mealy and Moore machines. We do this both for stochastic and the ordinary case—the procedure is exactly the same, but of course they are different objects, neither generalizing the other.

We end by constructing a stochastic dynamical systems theory for *smooth* dynamical systems—requiring a brief detour to construct a suitable Markov category of smooth kernels.

6.2 Double categories of stochastic charts and lenses

To construct the double category $\mathbb{A}\mathsf{rena}(\mathcal{A})$ of charts and lenses for an ordinary fibration $\mathcal{A} \to \mathcal{C}$, one can use the following procedure:

1. Form the square double category $\mathcal{A}^{\to} \rightrightarrows \mathcal{A}$
2. Take the fiberwise opposite of these objects: $(\mathcal{A}^{\to})^{\mathrm{fop}} \rightrightarrows \mathcal{A}^{\mathrm{fop}}$.
3. Observe that fiberwise opposite preserves pullbacks, and hence this is again a double category.

Here we used the following result: If $\mathcal{D} \to \mathcal{C}$ is a Grothendieck fibration and $\mathcal{A}$ is any category, then $\mathcal{D}^{\mathcal{A}} \to \mathcal{C}^{\mathcal{A}}$ is again a fibration (which classifies the lax limits of the composite $\mathcal{A}^{\mathrm{op}} \to \mathcal{C}^{\mathrm{op}} \to \mathsf{Cat}$), and a natural transformation is Cartesian iff it is levelwise Cartesian. It would be neat to obtain a similar result for Markov fibrations.

The first problem with this is that $\mathcal{C}^{\mathcal{A}}$ does not generally inherit a Markov structure from $\mathcal{C}$. As we noted when we introduced diagram Markov categories (Example 2.4.5), one has to consider the category $\mathsf{Fun}(\mathcal{A}, \mathcal{C})$ of deterministic diagrams instead.

First, we will obtain the result sketched above for Markov *prefibrations*—that is, if $\mathcal{D} \to \mathcal{C}$ is a prefibration, there is a natural prefibration over $\mathsf{Fun}(\mathcal{A}, \mathcal{C})$ which has

$\mathsf{Fun}(\mathcal{A}, \mathcal{D})$ as its underlying fibration. This implies that $(-)^{\mathcal{A}}$ lifts from fibrations to stochastic modules.

Proposition 6.2.1. *Let $\mathcal{D} \to \mathcal{C}$ be a Markov prefibration. Recall that by $\mathcal{C}^{\to}$ we denote the category of* deterministic *arrows in $\mathcal{C}$. Let $\mathsf{Ar}(\mathcal{C})$ denote the ordinary arrow category. Let now $\mathcal{D}^{\to}$ denote the category $\mathsf{Ar}(\mathcal{D}) \times_{\mathsf{Ar}(\mathcal{C})} \mathcal{C}^{\to}$ consisting of those arrows in $\mathcal{D}$ which lie over a deterministic base (but again, where the morphisms consist of commutative squares whose other sides do not necessarily have deterministic bases). Then $\mathcal{D}^{\to} \to \mathcal{C}^{\to}$ is a Markov prefibration.*

This yields a functor $\mathsf{MarkPreFib}(\mathcal{C}) \to \mathsf{MarkPreFib}(\mathcal{C}^{\to})$, so that $(\mathcal{D}^{\to})|_{\mathrm{det}} = (\mathcal{D}|_{\mathrm{det}})^{\to}$. This equation induces a natural transformation $\overline{(\mathcal{D}_0^{\to})}|_{\mathrm{det}} \to (\overline{\mathcal{D}_0}|_{\mathrm{det}})^{\to}$, which in turns gives a lift of $(-)^{\to}$ to the category of stochastic module fibrations, where the induced algebra structure acts pointwise.

Proof. Noting that $\mathcal{C}^{\to}$ is a Markov category with the "pointwise" structure, and the deterministic maps consist precisely of the pointwise deterministic maps, clearly $\mathcal{D}^{\to}|_{\mathrm{det}} = \mathcal{D}|_{\mathrm{det}}^{\to}$, and fibrations are stable under the formation of arrow categories, with Cartesian maps formed pointwise.

It is *not* trivial that this is a prefibration, because given a triangle in $\mathcal{C}^{\to}_{\mathrm{det}}$—a "prism"—and a Cartesian lifting, we only know that the maps "at the ends" are Cartesian, not the maps between the triangles.

Therefore we can not immediately apply the unique lifting property to say that the square between the induced lifts $\bar{Y}_{iX_i} \to \bar{Y}_{iZ_i}$ commute, given some map $\bar{Y}_0 \to \bar{Y}_1$ over $Y_0 \to Y_1$. However, by taking the pullback on both sides (and noting that pullbacks are functorial), we can factor this square into two which live entirely over a deterministic base, and where the Cartesian property therefore imply commutativity.

We have already argued that this commutes with restriction to the deterministic part. The natural transformation is induced for completely abstract reasons, by applying $(-)^{\to}$ to the unit to obtain a map $\mathcal{D}_0^{\to} \to (\overline{\mathcal{D}_0}|_{\mathrm{det}})^{\to} = (\overline{\mathcal{D}_0^{\to}})|_{\mathrm{det}}$, which by the universal property of $\overline{(-)}$ induces the desired map $\overline{(\mathcal{D}_0^{\to})}|_{\mathrm{det}} \to (\overline{\mathcal{D}_0}|_{\mathrm{det}})^{\to}$. □

Corollary 6.2.2. *Let $\mathcal{A}$ be a small category, and let $\mathcal{D} \to \mathcal{C}$ be a Markov prefibration. Let $\mathsf{Fun}(\mathcal{A}, \mathcal{D}) := \mathcal{D}^{\mathcal{A}} \times_{\mathcal{C}^{\mathcal{A}}} \mathsf{Fun}(\mathcal{A}, \mathcal{C})$. Then the functor $\mathsf{Fun}(\mathcal{A}, \mathcal{D}) \to \mathsf{Fun}(\mathcal{A}, \mathcal{C})$ is a Markov prefibration.*

Proof. Note that $\mathsf{Fun}(\mathcal{A}, \mathcal{D})$ is a limit of the categories $\mathcal{D}^{\to}$ and prefibrations are stable under these limits. □

The question is now

1. If $\mathcal{D}$ is a Markov fibration, we get a stochastic module structure on $\mathcal{D}|_{\mathrm{det}}^{\mathcal{A}}$—does it present a Markov fibration?
2. There is an induced map $\mathsf{SChart}(\mathsf{Fun}(\mathcal{A}, \mathcal{D}|_{\mathrm{det}})) \to \mathsf{Fun}(\mathcal{A}, \mathcal{D})$ (where the latter is taken by convention to mean the full subcategory of functors whose image in $\mathcal{C}$ consists of deterministic arrows). Is this an isomorphism? (If it is, clearly this implies point 1)

Unfortunately it's not clear that either of these are true—the surjectivity of $\mathsf{SChart}(\mathcal{D}|_{\mathrm{det}}) \to \mathcal{D}$ cannot a priori be lifted to the arrow category. The issue is that, given a map in $\mathcal{D}^{\to}$ consisting of, say ϕ_0, ϕ_1, it is not sufficient to find charts representing each of these—we must find a *chart of squares* representing the square.

This is not guaranteed by the Markov fibration structure, and a similar issue comes into play for the equivalence witnesses.

We may attempt to ignore this issue and simply try to form a double category $\mathsf{SLens}(\mathcal{D}^{\to}) \rightrightarrows \mathsf{SLens}(\mathcal{D})$, given a stochastic module $\mathcal{D}$, but here the problem is that SLens does not commute with limits in general. Hence we can not easily define a composition on the 2-cells of lenses obtained this way.

There are various ways we might attempt to remedy this problem. One approach would be to formulate a behavioural notion of "commutativity" for squares of stochastic lenses and charts, but the problem with this is that it is not obvious whether this property is stable under composition.

The basic problem stems from the fact that chart equivalences have a "directed" nature, and given a morphism of precharts $(M, \phi) \to (N, \psi)$ and an equivalence $(N, \psi) \leftarrow (N', \psi')$ (for example given by a stochastic section $N' \to N$ satisfying suitable conditions), there is not in general a way to lift this back into an equivalent (M', ϕ') with a map to N'.

This observation leads to the idea that we might define a double category of precharts which has the directed equivalences among its morphisms (going only in one direction). We will begin by constructing this double category.

Definition 6.2.3. Let $\mathcal{D} \to \mathcal{C}$ be a (Grothendieck) fibration, and let $\mathcal{C} \to \mathcal{C}'$ be a faithful, identity-on-objects functor. Suppose $\mathcal{C}$ admits pullbacks, and given a pair of morphisms $P \to X, Y \in \mathcal{C}'$ over Z, where $P \to X$ is in $\mathcal{C}$, there is a unique common factorization $P \to X \times_Z Y$.

1. The double category $\mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{C})$ has $\mathcal{C}$ as the vertical category, spans $X \leftarrow P \to Y$ in $\mathcal{C}$ equipped with a section $X \to P \in \mathcal{C}'$ as horizontal cells, and maps of spans which commute with the sections as 2-cells.

2. The double category $\mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{D}/\mathcal{C})$ lying over $\mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{C})$ has $\mathcal{D}$ as the vertical category, and spans $\bar{X} \xleftarrow{f} \bar{P} \xrightarrow{g} \bar{Y}$ where f is Cartesian, decorated with a section $X \to P$ in $\mathcal{C}'$ as the horizontal cells, and maps of such spans (so that the underlying thing commutes with the sections) as the 2-cells. We will call the horizontal cells *decorated spans*.

3. There is an apparent forgetful functor $\mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{D}/\mathcal{C}) \to \mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{C})$

To spell it out, a 2-cell in $\mathbb{S}\mathsf{pan}_{\mathcal{C}'}(\mathcal{D}/\mathcal{C})$ consists of a diagram of this form in $\mathcal{D}$, where f_1, f_2 are Cartesian, and (writing X_1 for the object underlying $\overline{X_1}$, and so on) two sections s_1, s_2 of the underlying maps in $\mathcal{C}'$, so that the second diagram also commutes in $\mathcal{C}'$:

$$\begin{array}{ccccc} \overline{X_1} & \xleftarrow{f_1} & \overline{P_1} & \longrightarrow & \overline{Y_1} \\ \downarrow & & \downarrow & & \downarrow \\ \overline{X_2} & \xleftarrow{f_2} & \overline{P_2} & \longrightarrow & \overline{Y_2} \end{array}$$

$$\begin{array}{ccccc} X_1 & \xrightarrow{s_1} & P_1 & \longrightarrow & Y_1 \\ \downarrow & \xrightarrow{s_2} & \downarrow & & \downarrow \\ X_2 & & P_2 & \longrightarrow & Y_2 \end{array}$$

Naturally, we are interested in the case of $\mathbb{S}\mathsf{pan}_{\mathcal{C}}(\mathcal{D}/\mathcal{C}_{\text{det}})$ for a stochastic module $\mathcal{D}$. Then the decorated spans are representatives for stochastic charts. We will start by introducing a loosed notion of 2-cell for these spans, which combines the directed equivalences with the ordinary deterministic 2-cells of spans.

Definition 6.2.4. Let $\mathcal{C}$ be a Markov category and let $\mathcal{D}$ be a stochastic module over $\mathcal{C}$. Let $(M_1, \phi_1) : \bar{X}_1 \leftrightarrows \bar{Y}_1, (M_2, \phi_2) : \bar{X}_2 \leftrightarrows \bar{Y}_2$ be decorated spans in $\mathbb{S}\mathsf{pan}_{\mathcal{C}}(\mathcal{D}/\mathcal{C}_{\text{det}})$, and let $f : \bar{X}_1 \rightrightarrows \bar{X}_2, g : \bar{Y}_1 \to \bar{Y}_2$ be morphisms in $\mathcal{D}$, so that we have a square

$$\begin{array}{ccccc}
\overline{X_1} & \longleftarrow & \overline{X_1}_{M_1} & \xrightarrow{\phi_1} & \overline{Y_1} \\
{\scriptstyle f}\big\downarrow & & & & \big\downarrow{\scriptstyle g} \\
\overline{X_2} & \longleftarrow & \overline{X_2}_{M_2} & \xrightarrow{\phi_2} & \overline{Y_2}
\end{array}$$

A *2-cell of decorated spans* for this data consists of a morphism $m : M_1 \to M_2 \in \mathcal{C}$ (that is, possibly stochastic), satisfying the following two conditions. First, the diagram

$$\begin{array}{ccccc}
X_1 & \longleftarrow & M_1 & \longrightarrow & Y_1 \\
\big\downarrow & & \big\downarrow & & \big\downarrow \\
X_2 & \longleftarrow & M_2 & \longrightarrow & Y_2
\end{array}$$

in the base must commute. Given this, there is an induced square

$$\begin{array}{ccc}
\overline{X_1}_{M_1} & \longrightarrow & \overline{Y_1}_{M_1} \\
\big\downarrow & & \big\downarrow \\
\overline{X_2}_{M_1} & \longrightarrow & \overline{Y_2}_{M_1}
\end{array}$$

in $\mathcal{D}_{M_1}$, where the bottom map is given by pulling back ϕ_2 along m, in the sense of Lemma 3.3.13. The second condition is that this square must also commute.

Lemma 6.2.5. *2-cells of decorated spans compose—that is, given a diagram*

$$\begin{array}{ccc}
\bar{X}_1 & \xrightarrow{\phi_1} & \bar{Y}_1 \\
\big\downarrow & & \big\downarrow \\
\bar{X}_2 & \xrightarrow{\phi_2} & \bar{Y}_2 \\
\big\downarrow & & \big\downarrow \\
\bar{X}_3 & \xrightarrow{\phi_3} & \bar{Y}_3
\end{array}$$

where the horizontal maps are decorated spans, and the vertical maps are maps in $\mathcal{D}$ with deterministic base, and given maps of decorated spans $\phi_1 \xrightarrow{\alpha} \phi_2 \xrightarrow{\beta} \phi_3$, there is a map of decorated spans $\phi_1 \to \phi_3$

Proof. We can easily compose the two maps to get $M_1 \to M_3$. Now, the question is whether the perimeter of this diagram commutes:

$$\begin{array}{ccc}
\bar{X}_{1M_1} & \longrightarrow & \bar{Y}_{1M_1} \\
\downarrow & & \downarrow \\
\bar{X}_{2M_1} & \longrightarrow & \bar{Y}_{2M_1} \\
\downarrow & & \downarrow \\
\bar{X}_{3M_1} & \longrightarrow & \bar{Y}_{3M_1}
\end{array}$$

Note that the top square is commutative by assumption, since pullbacks compose (even along stochastic maps) and this is the assumption that α is a cell. The bottom square is the result of pulling back a commutative square over M_2 again along α. Note that pullback along stochastic morphisms is not in general functorial—but since the vertical parts of this square are themselves pulled back from X_2, Y_2, this composition is preserved by pullback along $\alpha : M_1 \to M_2$. This finishes the proof. □

Proposition 6.2.6. *2-cells of decorated spans compose horizontally: Given a square*

$$\begin{array}{ccccc}
\bar{X}_1 & \xrightarrow{\phi_1} & \bar{Y}_1 & \xrightarrow{\psi_1} & \bar{Z}_1 \\
\downarrow & & \downarrow & & \downarrow \\
\bar{X}_2 & \xrightarrow{\phi_2} & \bar{Y}_2 & \xrightarrow{\psi_2} & \bar{Z}_2
\end{array}$$

where the horizontal maps are precharts and the vertical maps are morphisms in $\mathcal{D}|_{\mathrm{det}}$*, and given* $\alpha : \phi_1 \to \phi_2$ *and* $\beta : \psi_1 \to \psi_2$*, there is a prechart morphism* $\psi_1\phi_1 \to \psi_2\phi_2$*.*

Proof. Unlike the proof of Lemma 6.2.5, this is straightforward: If the carrier of ϕ_i is M_i, and of ψ_i, N_i (for $i = 1, 2$), then by definition the composites are carried by the pullback $M_i \times_{Y_i} N_i$. There is a canonical map $M_1 \times_{Y_1} N_1 \to M_2 \times_{Y_2} N_2$ over M_2, N_2, given by the independent pairing of α and β.

Since pullbacks compose, the square over $M_1 \times_{Y_1} N_1$ that must commute is given by the two commutative squares induced by α, β, pulled back and composed with each other. Here we are pulling back along the deterministic projections from the pullback, and hence these commutative squares are preserved, and hence the composite square commutes as well. □

Proposition 6.2.7. *Let* $\mathcal{D}$ *be a stochastic module over* $\mathcal{C}$*. There is a double category* $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{chart}}$ *which has decorated spans as its horizontal maps, morphisms in* $\mathcal{D}|_{\mathrm{det}}$ *as its vertical maps, and decorated span 2-cells as its 2-cells.*

Moreover, there is another double category $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{lens}}$ *which has decorated spans* in $\mathcal{D}^{\mathrm{fop}}$ *as its horizontal maps instead.*

(The modification of everything above to lenses instead of charts is obvious).

In fact, the globular 2-cells are in a sense exactly the equations defining the set of charts:

Lemma 6.2.8. *Let* $\phi, \psi : \bar{A} \to \bar{B}$ *be parallel decorated spans in* $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{chart}}$ *(or* $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{lens}}$ *).*

They represent the same stochastic chart (lens) if and only if there exists a zig-zag of globular 2-cells between them.

Proof. Unwinding the definitions, a 2-cell $M_\phi \to M_\psi$ is precisely a map exhibiting the equality of the two decorated spans, as in Lemma 3.4.7 □

The following straightforward proposition shows that charts over deterministic bases have initial representatives as spans. This will be important later:

Proposition 6.2.9. *Suppose given a square*

$$\begin{array}{ccc} \bar{X}_1 & \xrightarrow{\phi_1} & \bar{Y}_1 \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} \\ \bar{X}_2 & \xrightarrow{\phi_2} & \bar{Y}_2 \end{array}$$

in $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{lens}}$ *(or* $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{chart}}$*) Suppose further the underlying square in* $\mathcal{C}$ *is deterministic. Then:*

1. *If there exists a filling decorated span 2-cell, the image in* $\mathbb{A}\mathsf{rena}(\mathcal{D}|_{\mathrm{det}})$ *commutes.*
2. *If* M_1 *is the carrier of* ϕ_1 *and the left leg* $M_1 \to X_1$ *is an isomorphism, the implication of 1. is an equivalence.*

Proof. Given a stochastic chart carried by $X \leftarrow M \to Y$, so that $X \to Y$ is deterministic, recall that we can first pull back to the equalizer $M' \hookrightarrow M$ of the two maps $M \to X \to Y, M \to Y$, then along the prescribed section $X \to M'$. Note that this gives a 2-cell from this canonical representative with carrier X to the initial representative.

Given two such cells, we get a square

$$\begin{array}{ccc} M_1 & \longrightarrow & M_2 \\ \uparrow & & \uparrow \\ X_1 & \longrightarrow & X_2 \end{array}$$

But as part of the square surrounding the 2-cell, there is given a map $X_1 \to X_2$, which must make this square commute. The functoriality of base change (pulling back the map ϕ_2) proves this bottom map is again a 2-cell.

But the property for this map to be a 2-cell is exactly the property for this square to be a commutative square in $\mathbb{A}\mathsf{rena}\mathcal{D}|_{\mathrm{det}}$. This proves both parts of the statement. □

6.3 Double Categories of Stochastic Dynamical System

Definition 6.3.1 (Stochastic Dynamical Systems Theories)**.** A *stochastic dynamical systems theory* consists of a stochastic module $\mathcal{D}$ over $\mathcal{C}$ equipped with a section $T : \mathcal{C}_{\mathrm{det}} \to \mathcal{D}|_{\mathrm{det}}$ of the underlying fibration.

We adopt the notation $\widetilde{\mathbb{A}\mathsf{rena}}(\mathcal{D})$ for the double category which was denoted $\widetilde{\mathbb{S}\mathsf{pan}}(\mathcal{D})^{\mathrm{lens}}$ above. We use the symbol $\leftrightarrows$ for the horizontal (span) morphisms and $\rightrightarrows$ for the vertical morphisms (in $\mathcal{D}$).

Given a stochastic dynamical systems theory $(\mathcal{D}, T)$, the *double category of dynamical systems* $\mathsf{Sys}(\mathcal{D}, T)$ has

1. Objects triples $S \in \mathcal{C}, \bar{A} \in \mathcal{D}, \xi : TS \leftrightarrows \bar{A} \in \widetilde{\mathbb{A}\mathsf{rena}}$

2. Horizontal morphisms given by a pair $f : S \to S', g : \bar{A} \rightrightarrows \bar{B} \in \mathcal{D}$, c where c is a square filling ξ, ξ', Tf, g

3. Vertical morphisms given by a pair $f : S \to S', g : \bar{A} \leftrightarrows \bar{B} \in \overline{\mathcal{D}}$, c where c is a square filling $Tf, g\xi, \xi', 1_{\bar{B}}$

Proposition 6.3.2. *Let* $T, \mathcal{D}$ *be a stochastic dynamical systems theory. Then* $T, \mathcal{D}|_{\text{det}}$ *is an ordinary dynamical systems theory. There is a double functor* $\mathsf{Sys}(T, \mathcal{D}|_{\text{det}}) \to \mathsf{Sys}(T, \mathcal{C})$. *This functor is full on vertical morphisms, and on 2-cells.*

Let $\mathsf{Sys}(T, \mathcal{D})_{\text{det}}$ *denote the subcategory spanned by systems with deterministic readout, prelenses with deterministic base, and all the morphisms of* $\mathcal{D}$. *Then the restricted functor* $\mathsf{Sys}(T, \mathcal{D}|_{\text{det}}) \to \mathsf{Sys}(T, \mathcal{D})_{\text{det}}$ *admits a chartwise right adjoint, which assigns to each system or prelens its equivalence class.*

Proof. The functor simply acts as the functor in Proposition 6.3.3. Since that inclusion is full on 2-cells, this one is full on vertical morphisms, and since a 2-cell in Sys is merely a 2-cell in $\mathbb{A}\mathsf{rena}$ between specific objects, it is also full on 2-cells.

The right adjoint property likewise follows from the analogous property of the inclusion functor on arenas. □

Proposition 6.3.3. *Let* $\mathcal{D}$ *be a Markov prefibration. Then there is a double functor* $\mathbb{A}\mathsf{rena}(\mathcal{D}|_{\text{det}}) \to \widetilde{\mathbb{A}\mathsf{rena}}(\mathcal{D})$, *which acts as identity on the morphisms of* $\mathcal{D}$, *and carries each lens* $(f : X \to Y, \phi : f^*\bar{Y} \to \bar{X})$ *to the prelens* $(X \leftarrow X \to Y, \phi)$. *This double functor is full on 2-cells.*

The restriction to $\mathbb{A}\mathsf{rena}(\mathcal{D}|_{\text{det}}) \to \mathbb{A}\mathsf{rena}(\mathcal{D})_{\text{det}}$ *admits a right adjoint, which carries every prelens to its equivalence class.*

Proof. First we must verify functoriality with respect to lens composition. To that end, let $X \xrightarrow{f} Y \xrightarrow{g} Z$ be deterministic morphisms and ϕ, ψ be lenses over them. Their composite as prelenses has carrier $X \times_Y Y = X$ and $f^*(\psi)\phi$ as the morphism in the fiber, which is exactly the prelens associated to their composite as lenses.

Second, we must verify fullness on 2-cells. But this is a special case of Proposition 6.2.9.

Finally, Proposition 6.2.9 is precisely the statement that assigning a prelens to its equivalence class lens is right adjoint to this (and in particular that it forms a functor) □

Recall that in Myers' categorical dynamical systems theory, trajectories of a system $\xi : TS \leftrightarrows \bar{A}$ are identified with chart morphisms from a "clock" system—thus for example trajectories of a smooth dynamical system $M \to TM$ are exactly those maps $\gamma : \mathbb{R} \to M$ which, when $\mathbb{R}$ is equipped with the vector field $dx/dt = 1$, are homomorphisms.

In general this presents an issue for our replacement category $\mathsf{Sys}(T, \mathcal{D})$—since we wish to regard two systems given by equivalent lenses as equivalent, but their set of homomorphisms from a given clock system is not necessarily in bijection. In the general case, we do not presently have a way around this problem—but at least for clock systems with deterministic readout, the above presents a solution: choosing the initial representative for such a system, we find that the set of trajectories does not depend on the equivalence class of the target system.

6.4 The triple category of dynamical systems

Having done the prep-work in chapter 4 (and section 2.6), we can jump directly into the construction. For a category $\mathcal{C}$, let $\mathcal{C}^{\simeq}$ denote the maximal subgroupoid of $\mathcal{C}$ (that is, the subcategory containing only the isomorphisms.)

Definition 6.4.1 (BiSys)**.** Let $\mathcal{A} \to \mathcal{C}, \mathsf{T} : \mathcal{C} \to \mathcal{A}$ be a symmetric monoidal dynamical systems theory. Then we may define a symmetric monoidal double category by taking the vertical category to be $\mathcal{C}^{\simeq}$, and the horizontal category to be $\mathcal{C}$ (with the squares being the commutative squares in $\mathcal{C}$). There is a symmetric monoidal double functor from this to $\mathsf{Arena}(\mathcal{A})$ given by T in both directions (since T applied to an isomorphism gives an isomorphism in charts, which is the same as an isomorphism in lenses). The following diagram depicts this data:

$$\begin{array}{ccc} (\mathcal{C}^{\to})^{\simeq} & \xrightarrow[(\mathsf{T}^{\to})]{} & \mathsf{Arena}(\mathcal{A})_1 \\ \downarrow\downarrow & & \downarrow\downarrow \\ \mathcal{C}^{\simeq} & \xrightarrow[\mathsf{T}]{} & \mathsf{Arena}(\mathcal{A})_0 \end{array}$$

This induces an object of $\mathsf{SymMon}(\mathsf{Act}(\mathsf{DblCat}))$. Applying $\mathbb{P}\mathsf{ara}(-)$ under $\mathsf{SymMon}(-)$, we obtain a symmetric pseudomonoid in internal pseudocategories in DblCat. Denote by $\mathsf{BiSys}(\mathcal{C}, \mathcal{A}, \mathsf{T})$ this induced symmetric monoidal pseduocategory in strict double categories.

Remark 6.4.2 (The Structure of BiSys)**.** $\mathsf{BiSys}(\mathcal{C}, \mathcal{A}, \mathsf{T})$ has the following structure: The objects are simply the objects of $\mathcal{A}$—the bundles.

There are three types of 1-cell: Lenses, that is morphism in $\mathcal{A}^{\text{fop}}$, which we write $A \leftrightarrows B$, charts, that is morphisms in $\mathcal{A}$, which we write $A \rightrightarrows B$, and *bisystems,* which are pairs $(S \in \mathcal{C}, \mathsf{T}S \otimes A \leftrightarrows B)$, and which we write $A \twoheadrightarrow B$.

Moreover there are three types of 2-cell:

1. Lens-chart cells, which are the 2-cells of $\mathbb{A}\mathsf{rena}(\mathcal{A})$. Note that this double category is thin. We will say a square of lenses and charts *commutes* if it is filled by such a 2-cell.

2. Chart-bisystem cells—given charts $A_1 \rightrightarrows B_1, A_2 \rightrightarrows B_2$ and systems $\mathsf{T}S \otimes A_1 \leftrightarrows A_2, \mathsf{T}S' \otimes A_2 \leftrightarrows B_2$, a square filling this is a choice of map $S \to S' \in \mathcal{C}$ so that the resulting lens-chart square

$$\begin{array}{ccc} \mathsf{T}S \otimes A_1 & \leftrightarrows & A_2 \\ \downarrow\downarrow & & \downarrow\downarrow \\ \mathsf{T}S' \otimes B_1 & \leftrightarrows & B_2 \end{array}$$

   commutes.

3. Lens-bisystem cells. Given lenses $A_1 \leftrightarrows B_1, A_2 \leftrightarrows B_2$, and systems $\mathsf{T}S' \otimes A_1 \leftrightarrows A_2, \mathsf{T}S' \otimes B_1 \leftrightarrows B_2$, a 2-cell consists of an *isomorphism* $S' \xrightarrow{\sim} S'$ so that the resulting square of lenses commutes (recall that isomorphism charts are the same as isomorphism lenses).

$$\begin{array}{ccc} \mathsf{TS} \otimes \mathsf{A}_1 & \leftrightarrows & \mathsf{A}_2 \\ \downarrow\uparrow & & \downarrow\uparrow \\ \mathsf{TS}' \otimes \mathsf{B}_1 & \leftrightarrows & \mathsf{B}_2 \end{array}$$

The charts, bisystems and chart-bisystem cells form a pseudo double category with the obvious composition. So do the lenses, bisystems and lens-bisystem cells. Finally, there is a notion of 3-cell given by a box whose sides are 2-cells of each kind, so that the resulting diagram in $\mathcal{C}$ (with isomorphisms on two sides) commutes.

The lens-bisystem double category is $\mathbb{P}\mathsf{ara}_{\mathcal{C}^{\simeq}}(\mathbb{A}\mathsf{rena}\mathcal{A}_0)$ (recall that $\mathbb{A}\mathsf{rena}\mathcal{A}_0 = \mathcal{A}^{\mathrm{fop}}$), with the action given by the functor T (restricted to isomorphisms). The chart-bisystem double category is the result of taking $\mathsf{BiSys}(\mathcal{C}, \mathcal{A}, \mathsf{T})$, a pseudocategory in double categories, and applying $\mathsf{PsCat}((-)_h) : \mathsf{PsCat}(\mathsf{DblCat}) \to \mathsf{PsCat}(\mathsf{Cat})$ where $(-)_h$ takes the *horizontal* category of a (strict) double category.

The 3-cells, of course, are the 2-cells of $\mathbb{P}\mathsf{ara}_{(\mathcal{C}^{\to})^{\simeq}}(\mathbb{A}\mathsf{rena}(\mathcal{A})_1)$. Analogously to the above, for each class of 1-cells, there is a double category with those as the objects, the two types of cell as the two morphisms, and the 3-cells as the 2-cells.

All of these six double categories admit a symmetric monoidal structure.

We can recover the ordinary category of (Moore) systems as the slice over I, in the following sense:

Proposition 6.4.3. *There is a pullback diagram*

$$\begin{array}{ccc} \mathsf{Sys}(\mathsf{T}) & \longrightarrow & \mathsf{BiSys}(\mathsf{T})_1 \\ \big\downarrow & \lrcorner & \big\downarrow d \\ \bullet & \xrightarrow{\ \mathrm{I}\ } & \mathsf{BiSys}(\mathsf{T})_0 \end{array}$$

in DblCat

In a double category $\mathbb{C}$, there is a "horizontal slice" 1-category having objects the horizontal maps $\mathsf{A} \to \mathsf{B}$ and morphisms given by 2-cells that are identity on the left boundary, composed vertically. Similarly there is a "vertical slice". This is given by a similar pullback in Cat—we simply do this one level up.

Note that Myers' construction of the double fibration $\mathsf{Sys}(\mathcal{C}, \mathcal{A}, \mathsf{T}) \to \mathbb{A}\mathsf{renaCat}$ in fact uses the vertical slice in this sense.

This somewhat trivial observation means that any composition in $\mathsf{BiSys}(\mathcal{C}, \mathcal{A}, \mathsf{T})$ which produces 2-cell under I in fact produces a morphism of systems in the ordinary sense. Replacing I with another object, we may regard the slices as further-parametrized versions of $\mathsf{Sys}(\mathcal{C}, \mathcal{A}, \mathsf{T})$.

There is essentially no difficulty in applying this to the Markov case:

Definition 6.4.4 ($\mathsf{BiSys}^{\mathsf{M}}$)**.** Let $(\mathcal{C}, \mathcal{A}, \mathsf{T})$ be a stochastic dynamical systems theory. Consider the double category whose vertical category is $\mathcal{C}_{\mathrm{det}}$ and horizontal category is $\mathcal{C}$, with commutative squares as the 2-cells (this is the *transpose* of $\mathcal{C}^{\to} \rightrightarrows \mathcal{C}$). This carries an obvious symmetric monoidal structure, and acts on the double category of stochastic arenas $\widetilde{\mathbb{A}\mathsf{rena}}(\mathcal{A})$ via the functor T. Denote by $\mathsf{BiSys}^{\mathsf{M}}(\mathcal{C}, \mathcal{A}, \mathsf{T})$ the symmetric monoidal triple category induced as in Definition 6.4.1 by this data.

The information contained in $\mathsf{BiSys}^{\mathcal{M}}(\mathcal{C}, \mathcal{A}, T)$ is much as above, although the complications involved in the double category of stochastic arenas remain present.

Our bisystems are reminiscent of the *energy-driven systems* of Capucci, Lynch, and Spivak [CLS24]. Indeed, their $\mathbb{C}$org is essentially the bisystems in the (ordinary) doctrine of smooth dynamical systems. As we mentioned in the introduction, Shapiro and Spivak [SS23] have previous developed a structure $\mathbb{O}$rg which consists of the bisystems for the discrete dynamical systems doctrine (i.e $\mathsf{Set}^{\rightarrow} \to \mathsf{Set}$). $\mathbb{O}$rg has a tremendous amount of structure coming from the representation of lenses in this doctrine as the category of polynomial functors, which can't be replicated for a general systems theory (and certainly not for a general *stochastic* systems theory).

6.5 A Stochastic Dynamical Systems Theory of Smooth Manifolds

In this section, as the title suggests, we construct a stochastic dynamical systems theory of smooth manifolds, with the usual tangent bundle. The main point is to construct a Markov category containing the smooth manifolds which is pullback-positive. We do this by considering the larger category of *diffeological spaces*. In order to make the topology work, we need to complicate the notion of diffeological space a bit, but having done so, we obtain a representable Markov category which is pullback-positive, and contains SmMfd as a full subcategory of the deterministic maps. A kernel $p : M \to N$ is a Markov kernel valued in Radon measures which is weakly continuous—so induces a linear map $C(N) \to C(M)$ taking ϕ to the function $x \mapsto E_{p_x}\phi$ on the spaces of continuous functions—and which furthermore smooth in the sense that this operation preserves the smooth functions.

This Markov category of "smooth stochastic maps" may be of some independent interest.

Definition 6.5.1 (CartSp)**.** Let CartSp denote the full subcategory of SmMfd spanned by the objects $\mathbb{R}^n$ for each n. Note that CartSp has finite products, and is generated by the object $\mathbb{R}$ under finite products.

Definition 6.5.2 (Diffeological Space)**.** A *smooth space* is a sheaf on SmMfd in the standard topology of open covers. A smooth space X is a *diffeological space* if, for each $M \in \mathsf{SmMfd}$, the map $X(M) \to \prod_{p \in M} X(\{p\})$ is injective.

A diffeological space with underlying set X is called a *diffeology* on X, and consists of specifying which maps $\mathbb{R}^n \to X$ are smooth. We call these maps *smooth plots*.

Given a subset $X' \subset X$, there is an obvious canonical diffeology on X' given by taking the plots to be those functions whose image in X is smooth. We call this the *subspace diffeology*.

A morphism of diffeological spaces is called a *smooth map*. It is equivalently a function $X \to Y$ which carries smooth plots to smooth plots.

Definition 6.5.3 (Diffeo-Topological space)**.** A *diffeo-topological space* is a diffeological space X equipped with a topology τ so that all the smooth plots are continuous. A map of diffeo-topological spaces is a smooth map (for the diffeology) which is also continuous (for the topology).

Given a subset $X' \subset X$, there is an obvious canonical diffeo-topology on X' given by taking the plots to be those functions whose image in X is smooth, and equipping X' with the subspace topology.

Lemma 6.5.4. *The category of diffeo-topological spaces admits all limits, given by taking the limits in topological spaces and diffeological spaces (which have the same underlying set).*

Proposition 6.5.5. *Let* X *be a diffeo-topological space whose underlying space is Tychonoff. Then the space of probability measures* $\mathsf{P}(\mathsf{X})$ *has a canonical diffeology given by those plots* $f : U \to \mathsf{P}(\mathsf{X})$ *so that for each continuous function* g *on* X*, the resulting map* $u \mapsto E_{f(u)}g$ *is continuous, and if* g *is smooth, then this is smooth as well. With this diffeology, and the topology of weak convergence,* $\mathsf{P}(\mathsf{X})$ *is a diffeo-topological space. This defines a commutative affine monad on* $\mathsf{TychDiff}$*, the category of such diffeo-topological spaces.*

Proof. The topology of weak convergence on $\mathsf{P}(\mathsf{X})$ is such that $A \to \mathsf{P}(\mathsf{X})$ is continuous if and only if the expectation map carries continuous functions to continuous functions. But this is part of the requirement to be a smooth plot, so certainly this is a diffeo-topological space.

Since the linear operator associated to $x \in X$ under the unit $X \to \mathsf{P}(X)$ is merely evaluation at x, the unit is clearly smooth. Consider $\mu : PPX \to PX$. To test this is smooth, let $f : U \to PP(X)$ be a plot. We must show μf is a plot. So let $\nu : X \to \mathbb{R}$ be a continuous (resp. smooth) function. We must show its expectation a is continuous (resp. smooth) function of $u \in U$.

By construction $E(\nu) : PX \to \mathbb{R}$ is continuous (resp. smooth), and so since f is a plot, the map $u \mapsto E_{f(u)}E(\nu)$ is continuous (resp. smooth). But this is exactly what we wanted.

The monad laws follow from their holding in Tych. Since a commutative monad is equivalently a strong monad satisfying a certain equation (which holds for this monad in Tych and therefore also here), it suffices to show that the strength $P(X) \times Y \to P(X \times Y)$ is smooth. This follows by a completely analogous argument. □

Corollary 6.5.6. *The category* $\mathsf{TychDiffStoch}$ *of Tychonoff diffeological spaces and Kleisli maps for the monad* P *described in Proposition 6.5.5 is a pullback-positive Markov category. Its deterministic category is* $\mathsf{TychDiff}$*. There is a fully faithful functor* $\mathsf{SmMfd} \to \mathsf{TychDiff}$ *which preserves transverse pullbacks. The maps between smooth manifolds are given by weakly continuous families of Radon probability measures, so that the expectation operator carries smooth maps to smooth maps.*

Remark 6.5.7. Since we describe properties of kernels in terms of their corresponding linear operator on function spaces, it would seem natural to take the function spaces as the basic object. Hence we might consider the category of C^∞-algebras with some relaxed notion of maps between them. The tricky part there is to find some reasonable class of maps so that the tensor product (coproduct) of C^∞-algebras extends to these. (Since it is not the same as the tensor product of $\mathbb{R}$-algebras, linear maps do not automatically extend to the tensor product). In particular, when considering kernels $*_1 \to X$ where $*_1$ is a "fat point of order 1"—that is, functions on 1_* have a value at the point and a derivative—it is not clear what sort of continuity condition the derivative operation on $C^\infty(X)$ should satisfy, nor how to define this in a general way for all C^∞-algebras. It would certainly be of interest to synthetic computational geometry to have such a Markov category, but we leave this for future work.

We will now give an example of how to represent the training dynamics of a machine learning system using the tools developed so far. As discussed previously, given a parameterized function $F : P \times X \to Y$, its reverse derivative naturally becomes a parameterized lens, and the composition of these describe how gradient vectors are passed around to compute an update during training. It is natural to want to compose this lens with the data-generating distribution $I \to X \otimes Y$, (along with some more context describing the loss function, etc) to obtain the training dynamics of such a model. This requires a category of parameterized lenses which allows stochastic maps in the base. The goal of combining this feature with non-trivial tangent bundles was one of the original motivations for developing a theory of stochastic lenses.

Example 6.5.8. Consider the Markov prefibration $\mathsf{TychDiffStoch}^{\to} \to \mathsf{TychDiffStoch}$. Equip this with the section $T(X) = X \otimes X \xrightarrow{\pi_0} X$—this describes discrete-time systems (whose update is required to be smooth in the input and present state). This is clearly a symmetric monoidal functor and thus defines a systems theory. Note that this is completely different from the ordinary tangent bundle, despite the coincidence of notation.

Let $m_1 : TS_2 \otimes X_1 \leftrightarrows Y_1$ and $m_2 : TS_2 \otimes X_2 \leftrightarrows Y_2$ be two bisystems in this theory. As in section 5.3, we may define a parameterized lens $(TS_1 \& TS_2) \otimes (X_1 \oplus X_2) \leftrightarrows (Y_1 \oplus X_2)$, denoting by $\oplus$ the coproduct in lenses, and by $\&$ the Markov structure defined in Corollary 3.6.7. Observe that $TS_1 \& TS_2$ is simply the indexed set $(S_1 \coprod S_2) \otimes S_1 \otimes S_2 \to S_1 \otimes S_2$. There is an obvious indexed map from this to $T(S_1 \otimes S_2) = S_1 \otimes S_2 \otimes S_1 \otimes S_2$, given by $(\mathrm{inl}\, s_1', s_1, s_2) \mapsto (s_1', s_2, s_1, s_2)$ and $(\mathrm{inr}\, s_2', s_1, s_2) \mapsto (s_1, s_2', s_1, s_2)$. In words, we receive an update either to the S-state or to the S'-state. We apply this update to the relevant state and leave the other alone. This defines a lens $T(S_1 \otimes S_2) \leftrightarrows T(S_1) \& T(S_2)$, which we may compose with the above to obtain a bisystem $T(S_1 \otimes S_2) \otimes (X_1 \oplus X_2) \leftrightarrows Y_1 \oplus Y_2$. Let us denote this by $m_1 \oplus m_2$.

This has very much the same flavor as the external choice for open games, although we will not develop the theory of this operation in detail here.

Now, given a smooth (deterministic) map $F : P \times X \to Y$, where P, X, Y are smooth manifolds (not merely diffeo-toplogical spaces), we obtain a lens $T^*(F) : \binom{T^*P}{P} \otimes \binom{T^*X}{X} \leftrightarrows \binom{T^*Y}{Y}$, where $T^*(-)$ denotes the cotangent bundle. (This does not, prima facie, make sense for a general diffeo-topological space).

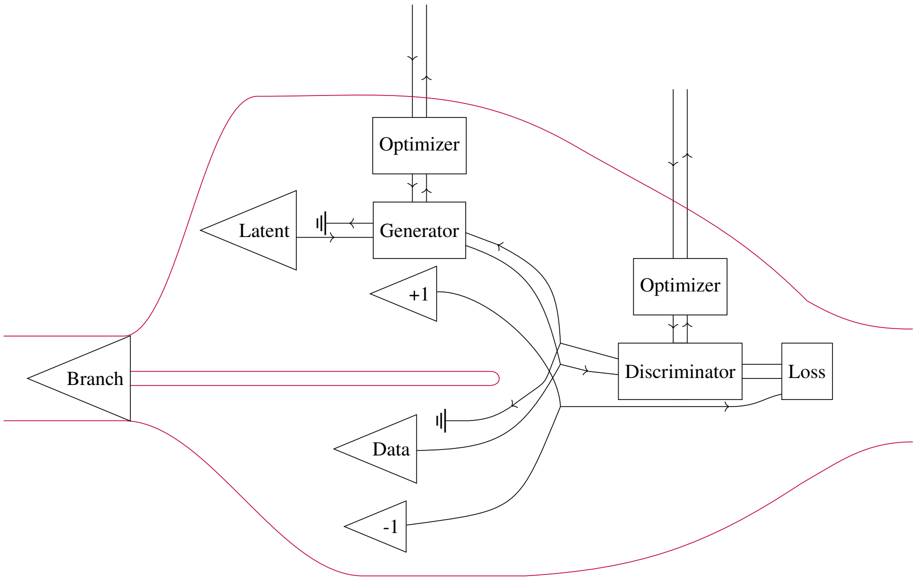
Let us take as given some family of lenses $T(S) = \binom{S \otimes S}{S} \leftrightarrows \binom{T^*S}{S}$ for various S. Such an operation amounts to choosing a way of updating $s \in S$ given a cotangent vector—hence we can see it as an optimization algorithm. One example of such would be gradient descent, which given a choice of Riemann structure on S, takes a step of a given length in the direction which most quickly decreases the given covector.

(It should be noted that there are more complicated optimization strategies which don't fit this particular pattern—for example, *momentum* algorithms have to maintain some extra internal state other than $s \in S$. But let's stick with this pattern for this example). Note also that we're not assuming the optimizers are a natural transformation or anything like that.

Now we are ready to build the neural network architecture known as a generative adversarial network, or GAN [Goo+14]. Let us first describe the idea. Our goal is to generate additional samples from some distribution, given a set of existing samples—for example, our goal may be to generate more pictures in the same style as an existing corpus. Suppose our data is of type X, and let $d : I \to X$ be the data distribution. We fix some *latent distribution* $\lambda : I \to L$, where L is any space of our choice—usually, L is $\mathbb{R}^n$ and λ is a Gaussian. Finally we choose two neural networks, the *generator* $G : P_G \otimes L \to X$, and the *discriminator* $D : P_D \otimes X \to \mathbb{R}$. The goal of the discriminator is to discriminate real samples from the data from generated samples, by providing a low value on the generated samples and a high value on the true samples.

The training process now goes as follows: for each step of training, we either sample from the latent distribution, and have the generator use this to generate a sample, or draw a sample from the data distribution (choosing between these with some probability p). Then in either case, we have the discriminator score the generated sample. If the sample was generated by the generator, the discriminator's loss is equal to its output, otherwise it is equal to -1 times its output. We update the discriminator according to the gradient of this loss (minimizing it), and update the generator (if we are in the branch where it was run) according to the negative of the gradient of this loss with respect to the generator parameters—this amounts to doing a gradient descent update on the generator for the negative of the discriminator loss.

We can represent this schematically using the following *tape diagram* (see [Bon+25]):



The two "tapes" branching off at the start represent two maps (bisystems) composed by $\oplus$, while the backwards wires indicate the flow of the gradients. The ground symbol indicates a value being discarded. Note that the theory of tape diagrams has only been developed formally for *distributive* categories, and for essentilly the same reason as in section 5.3, we do not have distributivity in this case. However, the interpretation of this diagram is still unambiguous—distributivity is required to make tape diagrams *complete*, not to make them sound. Concretely, if we tried to represent the tensoring of this system with another system, we would have no way of doing it, but if the category was distributive we could do so by adding this additional system to each of the branches. Still, the figure is best viewed as a visual aid rather than a formal representation.

Chapter 7

Conclusions & Future work

7.1 Contributions

We now outline briefly the contributions of the present thesis. We have introduced the novel theory of stochastic module fibrations, Markov fibrations, and -prefibrations. These capture in different ways the ways in which a indexed stochastic mapping between indexed sets can be reindexed along another stochastic mapping. Of these, stochastic module fibrations are the most general, while Markov (pre)fibrations give convenient ways of presenting the data of a stochastic module.

This theory allows for a slick definition of "stochastic dependent lenses" in the following way: by applying the construction which attempts to reconstruct a (pre)fibration from its underlying stochastic module to the fiberwise opposite of an indexed family fibration, we obtain a suitable category of stochastic lenses. This construction also recovers the ordinary category of optics as the similar fiberwise opposite of a stochastic module associated to the natural category of "trivially-indexed" objects, and thus the classical theory of optics really becomes a special case of this theory in the expected way. Similarly, the classical theory of dependent lenses is merely the case when the base is Cartesian, again the expected specialization.

In addition to the theory necessary to present the above examples, we also develop a few elaborations, such as the theory of monoidal stochastic modules/Markov fibrations, and some results on when these structures are preserved under limits of categories.

We go on to apply this theory to give a definition of a category of open games which supports an external choice operator, resolving an open problem in the theory of open games. Our construction is parametric in a Markov fibration, inserting the fibration $\mathsf{Kl}(\Delta)^{\rightarrow}$ of discrete probability recovers Bayesian open games with mixed strategies in the usual sense.

In parallel with this, we give a detailed construction of the double category $\mathbb{P}\mathsf{ara}$, generalized to actions of internal pseudomonoids in any 2-category. We further use this to present a triple categorical extension of Myers' bicategories of dynamical systems—to be precise, it is a symmetric monoidal pseudo double category internal to strict double categories. While the slices of this object capture the composition of open systems in the sense of Myers, composition in the third dimension captures Gavranović's composition of gradient based learners.

We also develop a distinct extension of Myers' theory, allowing for dynamical systems to compose along stochastic maps. Along the way we also develop a Markov category of diffeological spaces and "smooth stochastic maps" which may be of independent interest.

7.2 Future work

There are a number of obvious points for future development of the theory of Markov fibrations. First, the present theory requires the underlying category to be pullback-positive (Definition 2.4.9), and in particular the deterministic subcategory must *have pullbacks*. This is true in many cases of interest, but false in many others, for example smooth manifolds. In some cases this problem can be resolved by embedding in a larger category which does have pullbacks, but it would be interesting to see whether it could be developed without this assumption. One direction would be to try to construct free prefibrations without this assumption (after all, the universal property of $\overline{\mathcal{D}_0}$ still makes sense, we just don't know if an object with it exists). An alternative would be to try to make sense of the notion of stochastic module without using this assumption, which also seems tractable.

In some cases, namely Markov categories with fibred supports but not conditionals, we know that $\mathcal{C}^{\rightarrow} \rightarrow \mathcal{C}$ presents a Markov fibration, but not what it is. The expectation is that it will be some subcategory of $\mathcal{C}^{\rightarrow}$ (the assumption about conditionals is required to prove that all squares admit a representation as a chart). It would be interesting to understand this subcategory better.

The comparison to the various forms of linear dependent type theory is also very interesting. Of course, type theories involve a number of extra assumptions about which types can be formed (sum types and so on) whereas a Markov fibration only involves reindexing. Still it seems a plausible conjecture that for some of these type theories, there should be a Markov fibration (or stochastic module) over contexts which classifies the types in context.

When it comes to external choice in open games, we briefly touched on the connection to extensive form, but did not get into the details, nor did we prove any formal connection between these. It would be interesting to try to prove a formal comparison theorem, and consider modifications to the theory that would yield subgame perfect equilibrium rather than Nash equilibrium as the composite solution concept.

Bibliography

[ABK24] Nathanael Arkor, John Bourke, and Joanna Ko. "Enhanced 2-categorical structures, two-dimensional limit sketches and the symmetry of internalisation". In: arXiv:2412.07475 (Dec. 2024). DOI: 10.48550/arXiv.2412.07475. URL: http://arxiv.org/abs/2412.07475.

[Atk18] Robert Atkey. "Syntax and Semantics of Quantitative Type Theory". In: *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. July 2018, pp. 56–65. DOI: 10.1145/3209108.3209189.

[Atk+21] Robert Atkey, Bruno Gavranović, Neil Ghani, Clemens Kupke, Jérémy Ledent, and Fredrik Nordvall Forsberg. "Compositional Game Theory, Compositionally". In: *Electronic Proceedings in Theoretical Computer Science* 333 (2021), pp. 14–30. ISSN: 2075-2180. DOI: 10.4204/EPTCS.333.2. URL: https://arxiv.org/abs/2101.12045.

[BC20] John C. Baez and Kenny Courser. "Structured cospans". In: *Theory and Applications of Categories* 35.48 (Oct. 2020), pp. 1771–1822.

[Ben67] Jean Benabou. "Introduction to bicategories". en. In: *Reports of the Midwest Category Seminar*. Ed. by J. Benabou et al. Berlin, Heidelberg: Springer, 1967, pp. 1–77. ISBN: 978-3-540-35545-8. DOI: 10.1007/BFb0074299.

[Bén85] Jean Bénabou. "Fibered categories and the foundations of naive category theory". In: *The Journal of Symbolic Logic* 50.1 (1985), pp. 10–37. DOI: 10.2307/2273784.

[Boc+23] Guido Boccali, Bojana Femić, Andrea Laretto, Fosco Loregian, and Stefano Luneia. "The semibicategory of Moore automata". In: arXiv:2305.00272 (Apr. 2023). DOI: 10.48550/arXiv.2305.00272. URL: http://arxiv.org/abs/2305.00272.

[BHW18] J. Bolt, Jules Hedges, and Viktor Winschel. "The algebra of predicting agents". In: *ArXiv* (Mar. 2018). URL: https://www.semanticscholar.org/paper/The-algebra-of-predicting-agents-Bolt-Hedges/3dfd6ac5f10eb0a1bf83f61ccda940df86b5ce5c.

[BHZ19] Joe Bolt, Jules Hedges, and Philipp Zahn. "Bayesian open games". en. In: *arXiv:1910.03656 [cs, math]* (Oct. 2019). URL: http://arxiv.org/abs/1910.03656.

[Bon+25] Filippo Bonchi, Cipriano Junior Cioffo, Alessandro Di Giorgio, and Elena Di Lavore. "Tape Diagrams for Monoidal Monads". In: arXiv:2503.22819 (Mar. 2025). DOI: 10.48550/arXiv.2503.22819.

[Bor94] Francis Borceux. *Handbook of Categorical Algebra*. 1st ed. Cambridge University Press, Aug. 1994. ISBN: 978-0-521-44178-0. DOI: 10.1017/CBO9780511525858. URL: https://www.cambridge.org/core/product/identifier/9780511525858/type/book.

[Bor+16] Johannes Borgström, Ugo Dal Lago, Andrew D. Gordon, and Marcin Szymczak. "A lambda-calculus foundation for universal probabilistic programming". In: *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP 2016)*. 2016. DOI: 10.48550/arXiv.1512.08990. URL: https://arxiv.org/abs/1512.08990.

[Bra21] Edwin Brady. "Idris 2: Quantitative Type Theory in Practice". In: *LIPIcs, Volume 194, ECOOP 2021* (2021). DOI: 10.4230/LIPIcs.ECOOP.2021.9.

[Bra+21] Dylan Braithwaite, Matteo Capucci, Bruno Gavranović, Jules Hedges, and Eigil Fjeldgren Rischel. "Fibre optics". In: arXiv:2112.11145 (Dec. 2021). DOI: 10.48550/arXiv.2112.11145. URL: http://arxiv.org/abs/2112.11145.

[BH22] Dylan Braithwaite and Jules Hedges. "Dependent Bayesian Lenses: Categories of Bidirectional Markov Kernels with Canonical Bayesian Inversion". In: arXiv:2209.14728 (Sept. 2022). DOI: 10.48550/arXiv.2209.14728. URL: http://arxiv.org/abs/2209.14728.

[Buc14] Mitchell Buckley. "Fibred 2-categories and bicategories". In: *Journal of Pure and Applied Algebra* 218.6 (June 2014), pp. 1034–1074. DOI: 10.1016/j.jpaa.2013.11.002.

[CG23] Matteo Capucci and Bruno Gavranović. "Actegories for the Working Amthematician". In: arXiv:2203.16351 (Dec. 2023). DOI: 10.48550/arXiv.2203.16351. URL: http://arxiv.org/abs/2203.16351.

[Cap+22] Matteo Capucci, Bruno Gavranović, Jules Hedges, and Eigil Fjeldgren Rischel. "Towards Foundations of Categorical Cybernetics". In: *Electronic Proceedings in Theoretical Computer Science* 372 (Nov. 2022), pp. 235–248. ISSN: 2075-2180. DOI: 10.4204/EPTCS.372.17. URL: https://arxiv.org/abs/2105.06332.

[Cap+23] Matteo Capucci, Neil Ghani, Jérémy Ledent, and Fredrik Nordvall Forsberg. "Translating Extensive Form Games to Open Games with Agency". In: *Electronic Proceedings in Theoretical Computer Science* 380 (2023), pp. 145–158. ISSN: 2075-2180. DOI: 10.4204/EPTCS.380.9. URL: https://arxiv.org/abs/2105.06763.

[CLS24] Matteo Capucci, Owen Lynch, and David I. Spivak. "Organizing Physics with Open Energy-Driven Systems". In: arXiv:2404.16140 (Apr. 2024). DOI: 10.48550/arXiv.2404.16140. URL: http://arxiv.org/abs/2404.16140.

[Car24] Kevin Carlson. *Introducing CatColab*. en. Oct. 2024. URL: https://topos.institute/blog/2024-10-02-introducing-catcolab/.

[Con72] F. Conduché. "Au sujet de l'existence d'adjoints à droite aux foncteurs "image réciproque" dans la catégorie des catégories". In: *Comptes Rendus de l'Académie des Sciences, Série A* 275 (1972), A891–A894.

[Cru+21] G. S. H. Cruttwell, Bruno Gavranović, Neil Ghani, Paul Wilson, and Fabio Zanasi. "Categorical Foundations of Gradient-Based Learning". In: arXiv:2103.01931 (July 2021). DOI: 10.48550/arXiv.2103.01931. URL: http://arxiv.org/abs/2103.01931.

[DE11] Vincent Danos and Thomas Ehrhard. "Probabilistic coherence spaces as a model of higher-order probabilistic computation". In: *Information and Computation* 209.6 (2011), pp. 966–991.

[DGP09] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. "The Complexity of Computing a Nash Equilibrium". en. In: *SIAM Journal on Computing* 39.1 (Jan. 2009), pp. 195–259. ISSN: 0097-5397, 1095-7111. DOI: 10.1137/070699652.

[Dob07] Ernst-Erich Doberkat. *Stochastic Relations: Foundations for Markov Transition Systems*. Studies in Informatics Series. Chapman and Hall/CRC, 2007. ISBN: 978-1-58488-808-4.

[Dob09] Ernst-Erich Doberkat. *Stochastic Coalgebraic Logic*. Monographs in Theoretical Computer Science. An EATCS Series. Springer, 2009. ISBN: 978-3-642-02995-5. DOI: 10.1007/978-3-642-02995-5.

[Dor25] Maximilian Doré. "Linear Types with Dynamic Multiplicities in Dependent Type Theory". In: *Proceedings of the ACM on Programming Languages* 9.ICFP (Aug. 2025). DOI: 10.1145/3747531.

[Ehr63] Charles Ehresmann. "Categories structures". fra. In: *Annales scientifiques de l'Ecole Normale Superieure* 80.4 (1963), pp. 349–426. ISSN: 0012-9593.

[Fer06] Nelson Martins Ferreira. "Pseudo-categories". In: arXiv:math/0604549 (Apr. 2006). DOI: 10.48550/arXiv.math/0604549. URL: http://arxiv.org/abs/math/0604549.

[FST19] Brendan Fong, David I. Spivak, and Rémy Tuyéras. "Backprop as Functor: A compositional perspective on supervised learning". en. In: *arXiv:1711.10455 [cs, math]* (May 2019). URL: http://arxiv.org/abs/1711.10455.

[Fre+23] Seth Frey, Jules Hedges, Joshua Tan, and Philipp Zahn. "Composing games into complex institutions". In: arXiv:2108.05318 (Mar. 2023). DOI: 10.48550/arXiv.2108.05318. URL: http://arxiv.org/abs/2108.05318.

[Fri20] Tobias Fritz. "A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics". en. In: *Advances in Mathematics* 370 (Aug. 2020), p. 107239. ISSN: 00018708. DOI: 10.1016/j.aim.2020.107239. URL: https://arxiv.org/abs/1908.07021.

[Fri+25] Tobias Fritz, Tomas Gonda, Antonio Lorenzin, Paolo Perrone, and Areeb Shah Mohammed. "Empirical Measures and Strong Laws of Large Numbers in Categorical Probability". In: arXiv:2503.21576 (Mar. 2025). DOI: 10.48550/arXiv.2503.21576. URL: http://arxiv.org/abs/2503.21576.

[Fri+23a] Tobias Fritz, Tomas Gonda, Antonio Lorenzin, Paolo Perrone, and Dario Stein. "Absolute continuity, supports and idempotent splitting in categorical probability". In: arXiv:2308.00651 (Sept. 2023). DOI: 10.48550/arXiv.2308.00651. URL: http://arxiv.org/abs/2308.00651.

[Fri+23b] Tobias Fritz, Tomas Gonda, Paolo Perrone, and Eigil Fjeldgren Rischel. "Representable Markov categories and comparison of statistical experiments in categorical probability". In: *Theoretical Computer Science* 961 (June 2023), p. 113896. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2023.113896.

[FR20] Tobias Fritz and Eigil Fjeldgren Rischel. "Infinite products and zero-one laws in categorical probability". en. In: *Compositionality* 2 (Aug. 2020), p. 3. ISSN: 2631-4444. DOI: 10.32408/compositionality-2-3.

[FKS22] Peng Fu, Kohei Kishida, and Peter Selinger. "Linear Dependent Type Theory for Quantum Programming Languages". In: *Logical Methods in Computer Science* 18.3 (Sept. 2022). DOI: 10.46298/lmcs-18(3:28)2022. URL: https://arxiv.org/abs/2004.13472.

[GK13] Nicola Gambino and Joachim Kock. "Polynomial functors and polynomial monads". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 154.1 (Jan. 2013), pp. 153–192. ISSN: 0305-0041, 1469-8064. DOI: 10.1017/S0305004112000394. URL: https://arxiv.org/abs/0906.4931.

[Gav24] Bruno Gavranović. "Fundamental Components of Deep Learning: A category-theoretic approach". In: arXiv:2403.13001 (Mar. 2024). DOI: 10.48550/arXiv.2403.13001. URL: http://arxiv.org/abs/2403.13001.

[Gha25] Neil Ghani. "A Category Theoretic Approach to Approximate Game Theory". In: *Electronic Proceedings in Theoretical Computer Science* 429 (2025). ISSN: 2075-2180. URL: https://arxiv.org/abs/2509.20932.

[Gha+18a] Neil Ghani, Jules Hedges, Viktor Winschel, and Philipp Zahn. "Compositional game theory". In: arXiv:1603.04641 (Feb. 2018). DOI: 10.48550/arXiv.1603.04641. URL: http://arxiv.org/abs/1603.04641.

[Gha+18b] Neil Ghani, Clemens Kupke, Alasdair Lambert, and Fredrik Nordvall Forsberg. "A compositional treatment of iterated open games". In: *Theoretical Computer Science* 741 (2018), pp. 48–57. DOI: 10.1016/j.tcs.2018.05.026. URL: https://arxiv.org/abs/1711.07968.

[Gha+20] Neil Ghani, Clemens Kupke, Alasdair Lambert, and Fredrik Nordvall Forsberg. "Compositional Game Theory with Mixed Strategies: Probabilistic Open Games Using a Distributive Law". In: *Electronic Proceedings in Theoretical Computer Science* 323 (2020), pp. 71–90. ISSN: 2075-2180. DOI: 10.4204/EPTCS.323.5. URL: https://arxiv.org/abs/2009.06831.

[Ghi20] Enrico Ghiorzi. "Internal enriched categories". In: arXiv:2006.07997 (June 2020). DOI: 10.48550/arXiv.2006.07997. URL: http://arxiv.org/abs/2006.07997.

[Gir82] Michèle Giry. "A categorical approach to probability theory". In: *Categorical Aspects of Topology and Analysis*. Vol. 915. Lecture Notes in Mathematics. Springer, 1982, pp. 68–85.

[Gol02] Peter V Golubtsov. "Monoidal Kleisli Category as a Background for Information Transformers Theory". en. In: (2002), p. 24.

[Goo+14] Ian J. Goodfellow et al. "Generative Adversarial Networks". In: arXiv:1406.2661 (June 2014). DOI: 10.48550/arXiv.1406.2661. URL: http://arxiv.org/abs/1406.2661.

[Gra66] John W. Gray. "Fibred and Cofibred Categories". In: *Proceedings of the Conference on Categorical Algebra, La Jolla 1965*. Ed. by S. Eilenberg, D. K. Harrison, S. MacLane, and H. Röhrl. Springer, 1966, pp. 21–83. DOI: 10.1007/978-3-642-99902-4_2.

[Gro60] Alexander Grothendieck. "Descent technique and existence theorems in algebraic geometry. I. Generalities. Descent by faithfully flat morphisms". fr. In: *Bourbaki Seminar: Years 1958/59 - 1959/60, Presentations 169-204*. Bourbaki Seminar. French Mathematical Society, 1960, pp. 299–327. URL: https://www.numdam.org/item/SB_1958-1960__5__299_0/.

[Gro71] Alexander Grothendieck. *Revêtements étales et groupe fondamental (SGA 1)*. fr. Vol. 224. Lecture Notes in Mathematics. Séminaire de Géométrie Algébrique du Bois Marie 1960–61. Springer-Verlag, 1971. DOI: 10.1007/BFb0058656.

[Hed15] Jules Hedges. "String diagrams for game theory". In: arXiv:1503.06072 (Mar. 2015). DOI: 10.48550/arXiv.1503.06072. URL: http://arxiv.org/abs/1503.06072.

[Hed16] Jules Hedges. "Towards compositional game theory". en. PhD thesis. Oct. 2016. URL: https://qmro.qmul.ac.uk/xmlui/handle/123456789/23259.

[Hed17] Jules Hedges. "Morphisms of open games". In: *arXiv:1711.07059 [cs, math]* (Nov. 23, 2017). arXiv: 1711.07059. URL: http://arxiv.org/abs/1711.07059.

[Hed18] Jules Hedges. *Lenses for philosophers*. en. Aug. 2018. URL: https://julesh.com/2018/08/16/lenses-for-philosophers/.

[Hed24] Jules Hedges. *Enriched Closed Lenses*. Cybercat Insitute. Dec. 4, 2024. URL: https://cybercat.institute/2024/04/12/enriched-closed-lenses/ (visited on 08/25/2025).

[HS24] Jules Hedges and Riu Rodríguez Sakamoto. "Reinforcement Learning in Categorical Cybernetics". In: arXiv:2404.02688 (Apr. 2024). DOI: 10.48550/arXiv.2404.02688. URL: http://arxiv.org/abs/2404.02688.

[HT12] C. Hermida and R.D. Tennent. "Monoidal indeterminates and categories of possible worlds". en. In: *Theoretical Computer Science* 430 (Apr. 2012), pp. 3–22. ISSN: 03043975. DOI: 10.1016/j.tcs.2012.01.001.

[Her99] Claudio Hermida. "Some properties of Fib as a fibred 2-category". In: *Journal of Pure and Applied Algebra* 134.1 (1999), pp. 83–109. DOI: 10.1016/S0022-4049(97)00129-1.

[Heu+17] Chris Heunen, Ohad Kammar, Sam Staton, and Hongseok Yang. "A convenient category for higher-order probability theory". In: arXiv:1701.02547 (2017). DOI: 10.48550/arXiv.1701.02547. URL: https://arxiv.org/abs/1701.02547.

[Jac99] Bart Jacobs. *Categorical logic and type theory*. en. 1st ed. Studies in logic and the foundations of mathematics. Amsterdam; New York: Elsevier Science, 1999. ISBN: 978-0-444-50170-7.

[Jac08] Bart Jacobs. "Coalgebraic Trace Semantics for Combined Possibilistic and Probabilistic Systems". In: *Coalgebraic Methods in Computer Science (CMCS 2008)*. Vol. 203. Electronic Notes in Theoretical Computer Science 5. 2008, pp. 131–152. DOI: 10.1016/j.entcs.2008.05.023.

[Jac11] Bart Jacobs. "Probabilities, distribution monads, and convex categories". In: *Theoretical Computer Science* 412.28 (2011), pp. 3323–3336. DOI: 10.1016/j.tcs.2011.04.005. URL: https://arxiv.org/abs/0907.4249.

[Jac15] Bart Jacobs. "From Probability Monads to Commutative Effectuses". In: *Coalgebraic Methods in Computer Science (CMCS 2015)*. Vol. 9608. Lecture Notes in Computer Science. Springer, 2015, pp. 31–48. DOI: 10.1007/978-3-662-49630-5_3. URL: http://www.cs.ru.nl/B.Jacobs/PAPERS/probability-monads.pdf.

[Jac16] Bart Jacobs. *Introduction to Coalgebra: Towards Mathematics of States and Observation*. Vol. 59. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2016. ISBN: 978-1-107-17789-5. DOI: 10.1017/CBO9781316823187.

[JY20] Niles Johnson and Donald Yau. "2-Dimensional Categories". In: arXiv:2002.06055 (June 2020). DOI: 10.48550/arXiv.2002.06055. URL: http://arxiv.org/abs/2002.06055.

[Kel82] G M Kelly. *Basic Concepts of Enriched Category Theory*. en. Cambridge University Press, 1982.

[Koz81] Dexter Kozen. "Semantics of probabilistic programs". In: *Journal of Computer and System Sciences* 22.3 (1981), pp. 328–350.

[KPB18] Neelakantan R. Krishnaswami, Pierre Pratter, and Nick Benton. "Integrating Dependent and Linear Types". In: *Proceedings of the 45th ACM SIGPLAN Symposium on Principles of Programming Languages (POPL)*. 2018. URL: https://www.cl.cam.ac.uk/~nk480/dlnl-paper.pdf.

[Kuh53] H. W. Kuhn. "Extensive games and the problem of information". In: *Contributions to the Theory of Games*. Annals of Mathematics Studies 2.28 (1953), pp. 193–216.

[LP24] Michael Lambert and Evan Patterson. "Cartesian double theories: A double-categorical framework for categorical doctrines". In: *Advances in Mathematics* 444 (May 2024), p. 109630. ISSN: 00018708. DOI: 10.1016/j.aim.2024.109630. URL: https://arxiv.org/abs/2310.05384.

[LHS21] Elena Di Lavore, Jules Hedges, and Pawel Sobocinski. "Compositional modelling of network games". In: *LIPIcs, Volume 183, CSL 2021* 183 (2021), 30:1–30:24. ISSN: 1868-8969. DOI: 10.4230/LIPIcs.CSL.2021.30. URL: https://arxiv.org/abs/2006.03493.

[LM25] Sophie Libkind and David Jaz Myers. "Towards a double operadic theory of systems". In: arXiv:2505.18329 (May 2025). DOI: 10.48550/arXiv.2505.18329. URL: http://arxiv.org/abs/2505.18329.

[Lor21] Fosco Loregian. *(Co)end Calculus*. 1st ed. Cambridge University Press, June 2021. ISBN: 978-1-108-77865-7. DOI: 10.1017/9781108778657. URL: https://www.cambridge.org/core/product/identifier/9781108778657/type/book.

[LR20] Fosco Loregian and Emily Riehl. "Categorical notions of fibration". In: *Expositiones Mathematicae* 38.4 (2020), pp. 496–514. DOI: 10.1016/j.exmath.2019.02.004. URL: https://arxiv.org/abs/1806.06129.

[Lur09] Jacob Lurie. *Higher Topos Theory*. Vol. 170. Annals of Mathematics Studies. Princeton University Press, 2009. ISBN: 978-0-691-14049-0. DOI: 10.1515/9781400830558.

[Lyn+25] Owen Lynch, David Jaz Myers, Eigil Fjeldgren Rischel, and Sam Staton. "Clock systems for stochastic and non-deterministic categorical systems theories". Applied Category Theory 2025. 2025.

[Mas+20] Michael Bahir Maschler, Eilon Solan, Shmuel Zamir, Ziv Hellman, and Mike Borns. *Game theory*. eng. 2nd ed. New York: Cambridge University Press, 2020. ISBN: 978-1-108-49345-1.

[McB16] Conor McBride. "I Got Plenty o' Nuttin'". In: *A List of Successes That Can Change the World*. Vol. 9600. Lecture Notes in Computer Science. Springer, 2016, pp. 207–233. DOI: 10.1007/978-3-319-30936-1_12.

[Mea55] George H. Mealy. "A method for synthesizing sequential circuits". In: *The Bell System Technical Journal* 34.5 (Sept. 1955), pp. 1045–1079. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1955.tb03788.x.

[Mil21] Bartosz Milewski. *PolyLens*. en. Dec. 2021. URL: https://bartoszmilewski.com/2021/12/07/polylens/.

[MMS07] Stefan Milius, Lawrence S. Moss, and Daniel Schwencke. "Generic Trace Semantics via Coinduction". In: *Logical Methods in Computer Science* 3.4 (2007), pp. 1–36. DOI: 10.2168/LMCS-3(4:11)2007. URL: https://arxiv.org/abs/0803.3859.

[MV20] Joe Moeller and Christina Vasilakopoulou. "Monoidal Grothendieck Construction". en. In: *Theory and Applications of Categories* 35.31 (July 2020), pp. 1159–1207.

[Mog91] Eugenio Moggi. "Notions of computation and monads". In: *Information and Computation* 93.1 (1991), pp. 55–92.

[Moo56] Edward F. Moore. "Gedanken-Experiments on Sequential Machines". en. In: Princeton University Press, Mar. 1956, pp. 129–154. ISBN: 978-1-4008-8261-8.

[MP23] Sean Moss and Paolo Perrone. "A category-theoretic proof of the ergodic decomposition theorem". en. In: *Ergodic Theory and Dynamical Systems* 43.12 (Dec. 2023), pp. 4166–4192. ISSN: 0143-3857, 1469-4417. DOI: 10.1017/etds.2023.6.

[Mye23] David Jaz Myers. *Categorical Systems Theory*. en. 2023. URL: http://davidjaz.com/Papers/DynamicalBook.pdf.

[Nas51] John Nash. "Non-Cooperative Games". In: *Annals of Mathematics* 54.2 (1951), pp. 286–295. ISSN: 0003-486X. DOI: 10.2307/1969529.

[Nis+07] Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. en. Cambridge University Press, 2007. ISBN: 978-0-521-87282-9.

[Osb04] Martin J. Osborne. *An Introduction to Game Theory*. en. Oxford University Press, 2004. ISBN: 978-0-19-512896-3.

[OR94] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. en. MIT Press, 1994. ISBN: 978-0-262-65040-3.

[Pan09] Prakash Panangaden. *Labelled Markov Processes*. Imperial College Press, 2009. ISBN: 978-1-84816-392-8. DOI: 10.1142/p595.

[Pav97] Duško Pavlović. "Categorical logic of names and abstraction in action calculi". en. In: *Mathematical Structures in Computer Science* 7.6 (Dec. 1997), pp. 619–637. ISSN: 0960-1295, 1469-8072. DOI: 10.1017/S0960129597002296.

[PP02] Gordon Plotkin and John Power. "Notions of computation determine monads". In: *Foundations of Software Science and Computation Structures: 5th International Conference (FoSSaCS 2002)*. Vol. 2303. Lecture Notes in Computer Science. Springer, 2002, pp. 342–356.

[Pow99] John Power. "Enriched Lawvere Theories". English. In: *Theory and Applications of Categories* 6 (1999), pp. 83–93.

[Ril18] Mitchell Riley. "Categories of optics". In: *arXiv preprint arXiv:1809.00738* (2018).

[Rom21] Mario Román. "Open Diagrams via Coend Calculus". en. In: *Electronic Proceedings in Theoretical Computer Science* 333 (Feb. 2021), pp. 65–78. ISSN: 2075-2180. DOI: 10.4204/EPTCS.333.5.

[SS23] Brandon T. Shapiro and David I. Spivak. "Dynamic Operads, Dynamic Categories: From Deep Learning to Prediction Markets". In: *Electronic Proceedings in Theoretical Computer Science* 380 (Aug. 2023), pp. 183–202. ISSN: 2075-2180. DOI: 10.4204/EPTCS.380.11. URL: https://arxiv.org/abs/2205.03906.

[SL09] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. en. Cambridge University Press, 2009. ISBN: 978-0-521-89943-7.

[Shu09] Michael A. Shulman. "Framed bicategories and monoidal fibrations". en. In: *arXiv:0706.1286 [math]* (Jan. 2009). URL: http://arxiv.org/abs/0706.1286.

[Sil+13] Alexandra Silva, Filippo Bonchi, Marcello M. Bonsangue, and Jan J. M. M. Rutten. "Generalizing Determinization from Automata to Coalgebras". In: *Logical Methods in Computer Science* 9.1 (2013). DOI: 10.2168/LMCS-9(1:9)2013. URL: https://arxiv.org/abs/1202.3256.

[Sim56] H. A. Simon. "Rational choice and the structure of the environment". In: *Psychological Review* 63.2 (1956), pp. 129–138. ISSN: 1939-1471. DOI: 10.1037/h0042769.

[Sok05] Ana Sokolova. "Coalgebraic Analysis of Probabilistic Systems". PhD thesis. Technische Universiteit Eindhoven, 2005. ISBN: 90-386-2810-0. URL: https://research.tue.nl/en/publications/coalgebraic-analysis-of-probabilistic-systems.

[Spi20] David I. Spivak. "Poly: An abundant categorical setting for mode-dependent dynamics". In: arXiv:2005.01894 (June 2020). DOI: 10.48550/arXiv.2005.01894. URL: http://arxiv.org/abs/2005.01894.

[Sri08] S. M. Srivastava. *A Course on Borel Sets*. en. Springer Science & Business Media, Jan. 2008. ISBN: 978-0-387-22767-2.

[St 21] Toby St Clere Smithe. "Cyber Kittens, or Some First Steps Towards Categorical Cybernetics". In: *Electronic Proceedings in Theoretical Computer Science* 333 (2021), pp. 108–124. ISSN: 2075-2180. DOI: 10.4204/EPTCS.333.8. URL: https://arxiv.org/abs/2101.10483.

[Sta17] Sam Staton. "Commutative semantics for probabilistic programming". In: Lecture Notes in Computer Science 10201 (2017), pp. 855–879.

[Ste21] Dario Stein. "Structural foundations for probabilistic programming languages". PhD thesis. University of Oxford, 2021. URL: https://dario-stein.de/thesis.pdf.

[SS24] Dario Stein and Sam Staton. "Probabilistic programming with exact conditions". In: *Journal of the ACM* 71.1 (2024). URL: https://arxiv.org/abs/2312.17141.

[Str74] Ross Street. "Fibrations and Yoneda's lemma in a 2-category". In: *Category Seminar*. Ed. by G. M. Kelly. Vol. 420. Lecture Notes in Mathematics. Springer, 1974, pp. 104–133. DOI: 10.1007/BFb0063102.

[Str80] Ross Street. "Fibrations in bicategories". In: *Cahiers de Topologie et Géométrie Différentielle Catégoriques* 21.2 (1980), pp. 111–160. URL: http://www.numdam.org/item/CTGDC_1980__21_2_111_0/.

[V N28] J. V. Neumann. "Zur Theorie der Gesellschaftsspiele". de. In: *Mathematische Annalen* 100.1 (Dec. 1928), pp. 295–320. ISSN: 0025-5831, 1432-1807. DOI: 10.1007/BF01448847.

[Vák15] Matthijs Vákár. "Syntax and Semantics of Linear Dependent Types". In: arXiv:1405.0033 (Jan. 2015). URL: https://arxiv.org/abs/1405.0033.

[Vák17] Matthijs Vákár. "In search of effectful dependent types". PhD thesis. University of Oxford, 2017. URL: https://arxiv.org/abs/1706.07997.

[Ver23] Pietro Vertechi. "Dependent Optics". In: *Electronic Proceedings in Theoretical Computer Science* 380 (Aug. 2023), pp. 128–144. ISSN: 2075-2180. DOI: 10.4204/EPTCS.380.8. URL: https://arxiv.org/abs/2204.09547.

[VC22] Andre Videla and Matteo Capucci. "Lenses for Composable Servers". In: arXiv:2203.15633 (Mar. 2022). DOI: 10.48550/arXiv.2203.15633. URL: http://arxiv.org/abs/2203.15633.

[Vis04] Angelo Vistoli. "Notes on Grothendieck topologies, fibered categories and descent theory". In: arXiv:math/0412512 (2004). Published in: *Fundamental Algebraic Geometry*, Math. Surveys Monogr., vol. 123, AMS, 2005. URL: https://arxiv.org/abs/math/0412512.

[VM07] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior*. eng. 60. anniversary ed., 4. print., and 1. paperb. print. A Princeton classic edition. Princeton, NJ: Princeton University Press, 2007. ISBN: 978-0-691-13061-3.