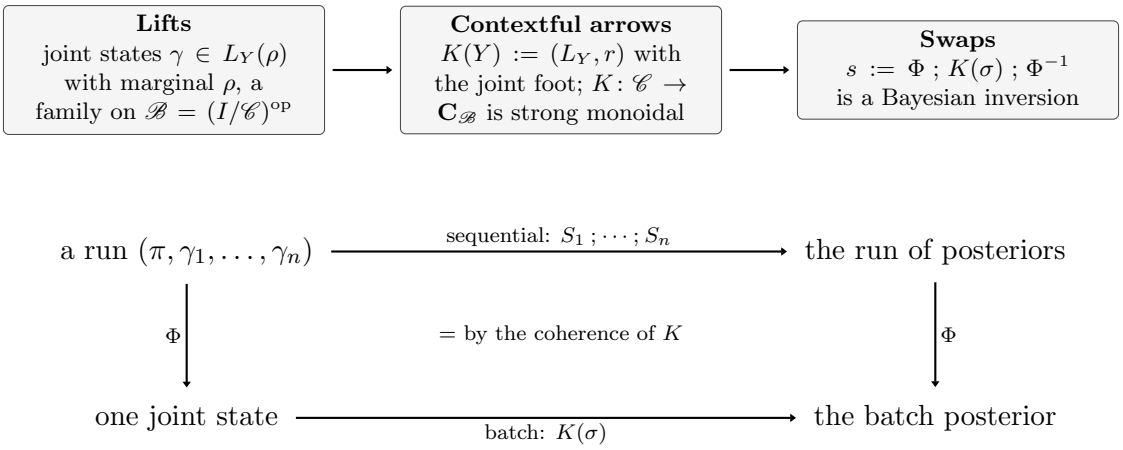
hexagon:2611.00985v1

BAYESIAN UPDATING WITH CONTEXTFUL ARROWS

KOTARO KAMIYA

Abstract. Many formulations of Bayesian updating with Markov categories are known, among them the dagger on probability spaces, Bayesian lenses and their dependent version, the recursion of the Bayes filter and conditional composition. A question common to many of them is how to make updating compositional, so that a posterior becomes the next prior; the author's earlier work also raised it, as the relation between sequential and batch updating. This paper gives one answer by composing contextful arrows. The states of a Markov category $\mathscr{C}$ form a base $\mathscr{B} = (I/\mathscr{C})^{\mathrm{op}}$, and their lifts, the joint states with a given marginal, form a presheaf L on $\mathscr{B}$. A foot makes L a contextful arrow $\mathscr{B} \nrightarrow \mathscr{B}$: the marginal foot recovers the chain rule of Bayesian lenses, and the joint foot makes each joint state the next state. With the joint foot, the lifts to an object Y give a strong monoidal functor K from $\mathscr{C}$ to the monoidal category of contextful arrows on $\mathscr{B}$, whose runs are conditional composites. Through K the symmetry of $\mathscr{C}$ becomes a swap of two lifts, and this swap is a Bayesian inversion. The coherence of K then gives the agreement of sequential and batch updating and recovers the recursion of the Bayes filter. These facts are known; here they follow from one strong monoidal functor.



The paper in one picture. The lifts of states form contextful arrows, and with the joint foot they give a strong monoidal functor K whose swaps are Bayesian inversions. A run of K is sequential updating; its comparison Φ collapses it into one joint state, and the square, the coherence of K (Lemma 5.3), says that sequential and batch updating agree (Theorem 5.5).

Date: November 2026.

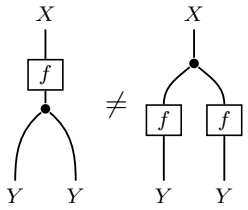
Contents

1. Introduction 2
2. Preliminaries 4
3. Lifts 7
4. The strong monoidal functor of lifts 8
5. Swaps and updating 10
6. The reading in probability spaces 13
Outlook 14
Acknowledgements 14
References 14

1. Introduction

1.1. The question

Let $\mathscr{C}$ be a Markov category, $\pi\colon I \to X$ a state and $f\colon X \to Y$ a channel. A Bayesian inversion $f^\dagger_\pi\colon Y \to X$ depends on the prior π, and inversions compose by the chain rule $(f\,;g)^\dagger_\pi \sim_{\pi;f;g} g^\dagger_{\pi;f}\,;f^\dagger_\pi$ up to almost sure equality [CJ19, Smi20]. Many formulations of Bayesian updating organise this rule, among them the dagger on the category $\mathrm{PS}(\mathscr{C})$ of probability spaces [KW21, Fri20], Bayesian lenses and their dependent version [Smi20, BH22, BHS23], the recursion of the Bayes filter [FKMSW24] and conditional composition [DLRS25]. A question common to many of them is how to make updating compositional, so that a posterior becomes the next prior, and how sequential updating, one datum at a time, then relates to batch updating on all data at once; Kamiya and Welliaveetil left it open beyond finite examples [KW21]. In a Bayesian lens $(f, f^\sharp)$, for instance, the prior is the index π of the backward part $f^\sharp_\pi$, fixed before the data and moved only by $\pi \mapsto \pi\,;g$ along a channel g, so the posterior $\delta_y\,;f^\dagger_\pi$, which is not a pushforward of π, never becomes the index of the next backward part. Ordinary lenses do not help, because composing lenses copies the input, and



in general: copying is natural only for deterministic channels.

The semicolons of this paper treat the prior differently. The black semicolon ; composes channels: it fixes π and pushes it forward to $\pi\,;f$. The white semicolon ⨟ is the dependent sum, of channels (Section 2.1) and of families (Section 2.2): it keeps what the first part produced and lets the second part read it. The composite ⨟ of contextful arrows (Definition 2.6) reads the second at the output of the first, its foot. When the foot returns the joint state, each joint state is passed on as the next state, and this is sequential updating.

1.2. Contributions

Section 3 takes the states of $\mathscr{C}$ as the base $\mathscr{B} := (I/\mathscr{C})^{\mathrm{op}}$ and the lifts of a state ρ to Y, the states with marginal ρ, as a family L_Y on $\mathscr{B}$. A Bayesian lens whose backward part is

topic	known	here
Bayesian lenses	compose by the chain rule [Smi20, Thm. 5.2], [Smi22, Thm. 4.3.14]; inversion is not monoidal [Smi22, Rem. 4.3.18]	the marginal foot m (Proposition 3.3, Table 2); K sends $\otimes$ to $\fatsemi$ (Theorem 4.2), not to a parallel product
probability spaces	the dagger on $\mathrm{PS}(\mathscr{C})$ [Fri20, Rem. 13.10], [KW21]; disintegrating in the other direction [CJ19, Sec. 3, Cor. 5.6]; pushing a coupling along the permutation [CDDG17, Sec. 3.3, Thm. 2]	the contextad of disintegrations (Proposition 6.1); the swap (Proposition 5.2)
conditional composition	keeps the intermediate outputs, associative and unital [DLRS25, Def. 2.3, Prop. 2.5]	the runs of K (Proposition 4.3)
sequential and batch	double conditionals are conditionals [Fri20, Lemma 11.11], and conditioning commutes with composition [Fri20, Lemma 11.12]; measure-theoretically for supervised models [Lê25, Thm. 3.1]; left open beyond finite examples [KW21, Secs. 4.1–4.2, Rem. 4.6]; finite Pearl updates [Jac19, Lemma 4.1(2), Prop. 4.3(3)]	the coherence of K (Theorem 5.5)
Bayes filter	its recursion computes the batch conditional [FKMSW24, Prop. 4.2, 4.7]	recovered by swaps (Remark 5.8)
dependent Bayesian lenses	backward objects depend on the state [BH22, Sec. 4, Prop. 15], [BHS23, Def. 15, Prop. 21]; joint priors set aside [Smi22, Rem. 5.2.13]; updating in $\mathrm{PS}(\mathscr{C})$ is not a process [BHS23, Rem. 23]	not as a category; their canonical inversion (Remark 6.3); the joint foot r: a lift is again a state
contextads	contextful arrows [CM24]	on one base, right-colax cells

TABLE 1. Known formulations and results, and where they arise in this paper.

an inversion is a section of lifts (Proposition 3.3), and a foot makes L_Y a contextful arrow $\mathscr{B} \nrightarrow \mathscr{B}$ in the sense of Capucci and Myers [CM24]: the marginal foot m gives the chain rule, the joint foot r the next state. Section 4 shows that $K(Y) := (L_Y, r)$ is a strong monoidal functor

$$K \colon (\mathscr{C}, \otimes, I) \longrightarrow (\mathbf{C}_{\mathscr{B}}, \fatsemi, (\mathbf{1}_{\mathscr{B}}, \varepsilon))$$

into the contextful arrows on $\mathscr{B}$ with right-colax cells (Theorem 4.2). Section 5 conjugates $K(\sigma)$ by the comparison Φ of K: the swap is a Bayesian inversion (Proposition 5.2), and the coherence of K gives the agreement of sequential and batch updating (Theorem 5.5). Section 6 reads this in $\mathrm{PS}(\mathscr{C})$, where the disintegrations form a contextad on $\mathrm{PS}(\mathscr{C})$ (Proposition 6.1), which carries the sequential updates as a dependent sum of predictives and posteriors (Proposition 6.4). Several of the formulations above arise from this structure (Table 1): Bayesian lenses with the marginal foot, probability spaces and their dagger as the contextad of disintegrations, and conditional composition as the runs of K (Proposition 4.3); the recursion of the Bayes filter is recovered (Remark 5.8). Dependent

Bayesian lenses do not arise as a category, since they widen the backward part whereas contextful arrows move the index; only their canonical inversion does (Remark 6.3). The paper concerns the structure that yields known results, in the manner of abstract nonsense.

1.3. Related work

Table 1 lists the known formulations and where they arise in this paper. The equation of Theorem 5.5 and the content of Proposition 5.2 are known; the paper derives both from one strong monoidal functor.

2. Preliminaries

2.1. Markov categories

Composition $f\,;g$ is diagrammatic, f then g, also for functors and natural transformations. String diagrams of channels run top to bottom and draw a state as a triangle, a channel as a box, and copy and delete as dots. In a symmetric monoidal category an equation holds if and only if it holds up to isomorphism of diagrams [JS91, Sel11], and every Markov category is comonoid equivalent to a strict one [Fri20, Thm. 10.17]. Hence the diagrams prove equations in every Markov category, and the propositions below take $\mathscr{C}$ strict, which loses no generality.

Definition 2.1 (Markov categories, [Fri20])**.** A *Markov category* is a symmetric monoidal category $(\mathscr{C}, \otimes, I)$ with a commutative comonoid $(\mathrm{copy}_X, \mathrm{del}_X)$ on every object, compatible with $\otimes$, such that del is natural. Its morphisms are *channels*, a channel $\pi\colon I \to X$ is a *state* on X, and $\langle f_1, \dots, f_n\rangle := \mathrm{copy}^{(n)}\,;(f_1 \otimes \cdots \otimes f_n)$ for channels with a common domain. The symmetry is σ.

Naturality of del makes I terminal. For a state γ on $A \otimes Y$, its *marginal* on A is $\gamma\,;(\mathrm{id}_A \otimes \mathrm{del}_Y)$.

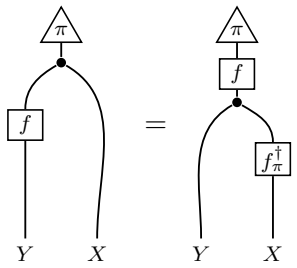
Definition 2.2 (Almost sure equality and conditionals, [Fri20])**.** Let $\mathscr{C}$ be a Markov category.

(i) *Almost sure equality.* For a state π on X and $g, g'\colon X \to Z$, $g \sim_\pi g'$ when $\pi\,;\langle \mathrm{id}_X, g\rangle = \pi\,;\langle \mathrm{id}_X, g'\rangle$.

(ii) *Conditionals.* For a state ω on $A \otimes Y$ with marginal ρ on A, a conditional of Y given A is $d\colon A \to Y$ with $\omega = \rho\,;\langle \mathrm{id}_A, d\rangle$, and $\mathscr{C}$ *has conditionals* when every such ω has one. Conditionals of other outputs are taken after the symmetry that brings the conditioned outputs to the left.

Definition 2.3 (Inversion at a state)**.** Let $\pi\colon I \to X$ and $f\colon X \to Y$. An *inversion of f at π* is $f^\dagger_\pi\colon Y \to X$ with

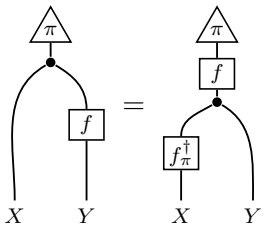
$$\pi\,;\langle f, \mathrm{id}_X\rangle = \pi\,;f\,;\langle \mathrm{id}_Y, f^\dagger_\pi\rangle, \qquad \text{that is,}$$



(2.1)

An inversion of f at π is the Bayesian inversion of Cho and Jacobs and of Fritz [CJ19, Sec. 3], [Fri20, Rem. 13.10]: it is a conditional of X given Y for the joint state $\pi \mathbin{;} \langle f, \mathrm{id}_X\rangle$. Hence any two inversions of f at π agree $\sim_{\pi;f}$ [Fri20, Prop. 13.7], and deleting Y in (2.1) gives $\pi \mathbin{;} f \mathbin{;} f^\dagger_\pi = \pi$. After the symmetry of the outputs, (2.1) is the conditionals equation

$$\pi \mathbin{;} \langle \mathrm{id}_X, f\rangle = (\pi \mathbin{;} f) \mathbin{;} \langle f^\dagger_\pi, \mathrm{id}_Y\rangle, \qquad \text{that is,} \qquad$$



(2.2)

is the root of this paper: one joint state, two conditionals.

Definition 2.4 (Probability spaces, [Fri20])**.** Let $\mathscr{C}$ be a Markov category with conditionals, hence causal [Fri20, Prop. 11.34]. Its probability spaces form the category $\mathrm{PS}(\mathscr{C})$ whose objects are the pairs (X, π) of an object and a state on it, and whose morphisms $[h]_\pi \colon (X, \pi) \to (Y, \pi \mathbin{;} h)$ are the classes of channels $h \colon X \to Y$ modulo $\sim_\pi$, with $[h]_\pi \mathbin{;} [k]_{\pi;h} = [h \mathbin{;} k]_\pi$ [Fri20, Lemma 13.4, Rem. 13.6, Prop. 13.9(a)].

Proposition 2.5 (Inversion with supports)**.** *Let every state π on X have a chosen support, a pair $X_\pi \xrightarrow{i_\pi} X \xrightarrow{j_\pi} X_\pi$ with $i_\pi \mathbin{;} j_\pi = \mathrm{id}$ such that $h \sim_\pi h'$ if and only if $i_\pi \mathbin{;} h = i_\pi \mathbin{;} h'$* [Fri20, Def. 13.20], [BH22, Def. 4]*. Then $\mathit{Supp}(X, \pi) := X_\pi$ and $\mathit{Supp}[h]_\pi := i_\pi \mathbin{;} h \mathbin{;} j_{\pi;h}$ define a functor $\mathit{Supp} \colon \mathrm{PS}(\mathscr{C}) \to \mathscr{C}$, which sends the class $[f^\dagger_\pi]_{\pi;f}$ of an inversion to the inversion $i_{\pi;f} \mathbin{;} f^\dagger_\pi \mathbin{;} j_\pi$ with support of Braithwaite and Hedges* [BH22, Prop. 15]*. Hence every equation between classes in $\mathrm{PS}(\mathscr{C})$ holds strictly between supports.*

Proof. *Supp* is well defined because $i_\pi \mathbin{;} h$ depends only on $[h]_\pi$. For $h \colon X \to Y$, $k \colon Y \to Z$ and $\rho := \pi \mathbin{;} h$, the support property turns $i_\rho \mathbin{;} r_\rho \mathbin{;} i_\rho = i_\rho$ into $[r_\rho \mathbin{;} i_\rho]_\rho = [\mathrm{id}]_\rho$, hence $[h \mathbin{;} r_\rho \mathbin{;} i_\rho \mathbin{;} k]_\pi = [h \mathbin{;} k]_\pi$, and applying $i_\pi \mathbin{;} - \mathbin{;} r_{\rho;k}$ gives $\mathit{Supp}[h]_\pi \mathbin{;} \mathit{Supp}[k]_\rho = \mathit{Supp}[h \mathbin{;} k]_\pi$, while identities go to $i_\pi \mathbin{;} r_\pi = \mathrm{id}$. The value at $[f]^\dagger_\pi$ is the definition. □

The *dependent sum* of channels $c \colon B \to Y$ and $d \colon B \otimes Y \to Z$ is

$$c \mathbin{⨟} d := \langle \mathrm{id}_B, c\rangle \mathbin{;} \langle \mathrm{del}_B \otimes \mathrm{id}_Y, d\rangle \colon B \to Y \otimes Z,$$

which keeps the output of c and feeds both B and it to d; it is the conditional composition of [DLRS25, Def. 2.3] and satisfies

$$\langle \mathrm{id}_B, c \mathbin{⨟} d\rangle = \langle \mathrm{id}_B, c\rangle \mathbin{;} \langle \mathrm{id}_{B\otimes Y}, d\rangle, \tag{2.3}$$

since both sides copy B, apply c once and give copies of B and Y to d; this identity gives its unit and associativity laws [DLRS25, Prop. 2.5]. The proofs below use three rules for a state ω [Fri20, Lemma 13.4, Rem. 13.6, Lemma 13.5]: (a) $a \sim_\omega b$ implies $a \mathbin{;} k \sim_\omega b \mathbin{;} k$, (b) $a \sim_{\omega;h} b$ implies $h \mathbin{;} a \sim_\omega h \mathbin{;} b$, and (c) $a \sim_\omega b$ and $a' \sim_\omega b'$ imply $\langle a, a'\rangle \sim_\omega \langle b, b'\rangle$.

2.2. Contextful arrows on one base

A *family* on a category $\mathscr{B}$ is here a presheaf $F \colon \mathscr{B}^{\mathrm{op}} \to \mathrm{Set}$ of large sets, and $\mathscr{B} \odot F$ is its category of elements: objects (b, x) with $x \in F(b)$, and a morphism $(b', x') \to (b, x)$ is $u \colon b' \to b$ with $F(u)(x) = x'$. Its projection to $\mathscr{B}$ is a discrete fibration. With the dependent sum $(F \mathbin{⨟} G)(b) := \coprod_{x \in F(b)} G(b, x)$ of F and a family G on $\mathscr{B} \odot F$, the families on all categories form a contextad on **Cat** in the sense of Capucci and Myers [CM24]. This paper uses only its contextful arrows on one base.

Definition 2.6 (Contextful arrows on one base)**.** Let $\mathscr{B}$ be a category. The category $\mathbf{C}_{\mathscr{B}}$ has the following parts.

(i) *Objects. Contextful arrows* $(F, f)\colon \mathscr{B} \nrightarrow \mathscr{B}$: a family F on $\mathscr{B}$ and any functor $f\colon \mathscr{B} \odot F \to \mathscr{B}$, its *foot*.

(ii) *Cells.* A cell $(\theta, \varphi)\colon (F, f) \Rightarrow (F', f')$ is a natural transformation $\theta\colon F \Rightarrow F'$ together with a natural transformation $\varphi\colon (\mathscr{B} \odot \theta)\,;\, f' \Rightarrow f$, its *foot cell*, where $\mathscr{B} \odot \theta$ sends (b, x) to $(b, \theta_b x)$. Cells compose by $(\theta, \varphi)\,;\,(\theta', \varphi') = (\theta\,;\,\theta', ((\mathscr{B} \odot \theta)\varphi')\,;\,\varphi)$.

(iii) *Composite.* $(F, f) \fatsemi (G, g) := \big(F \fatsemi f^*G,\ (b, x, y) \mapsto g(f(b, x), y)\big)$, where $f^*G := f^{\mathrm{op}}\,;\,G$ is the base change of G along f, so that $(F \fatsemi f^*G)(b) = \{(x, y) \mid x \in F(b),\ y \in G(f(b, x))\}$ and a morphism $u\colon b' \to b$ acts by $(x, y) \mapsto (F(u)x,\ G(f(u))y)$ for the morphism $f(u)\colon f(b', F(u)x) \to f(b, x)$.

(iv) *Unit.* $(\mathbf{1}_{\mathscr{B}}, \varepsilon)$ with $\mathbf{1}_{\mathscr{B}}(b) = \{*\}$ and $\varepsilon(b, *) = b$.

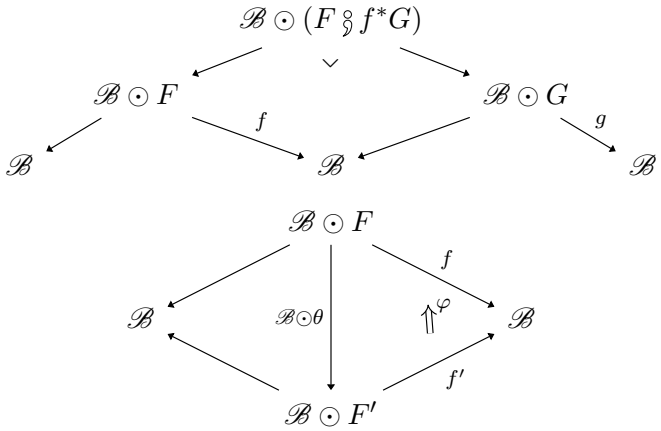


FIGURE 1. In **Cat**. Top: the composite $(F, f) \fatsemi (G, g)$, whose elements (b, x, y) form the pullback of f and the projection of $\mathscr{B} \odot G$. Bottom: a cell $(\theta, \varphi)\colon (F, f) \Rightarrow (F', f')$, strict on the projections and with $\varphi\colon (\mathscr{B} \odot \theta)\,;\, f' \Rightarrow f$ on the feet.

The cells are right-colax (Figure 1): the family is mapped on the nose and the foot up to φ, as for the 1-cells of the tricategory of spans of Capucci and Myers [CM24, Sec. 1.3]. Their double category of contextful arrows asks φ to be an identity; the strong monoidal functor of Section 4 needs the right-colax cells, because its action on channels moves the foot.

Lemma 2.7 (A monoidal category)**.** $(\mathbf{C}_{\mathscr{B}}, \fatsemi, (\mathbf{1}_{\mathscr{B}}, \varepsilon))$ *is a monoidal category whose associator and unitors are the rebracketings* $((x, y), z) \leftrightarrow (x, (y, z))$ *and* $(*, x) \leftrightarrow x \leftrightarrow (x, *)$ *with identity foot cells. On cells,*

$$(\theta, \varphi) \fatsemi (\zeta, \psi) := \big((x, y) \mapsto (\theta x,\ \zeta\, G(\varphi_{(b,x)})y),\quad \psi_{(f'(b,\theta x),\, G(\varphi)y)}\,;\, g(\varphi_{(b,x)}, \mathrm{id})\big)$$

for $(\theta, \varphi)\colon (F, f) \Rightarrow (F', f')$ *and* $(\zeta, \psi)\colon (G, g) \Rightarrow (G', g')$*, where* $G(\varphi_{(b,x)})\colon G(f(b, x)) \to G(f'(b, \theta x))$ *restricts along* $\varphi_{(b,x)}\colon f'(b, \theta x) \to f(b, x)$ *and* $g(\varphi_{(b,x)}, \mathrm{id})$ *is* g *applied to the morphism* $(f'(b, \theta x), G(\varphi)y) \to (f(b, x), y)$ *of* $\mathscr{B} \odot G$*.*

Proof. Composite. $F \fatsemi f^*G$ is a family since F, G and f are functors, and $(b, x, y) \mapsto g(f(b, x), y)$ is a functor on its elements, so $\fatsemi$ is well defined on objects. *Cells.* The first component of $(\theta, \varphi) \fatsemi (\zeta, \psi)$ is natural in b because θ, ζ and φ are natural and G is a functor, and it lands in $F' \fatsemi f'^*G'$ since ζ maps $G(f'(b, \theta x))$ to $G'(f'(b, \theta x))$. Its second

component goes from $g'(f'(b,\theta x),\zeta G(\varphi)y)$ through $g(f'(b,\theta x),G(\varphi)y)$ to $g(f(b,x),y)$ and is natural as a composite of natural transformations. *Functoriality.* For identity cells both components are identities. For composable cells, both sides of the interchange law apply $\theta\,;\theta'$ to x and, to y, restrict along the composite foot cell and then apply $\zeta\,;\zeta'$; this agrees because restriction along φ commutes with ζ by naturality of ζ, and the foot cells agree because g is a functor on $\mathscr{B}\odot G$. *Coherence.* Both bracketings of a triple composite have the elements (x,y,z) with $y\in G(f(b,x))$ and $z\in H(g(f(b,x),y))$, the same foot and the same action of morphisms, so the rebracketings are isomorphisms of families with identity foot cells, natural in cells by the formula above. The pentagon and triangle hold because they compare rebracketings of the same tuples. □

A monoid in $\mathbf{C}_{\mathscr{B}}$ is a contextful arrow (F,f) with cells $\eta\colon(\mathbf{1}_{\mathscr{B}},\varepsilon)\Rightarrow(F,f)$ and $\mu\colon(F,f)\fatsemi(F,f)\Rightarrow(F,f)$ satisfying the monoid laws.

3. Lifts

From now on $\mathscr{C}$ is a strict Markov category with conditionals.

Definition 3.1 (States and lifts)**.**

(i) *The base.* $\mathscr{B}:=(I/\mathscr{C})^{\mathrm{op}}$. Its objects are the states $\rho\colon I\to A$, and a morphism $\rho'\to\rho$ from a state ρ' on A' is a channel $u\colon A\to A'$ with $\rho\,;u=\rho'$.

(ii) *Lifts.* For an object Y, the *lifts* of ρ to Y are

$$L_Y(\rho):=\{\gamma\colon I\to A\otimes Y\mid \gamma\,;(\mathrm{id}_A\otimes \mathrm{del}_Y)=\rho\},$$

and a morphism $u\colon\rho'\to\rho$ acts by $L_Y(u)(\gamma):=\gamma\,;(u\otimes\mathrm{id}_Y)$.

(iii) *All lifts.* $L(\rho):=\{(Y,\gamma)\mid Y\in\mathscr{C},\ \gamma\in L_Y(\rho)\}=\coprod_{Y\in\mathscr{C}}L_Y(\rho)$, with the same action.

$L_Y(u)$ lands in $L_Y(\rho')$ because $(u\otimes\mathrm{id}_Y)\,;(\mathrm{id}_{A'}\otimes\mathrm{del}_Y)=(\mathrm{id}_A\otimes\mathrm{del}_Y)\,;u$, so L_Y and L are families on $\mathscr{B}$. A lift fixes one marginal only: fixing also the marginal ν on Y gives the Fréchet class of couplings of ρ and ν [Fré51], and $L_Y(\rho)$ is the disjoint union of these classes over the states ν on Y. The sum in L runs over the objects of $\mathscr{C}$ only; channels between the Y act in Section 4.

Proposition 3.2 (Lifts as kernels)**.** *Let ρ be a state on A. Then $[d]_\rho\mapsto\rho\,;\langle\mathrm{id}_A,d\rangle$ is a bijection from the classes of channels $d\colon A\to Y$ modulo $\sim_\rho$ onto $L_Y(\rho)$* [CJ19, Thm. 5.5], [Fri20, Prop. 13.9(b)].

Proposition 3.3 (Bayesian lenses as sections)**.** *Let $f\colon X\to Y$ be a channel.*

(i) A lift. *At a state π on X, the lens component $(f,f_\pi^\dagger)$ of a Bayesian lens whose backward part is an inversion determines the lift $\pi\,;\langle\mathrm{id}_X,f\rangle=(\pi\,;f)\,;\langle f_\pi^\dagger,\mathrm{id}_Y\rangle\in L_Y(\pi)$, and every almost sure choice of $f_\pi^\dagger$ gives this lift.*

(ii) A section. *Bundled over all states, the lens $(f,f^\sharp)$ with $f_\pi^\sharp=f_\pi^\dagger$ determines $(\pi\,;\langle\mathrm{id}_X,f\rangle)_\pi\in\prod_{\pi\in\mathrm{Hom}_{\mathscr{C}}(I,X)}L_Y(\pi)$, a section of the projection $\coprod_\pi L_Y(\pi)\to\mathscr{C}(I,X)$.*

Proof. (i) is (2.2), and its right-hand side depends only on the class of $f_\pi^\dagger$ modulo $\sim_{\pi;f}$ by Definition 2.2(i). (ii) follows by taking (i) at every π. □

Definition 3.4 (Two feet)**.** The family L_Y has two feet $\mathscr{B}\odot L_Y\to\mathscr{B}$:

(i) *the marginal foot* $m(\rho,\gamma):=\gamma\,;(\mathrm{del}_A\otimes\mathrm{id}_Y)$, a state on Y, with $m(u):=\mathrm{id}_Y$;

(ii) *the joint foot* $r(\rho,\gamma):=\gamma$, a state on $A\otimes Y$, with $r(u):=u\otimes\mathrm{id}_Y$.

Both are functors: $m(u)$ is a morphism because $\gamma\,;(u\otimes \mathrm{id}_Y)\,;(\mathrm{del}_{A'}\otimes \mathrm{id}_Y) = \gamma\,;(\mathrm{del}_A\otimes \mathrm{id}_Y)$, and $r(u)$ is one by the definition of $L_Y(u)$. Hence (L_Y, m) and (L_Y, r) are contextful arrows $\mathscr{B} \nrightarrow \mathscr{B}$ (Figure 2), and Table 2 compares their composites at π for $\gamma_1 := \pi\,;\langle \mathrm{id}_X, f\rangle$, $\gamma_2' := (\pi\,; f)\,;\langle \mathrm{id}_Y, g\rangle$ and $\gamma_2 := \gamma_1\,;\langle \mathrm{id}_{X\otimes Y}, \mathrm{del}_X\otimes g\rangle$.

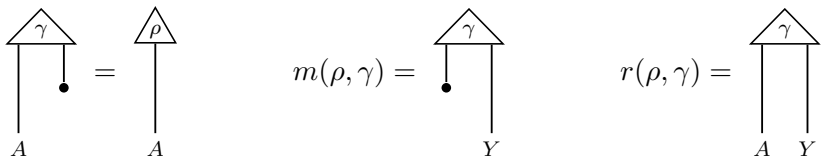


FIGURE 2. In $\mathscr{C}$. A lift $\gamma \in L_Y(\rho)$ has marginal ρ (left); its marginal foot deletes A, its joint foot is γ itself (right).

	marginal foot m	joint foot r
foot at π	$\pi\,; f$	γ_1
element at π	(γ_1, γ_2')	(γ_1, γ_2)
to one lift	glue on Y, delete Y	keep the last lift
result	$\pi\,;\langle \mathrm{id}_X, f\,; g\rangle \in L_Z(\pi)$	$\gamma_2 \in L_{Y\otimes Z}(\pi)$
inversion	$(f\,; g)^\dagger_\pi \sim_{\pi;f;g} g^\dagger_{\pi;f}\,; f^\dagger_\pi$, at each step	one inversion, at the end

TABLE 2. The composites $(L_Y, m) \fatsemi (L_Z, m)$ and $(L_Y, r) \fatsemi (L_Z, r)$ at a state π on X, for $f\colon X \to Y$ and $g\colon Y \to Z$. Gluing (γ_1, γ_2') on Y gives $\pi\,;\langle \mathrm{id}_X, f\,;\langle \mathrm{id}_Y, g\rangle\rangle$, the dependent sum of Proposition 6.2(ii), and deleting Y reads off the chain rule of Bayesian lenses (Proposition 6.2(iii)).

4. The strong monoidal functor of lifts

Definition 4.1 (The functor K)**.** For an object Y, $K(Y) := (L_Y, r)\colon \mathscr{B} \nrightarrow \mathscr{B}$. For a channel $g\colon Y \to Y'$, $K(g)\colon K(Y) \Rightarrow K(Y')$ is the cell with $\theta_g(\gamma) := \gamma\,;(\mathrm{id}_A \otimes g)$ and foot cell $(\varphi_g)_{(\rho,\gamma)} := \mathrm{id}_A \otimes g\colon r(\rho, \theta_g\gamma) \to r(\rho, \gamma)$, the morphism of $\mathscr{B}$ given by the channel $\mathrm{id}_A \otimes g\colon A \otimes Y \to A \otimes Y'$.

θ_g is natural since $(u \otimes \mathrm{id}_Y)\,;(\mathrm{id}_{A'} \otimes g) = (\mathrm{id}_A \otimes g)\,;(u \otimes \mathrm{id}_{Y'})$, $\theta_g(\gamma)$ is again a lift of ρ by naturality of del, and φ_g is a morphism $\gamma\,;(\mathrm{id}_A \otimes g) \to \gamma$ of $\mathscr{B}$ by its definition. The foot cell is not an identity, which is why $\mathbf{C}_{\mathscr{B}}$ has right-colax cells.

Theorem 4.2 (Lifts form a strong monoidal functor)**.** *The data of Table 3 form a strong monoidal functor $K\colon (\mathscr{C}, \otimes, I) \to (\mathbf{C}_{\mathscr{B}}, \fatsemi, (\mathbf{1}_{\mathscr{B}}, \varepsilon))$ with comparisons Φ and Φ_0, and a monoid $((L, r), \eta, \mu)$ in $\mathbf{C}_{\mathscr{B}}$ whose product μ restricts on the summands to Φ.*

Proof. Functor. $K(\mathrm{id}_Y)$ is the identity cell and $K(g)\,;K(g') = K(g\,; g')$, since $(\mathrm{id}_A \otimes g)\,;(\mathrm{id}_A \otimes g') = \mathrm{id}_A \otimes (g\,; g')$ on both components. *Comparison on $\otimes$.* An element of $(L_X \parr r^*L_Y)(\rho)$ is a pair (γ, γ') with $\gamma \in L_X(\rho)$ and $\gamma' \in L_Y(r(\rho, \gamma)) = L_Y(\gamma)$, that is, a state γ' on $A \otimes X \otimes Y$ with marginal γ on $A \otimes X$, whose marginal on A is ρ. Hence $\gamma' \in L_{X\otimes Y}(\rho)$, and γ is recovered from γ' as its marginal, so $\Phi_{X,Y}$ is a bijection at each ρ. It is natural in ρ because both sides act by $-\,;(u \otimes \mathrm{id}_{X\otimes Y})$, and the feet agree: the

	datum	on lifts at ρ	foot cell
object	$K(Y)$	$L_Y(\rho)$, foot r	
channel $g\colon Y \to Y'$	$K(g)$	$\gamma \mapsto \gamma\,;(\mathrm{id}_A \otimes g)$	$\mathrm{id}_A \otimes g$
symmetry	$K(\sigma_{X,Y})$	$\gamma \mapsto \gamma\,;(\mathrm{id}_A \otimes \sigma_{X,Y})$	$\mathrm{id}_A \otimes \sigma_{X,Y}$
comparison on $\otimes$	$\Phi_{X,Y}$	$(\gamma,\gamma') \mapsto \gamma'$	identity
its inverse	$\Phi^{-1}_{X,Y}$	$\gamma' \mapsto (\gamma'\,;(\mathrm{id}_{A\otimes X} \otimes \mathrm{del}_Y),\gamma')$	identity
comparison on I	Φ_0	$* \mapsto \rho \in L_I(\rho) = \{\rho\}$	identity
unit of (L,r)	η	$* \mapsto (I,\rho)$	identity
product of (L,r)	μ	$((X,\gamma),(Y,\gamma')) \mapsto (X\otimes Y,\gamma')$	identity

TABLE 3. The strong monoidal functor $K\colon (\mathscr{C},\otimes,I) \to (\mathbf{C}_{\mathscr{B}},\fatsemi,(\mathbf{1}_{\mathscr{B}},\varepsilon))$ and the monoid of lifts, on lifts $\gamma \in L_X(\rho)$ and $\gamma' \in L_Y(\gamma)$ of a state ρ on A.

composite foot sends (ρ,γ,γ') to $r(\gamma,\gamma') = \gamma' = r(\rho,\gamma')$. *Naturality.* For $g\colon X \to X'$ and $h\colon Y \to Y'$, the cell $K(g) \fatsemi K(h)$ of Lemma 2.7 sends (γ,γ') to $(\gamma\,;(\mathrm{id}\otimes g),\gamma'\,;(\mathrm{id}\otimes g\otimes \mathrm{id}_Y)\,;(\mathrm{id}\otimes\mathrm{id}\otimes h))$, since restriction along (φ_g) is $L_Y(\mathrm{id}_A\otimes g)$. Its image under Φ is $\gamma'\,;(\mathrm{id}_A\otimes g\otimes h) = \theta_{g\otimes h}(\gamma')$, and its foot cell is $\mathrm{id}_A\otimes g\otimes h$, which is that of $K(g\otimes h)$. *Unit.* $L_I(\rho) = \{\rho\}$ since $\mathrm{del}_I = \mathrm{id}_I$, and the foot of $K(I)$ at ρ is $\rho = \varepsilon(\rho,*)$. *Coherence.* Both sides of the associativity axiom send a triple $(\gamma,\gamma',\gamma'')$ of consecutive lifts to γ'', and both sides of the unit axioms send $(*,\gamma)$ and $(\gamma,*)$ to γ, since $\mathscr{C}$ is strict. *Monoid.* (L,r) is the sum of the $K(Y)$ over the objects of $\mathscr{C}$, and $\fatsemi$ preserves these sums in each variable because the elements of a composite are pairs. Hence η and μ are cells, and the monoid laws are those of Φ and Φ_0 on summands, together with the associativity and unit of $\otimes$ on objects. □

An element of $K(Y_1)\fatsemi\cdots\fatsemi K(Y_n)$ at ρ is a *run*: a chain $(\gamma_1,\dots,\gamma_n)$ of lifts with $\gamma_k \in L_{Y_k}(\gamma_{k-1})$ and $\gamma_0 = \rho$, and the n-fold comparison Φ sends it to $\gamma_n \in L_{Y_1\otimes\cdots\otimes Y_n}(\rho)$ (Figure 3). Sequential updating keeps the run; batch updating takes Φ.

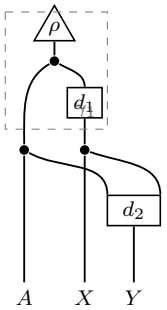


FIGURE 3. In $\mathscr{C}$. A run (γ_1,γ_2) of $K(X)\fatsemi K(Y)$ at ρ, with $\gamma_1 = \rho\,;\langle \mathrm{id}_A, d_1\rangle$ (dashed) and $\gamma_2 = \gamma_1\,;\langle \mathrm{id}_{A\otimes X}, d_2\rangle$ by Proposition 3.2; Φ sends it to the whole diagram γ_2.

Proposition 4.3 (Runs are conditional composites)**.** *Let ρ be a state on A, and for $1 \le k \le n$ let $d_k\colon A\otimes Y_1\otimes\cdots\otimes Y_{k-1} \to Y_k$ be channels. Put $\gamma_0 := \rho$ and $\gamma_k := \gamma_{k-1}\,;\langle \mathrm{id}, d_k\rangle$. Then $(\gamma_1,\dots,\gamma_n)$ is a run of $K(Y_1)\fatsemi\cdots\fatsemi K(Y_n)$ at ρ, every run is of this form with each d_k unique up to $\sim_{\gamma_{k-1}}$, and Φ sends it to $\gamma_n = \rho\,;\langle \mathrm{id}_A, d_1 \mathbin{⨟} \cdots \mathbin{⨟} d_n\rangle$, where $⨟$ is the dependent sum of channels of Section 2.1, the conditional composition of* [DLRS25, Def. 2.3]*, in any bracketing* [DLRS25, Prop. 2.5].

Proof. Deleting Y_k from $\gamma_{k-1}\,;\langle \mathrm{id}, d_k\rangle$ gives γ_{k-1} by naturality of del, so $\gamma_k \in L_{Y_k}(\gamma_{k-1})$, and γ_{k-1} is the point $r(\gamma_{k-2},\gamma_{k-1})$ (for $k\geq 2$) at which the k-th family is read; conversely Proposition 3.2 at γ_{k-1} gives d_k. The comparison Φ sends a run to its last lift (Table 3), and induction on k with (2.3) gives $\gamma_n = \rho\,;\langle \mathrm{id}_A, d_1 \fatsemi \cdots \fatsemi d_n\rangle$. □

With the marginal foot m instead, the k-th lift is read at the marginal on Y_{k-1} ($Y_0 := A$), so each d_k reads only Y_{k-1}, and a run is a chain of couplings that glues to a Markov chain.

5. Swaps and updating

Definition 5.1 (Swaps)**.** For objects X and Y, the *swap* is the invertible cell

$$s_{X,Y} := \Phi_{X,Y}\,;K(\sigma_{X,Y})\,;\Phi^{-1}_{Y,X}\colon K(X)\fatsemi K(Y)\Rightarrow K(Y)\fatsemi K(X).$$

Drawing a run as a tower whose k-th *floor* is the factor $K(Y_k)$ carrying the lift γ_k (Figure 4), $s_{X,Y}$ moves the floor X above the floor Y; it *elevates* X over Y.

Proposition 5.2 (Swaps are inversions)**.**

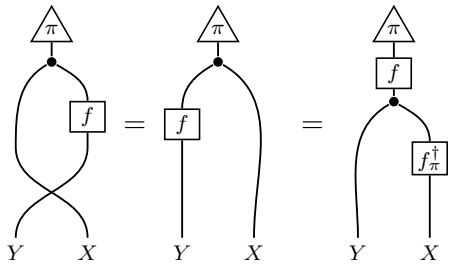
(i) At any state. *At ρ on A, $s_{X,Y}(\gamma,\gamma') = (\delta, \gamma'\,;(\mathrm{id}_A\otimes\sigma_{X,Y}))$, where δ is the marginal of $\gamma'\,;(\mathrm{id}_A\otimes\sigma_{X,Y})$ on $A\otimes Y$.*

(ii) At the unit. *At the state id_I, a pair is $(\pi,\pi\,;\langle\mathrm{id}_X, f\rangle)$ for a state π on X and a channel $f\colon X\to Y$, unique up to $\sim_\pi$, and*

$$s_{X,Y}(\pi,\ \pi\,;\langle\mathrm{id}_X,f\rangle) = \big(\pi\,;f,\ (\pi\,;f)\,;\langle\mathrm{id}_Y, f^\dagger_\pi\rangle\big).$$

The swap exchanges a prior and its lift along f for the predictive and its lift along the inversion $f^\dagger_\pi$.

Proof. (i) Φ sends (γ,γ') to γ', $K(\sigma)$ sends γ' to $\gamma'\,;(\mathrm{id}_A\otimes\sigma_{X,Y})$, and Φ^{-1} splits off the marginal. (ii) By Proposition 3.2 at $\rho=\mathrm{id}_I$, every pair is of the stated form. By (i), s sends it to the marginal $\pi\,;f$ and the state $\pi\,;\langle\mathrm{id}_X,f\rangle\,;\sigma_{X,Y}$, and



by the commutativity of copy and by (2.1). □

Lemma 5.3 (Coherence of elevation)**.** *For objects $X, Y_1,\dots,Y_n$, let $S_k := \mathrm{id}\fatsemi s_{X,Y_k}\fatsemi\mathrm{id}$ elevate X over the k-th floor in $K(Y_1)\fatsemi\cdots\fatsemi K(Y_{k-1})\fatsemi K(X)\fatsemi K(Y_k)\fatsemi\cdots\fatsemi K(Y_n)$. Then*

$$S_1\,;S_2\,;\cdots;S_n\,;\Phi = \Phi\,;K(\sigma_{X,\,Y_1\otimes\cdots\otimes Y_n})$$

as cells $K(X)\fatsemi K(Y_1)\fatsemi\cdots\fatsemi K(Y_n)\Rightarrow K(Y_1\otimes\cdots\otimes Y_n\otimes X)$, where Φ is the $(n+1)$-fold comparison.

Proof. By naturality of Φ (Theorem 4.2), $S_k\,;\Phi = \Phi\,;K(\mathrm{id}_{Y_{<k}}\otimes\sigma_{X,Y_k}\otimes\mathrm{id}_{Y_{>k}})$ with $Y_{<k} := Y_1\otimes\cdots\otimes Y_{k-1}$ and $Y_{>k} := Y_{k+1}\otimes\cdots\otimes Y_n$, since s_{X,Y_k} is $\Phi\,;K(\sigma)\,;\Phi^{-1}$ and Φ is coherent. Composing over k and using functoriality of K gives $\Phi\,;K$ applied to $(\sigma_{X,Y_1}\otimes\mathrm{id})\,;\cdots\,;(\mathrm{id}\otimes\sigma_{X,Y_n})$, which is $\sigma_{X,\,Y_1\otimes\cdots\otimes Y_n}$ by the hexagon axiom of the symmetry. □

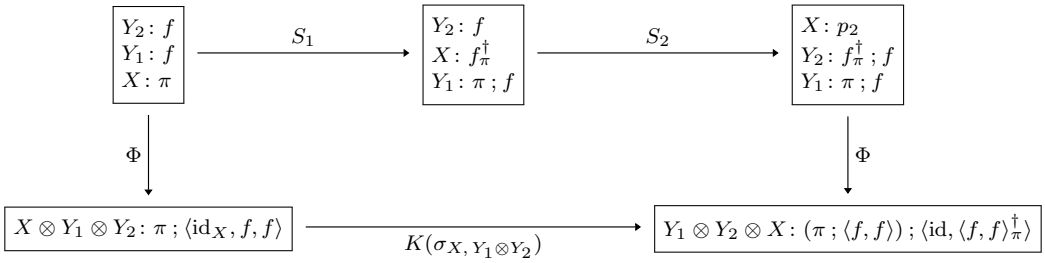


FIGURE 4. In $\mathbf{C}_{\mathscr{B}}$ at the unit state id_I, for $n = 2$ and $f_1 = f_2 = f$. A run is drawn as a tower of floors, each floor the channel that gives its object from the floors below (Proposition 3.2); each S_k is an inversion (Proposition 5.2), p_2 is the second sequential update of Definition 5.4, and the square is Lemma 5.3.

Updating reads Lemma 5.3 at the unit state (Figure 4).

Kamiya and Welliaveetil define sequential and batch updates at data [KW21, Secs. 4.1–4.2]; the following definition keeps the data as inputs.

Definition 5.4 (Sequential and batch updates)**.** Let π be a state on X and $f_k \colon X \to Y_k$ channels for $1 \le k \le n$, with $Y_{\le k} := Y_1 \otimes \cdots \otimes Y_k$ and $F_k := \langle f_1, \dots, f_k \rangle \colon X \to Y_{\le k}$, $F_0 := \mathrm{del}_X$.

(i) *Batch update.* An inversion $F_n^\dagger$ of F_n at π.

(ii) *Sequential updates.* $\zeta_0 := \pi$ and $p_0 := \pi$, and for $k \ge 1$ a conditional $p_k \colon Y_{\le k} \to X$ of X given $Y_{\le k}$ for the state $\zeta_{k-1} ; (\mathrm{id}_{Y_{<k}} \otimes \langle f_k, \mathrm{id}_X \rangle)$ on $Y_{\le k} \otimes X$, with $\zeta_k := (\pi ; F_k) ; \langle \mathrm{id}_{Y_{\le k}}, p_k \rangle$.

The k-th step conditions the joint state ζ_{k-1} of the data so far and X, so the posterior joint state is the next state, as for the joint foot.

Theorem 5.5 (Sequential and batch updating)**.** *Let π, f_k, F_k, p_k and ζ_k be given by Definition 5.4.*

(i) Agreement. *$\zeta_k = \pi ; \langle F_k, \mathrm{id}_X \rangle$ for $0 \le k \le n$, and p_k is an inversion of F_k at π, so $p_k \sim_{\pi ; F_k} F_k^\dagger$ for $1 \le k \le n$: the sequential update is a batch update.*

(ii) From coherence. *The lifts $\gamma_k := \pi ; \langle \mathrm{id}_X, F_k \rangle$, $0 \le k \le n$, form a run $(\gamma_0, \dots, \gamma_n)$ of $K(X) \fatsemi K(Y_1) \fatsemi \cdots \fatsemi K(Y_n)$ at id_I. The swaps $S_1 ; \cdots ; S_k$ put ζ_k on the floor X, and Lemma 5.3 puts $\pi ; \langle F_k, \mathrm{id}_X \rangle$ there, which proves (i) again without the induction: the swaps produce the sequential updates floor by floor, and $K(\sigma)$ the batch update at once.*

(iii) Data. *Let y_k be deterministic states on Y_k, $z_k := y_1 \otimes \cdots \otimes y_k$ and $z_0 := \mathrm{id}_I$, and suppose that the data determine the classes: $z_k ; g = z_k ; g'$ whenever $g \sim_{\pi ; F_k} g'$, for $1 \le k \le n$. Then for $1 \le k \le n$, $(z_{k-1} \otimes \mathrm{id}_{Y_k}) ; p_k$ is an inversion of f_k at $\pi_{k-1} := z_{k-1} ; p_{k-1}$, so the posteriors $\pi_k = z_k ; p_k$ are obtained one datum at a time, and the last one, π_n, equals $z_n ; g$ for every batch update g.*

Proof. (i) By induction on k. At $k = 0$, $\zeta_0 = \pi$. If $\zeta_{k-1} = \pi ; \langle F_{k-1}, \mathrm{id}_X \rangle$, then $\zeta_{k-1} ; (\mathrm{id} \otimes \langle f_k, \mathrm{id}_X \rangle) = \pi ; \langle F_k, \mathrm{id}_X \rangle$ by coassociativity of copy, and its conditional p_k gives $\pi ; \langle F_k, \mathrm{id}_X \rangle = (\pi ; F_k) ; \langle \mathrm{id}, p_k \rangle = \zeta_k$. This equation is (2.1) for p_k as an inversion of F_k at π, and inversions agree $\sim_{\pi ; F_k}$.

(ii) γ_k has marginal γ_{k-1}, so $(\gamma_0, \dots, \gamma_n)$ with $\gamma_0 = \pi$ is a run. By Proposition 5.2(i), S_1 replaces the first two lifts by $\pi ; f_1$ and $\pi ; \langle f_1, \mathrm{id}_X \rangle$, which is ζ_1, and by induction $S_1 ; \cdots ; S_k$ leaves the run $(\pi ; F_1, \dots, \pi ; F_k, \zeta_k, \dots)$: each S_k applies σ_{X, Y_k} to the lift over ζ_{k-1}, whose

conditional of X is p_k. By Lemma 5.3 for the first $k+1$ floors, whiskered by the identity on the rest, the same lift is $\gamma_k \mathbin{;} \sigma_{X,Y_{\le k}} = \pi \mathbin{;} \langle F_k, \mathrm{id}_X \rangle$, which by Proposition 5.2(ii) for F_k is $(\pi \mathbin{;} F_k) \mathbin{;} \langle \mathrm{id}, F_k^\dagger \rangle$. Hence the two readings of one lift give $\zeta_k = \pi \mathbin{;} \langle F_k, \mathrm{id}_X \rangle$ and $p_k \sim_{\pi;F_k} F_k^\dagger$ by Proposition 3.2, independently of the induction in (i).

(iii) For $k = 1$ the claim is the defining equation of p_1, since $z_0 = \mathrm{id}_I$ and $p_0 = \pi$. For $k \ge 2$, by (i) both sides of the defining equation of p_k are states $(\pi \mathbin{;} F_{k-1}) \mathbin{;} \langle \mathrm{id}, a \rangle$ and $(\pi \mathbin{;} F_{k-1}) \mathbin{;} \langle \mathrm{id}, b \rangle$ on $Y_{<k} \otimes Y_k \otimes X$, where $a := p_{k-1} \mathbin{;} \langle f_k, \mathrm{id}_X \rangle$ and $b \colon Y_{<k} \to Y_k \otimes X$ draws y from $p_{k-1} \mathbin{;} f_k$ and then x from p_k at y. Hence $a \sim_{\pi;F_{k-1}} b$ by Proposition 3.2, and the hypothesis at z_{k-1} gives $z_{k-1} \mathbin{;} a = z_{k-1} \mathbin{;} b$, which is (2.1) for $(z_{k-1} \otimes \mathrm{id}_{Y_k}) \mathbin{;} p_k$ as an inversion of f_k at π_{k-1}. Finally $p_n \sim_{\pi;F_n} g$ by (i), and the hypothesis at z_n gives $z_n \mathbin{;} g = z_n \mathbin{;} p_n = \pi_n$. The hypothesis cannot be dropped: at a datum of predicted probability zero, a conditional may take any value. □

Example 5.6 (The coin)**.** In FinStoch, the Markov category of finite sets and stochastic matrices, let $X = \{\mathrm{fair}, \mathrm{biased}\}$ with $\pi = (\frac{1}{2}, \frac{1}{2})$, and let $f \colon X \to Y = \{\mathrm{head}, \mathrm{tail}\}$ give head with probability $\frac{1}{2}$ and $\frac{3}{4}$. The first flip is predicted by $\pi \mathbin{;} f = (\frac{5}{8}, \frac{3}{8})$, and after head the posterior is $\pi_1 := \delta_{\mathrm{head}} \mathbin{;} f_\pi^\dagger = (\frac{2}{5}, \frac{3}{5})$. The second flip is predicted by $\pi_1 \mathbin{;} f$, which gives head the probability $\frac{13}{20}$, and after a second head the posterior is $\pi_2 = (\frac{4}{13}, \frac{9}{13})$. The batch update gives $(\delta_{\mathrm{head}} \otimes \delta_{\mathrm{head}}) \mathbin{;} \langle f, f \rangle_\pi^\dagger$, proportional to $(\frac{1}{2} \cdot \frac{1}{4}, \frac{1}{2} \cdot \frac{9}{16})$, which is $(\frac{4}{13}, \frac{9}{13})$ as well. Plugging in is determined since $\frac{5}{8}$ and $\frac{13}{20}$ are positive, so Theorem 5.5(iii) applies. Each head moves the prior towards the biased coin.

For $n = 2$ and $f_1 = f_2 = f$, Theorem 5.5(i) is $p_2 \sim_{\pi;\langle f,f \rangle} \langle f, f \rangle_\pi^\dagger$, the inversion of f at the first posterior agreeing with the batch inversion, known from Lemma 11.11 of Fritz [Fri20], and measure-theoretically for supervised models from Theorem 3.1 of Lê [Lê25]; (ii) obtains it from the coherence of K. The observations may also depend on earlier data, $f_k \colon Y_{<k} \otimes X \to Y_k$, with the same proof.

Remark 5.7 (Deferred evidence)**.** In Theorem 5.5(i) the data are open inputs of p_k and $F_n^\dagger$. In FinStoch a state ω on $Y_{\le n}$ that weighs only data of positive predicted probability under $\pi \mathbin{;} F_n$ thus gives $\omega \mathbin{;} p_n = \omega \mathbin{;} F_n^\dagger$. Plugging ω_1 in before the second step is another update, since an inversion is not affine in its prior [Jac19, Prop. 5.3(2)]: with $\pi = (\frac{1}{2}, \frac{1}{2})$, $f(0) = (\frac{4}{5}, \frac{1}{5})$, $f(1) = (\frac{1}{2}, \frac{1}{2})$, $\omega_1 = (\frac{7}{10}, \frac{3}{10})$ and $\omega_2 = (\frac{3}{5}, \frac{2}{5})$, the first point of X gets 0.49809 from the batch update and 0.49827 from early plugging.

Remark 5.8 (The Bayes filter)**.** A hidden Markov model with initial state π on X_1, transitions $h_k \colon X_{k-1} \to X_k$ and emissions $e_k \colon X_k \to Y_k$ [FKMSW24, Prop. 3.7(iv), Def. 3.8] is, by Proposition 4.3, a run of $K(X_1) \fatsemi K(Y_1) \fatsemi \cdots \fatsemi K(X_n) \fatsemi K(Y_n)$ at id_I in which the floor X_k reads X_{k-1} and the floor Y_k reads X_k; this needs the joint foot r, since with m a floor reads only the floor below it. Deleting the floor X_{k-1} by $K(\mathrm{del}_{X_{k-1}})$ followed by Φ_0^{-1} predicts, and the swap s_{X_k,Y_k} updates, and naturality of Φ, as in the proof of Lemma 5.3, recovers the recursion of the Bayes filter [FKMSW24, Prop. 4.2]. The Markov property enters only through the factorisation of the model, and with identity transitions the recursion is Definition 5.4(ii), and its agreement with the batch update is Theorem 5.5(i).

6. The reading in probability spaces

Recall the dependent sum of channels (2.3) and the rules (a)–(c) of Section 2.1.

Proposition 6.1 (The contextad of disintegrations)**.** *A* grade *at* (B,ω) *in* $\mathrm{PS}(\mathscr{C})$ *is a morphism* $[c]_\omega\colon (B,\omega)\to(Y,\omega\,;c)$*, and its base change along* $[h]_{\omega'}\colon(B',\omega')\to(B,\omega)$ *is the composite* $[h\,;c]_{\omega'}$*. By Proposition 3.2,* $[c]_\omega\mapsto\omega\,;\langle\mathrm{id}_B,c\rangle$ *identifies the grades at* (B,ω) *with values in* Y *with the lifts* $L_Y(\omega)$*.*

(i) Transport. *For* $\mu:=\omega'\,;\langle\mathrm{id}_{B'},h\,;c\rangle$ *there is a unique morphism* $[E]_\mu\colon(B'\otimes Y,\mu)\to(B\otimes Y,\omega\,;\langle\mathrm{id}_B,c\rangle)$ *with* $\langle\mathrm{id}_{B'},h\,;c\rangle\,;\langle\mathrm{del}_{B'}\otimes\mathrm{id}_Y,E\rangle\sim_{\omega'}h\,;\langle\mathrm{id}_B,c\rangle\,;\langle\mathrm{del}_B\otimes\mathrm{id}_Y,\mathrm{id}_{B\otimes Y}\rangle$.

(ii) Contextad. *The grades form a split discrete fibration over* $\mathrm{PS}(\mathscr{C})$*, and with the action* $(B,\omega)\odot[c]_\omega:=(B\otimes Y,\omega\,;\langle\mathrm{id}_B,c\rangle)$*, (i) on morphisms, the unit grades* $[\mathrm{del}_B]_\omega$ *and the dependent sum* $[c]_\omega\fatsemi[d]_{\omega;\langle\mathrm{id}_B,c\rangle}:=[c\fatsemi d]_\omega$*, they form a strict contextad on* $\mathrm{PS}(\mathscr{C})$ [CM24].

Proof. (i) The right-hand side is a state on $B'\otimes Y\otimes B\otimes Y$ whose marginal on $B'\otimes Y$ is μ, so Proposition 3.2 for that marginal gives a unique class $[E]_\mu$, whose codomain is $\omega'\,;h\,;\langle\mathrm{id}_B,c\rangle=\omega\,;\langle\mathrm{id}_B,c\rangle$; it depends only on $[h]_{\omega'}$ and $[c]_\omega$ by (a)–(c). (ii) Since composition in $\mathrm{PS}(\mathscr{C})$ is well defined and associative, base change is a functor from $\mathrm{PS}(\mathscr{C})^{\mathrm{op}}$ to large sets, and the grades are its category of elements. The transport of (i) is functorial because the transport along $h'\,;h$ satisfies the defining relation of the composite by (a) and (b). By (2.3), $[c\fatsemi d]_\omega$ depends only on the two classes, $(B,\omega)\odot([c]\fatsemi[d])=((B,\omega)\odot[c])\odot[d]$, and $(h\,;c)\fatsemi(E\,;d)\sim_{\omega'}h\,;(c\fatsemi d)$ by (a) and (i), so the dependent sum preserves the chosen lifts on the nose. Finally $\langle\mathrm{id}_B,\mathrm{del}_B\rangle=\mathrm{id}_B$, $\mathrm{del}_B\fatsemi c=c=c\fatsemi\mathrm{del}_{B\otimes Y}$, and both bracketings of $c\fatsemi d\fatsemi e$ copy B and apply c, d and e once. With discrete fibres the structure cells are identities and the axioms hold. □

Base change in this contextad is precomposition, not pullback, since the marginal projections of $\mathrm{PS}(\mathscr{C})$ need not have pullbacks. Inversion is a grade of this contextad, and reversing a dependent sum inverts it.

Proposition 6.2 (Reversed sums, [CJ19, Fri20])**.** *Let* π *be a state on* X*,* $f\colon X\to Y$ *and* $g\colon Y\to Z$*, with* $\rho:=\pi\,;f$ *and* $\tau:=\rho\,;g$*.*

(i) Inversion. *For* $F\colon X\to W$*, the inversions of* F *at* π *form the grade* $[F]^\dagger_\pi$ *at* $(W,\pi\,;F)$ *that Proposition 3.2 sends to* $\pi\,;\langle F,\mathrm{id}_X\rangle$*. It depends only on* $[F]_\pi$*,* $[\mathrm{id}_X]^\dagger_\pi=[\mathrm{id}_X]_\pi$*, and* $[F^\dagger_\pi]^\dagger_{\pi;F}=[F]_\pi$*; with (iii) the classes form an identity-on-objects functor* $\dagger\colon\mathrm{PS}(\mathscr{C})^{\mathrm{op}}\to\mathrm{PS}(\mathscr{C})$*.*

(ii) Reversal. *The dependent sums* $\pi\fatsemi f\fatsemi g$ *and* $\tau\fatsemi g^\dagger_\rho\fatsemi f^\dagger_\pi$*, read with* g *and* $f^\dagger_\pi$ *on the copy of* Y*, are* $\pi\,;\langle\mathrm{id}_X,f\,;\langle\mathrm{id}_Y,g\rangle\rangle$ *and* $\tau\,;\langle\mathrm{id}_Z,g^\dagger_\rho\,;\langle\mathrm{id}_Y,f^\dagger_\pi\rangle\rangle$*, and they agree up to the symmetry* $X\otimes Y\otimes Z\cong Z\otimes Y\otimes X$*.*

(iii) Functoriality. $[f\,;g]^\dagger_\pi=[g]^\dagger_\rho\,;[f]^\dagger_\pi$.

This is the dagger of Cho and Jacobs and of Fritz [CJ19, Cor. 5.6], [Fri20, Rem. 13.10]: (ii) reads the conditional product of couplings [Fri20, (12.3), Rem. 12.10] in both directions, and deleting Y gives the chain rule (iii) [Smi20, Thm. 5.2], [Smi22, Thm. 4.3.14].

Remark 6.3 (Dependent Bayesian lenses)**.** Dependent Bayesian lenses let the backward object depend on the state and allow any backward morphism [BH22, Sec. 4], [BHS23, Def. 15], and so they do not arise from contextful arrows on $\mathscr{B}$, which move the state instead. Their canonical inversion does. By Proposition 3.2, a morphism $[h]_\nu\colon(Y,\nu)\to(X,\pi)$ of $\mathrm{PS}(\mathscr{C})$ is the lift $\nu\,;\langle\mathrm{id}_Y,h\rangle\in L_X(\nu)$ whose marginal foot is π, and $f\mapsto$

$(f, ([f]^\dagger_\pi)_\pi)$ is a functor, by Proposition 6.2(i) and (iii), into the category whose objects are those of $\mathscr{C}$ and whose morphisms $X \to Y$ are pairs $(f, (f^\sharp_\pi)_\pi)$ with $f^\sharp_\pi \colon (Y, \pi\,;\, f) \to (X, \pi)$ in $\mathrm{PS}(\mathscr{C})$, composed by $(g^\sharp_{\pi;f}\,;\, f^\sharp_\pi)_\pi$; under the bijection above its backward parts are the second components of the swaps of Proposition 5.2(ii). This is the category proposed in [BHS23, Rem. 23]; with supports, *Supp* (Proposition 2.5) sends the functor to the canonical inversions of [BH22, Prop. 15] and [BHS23, Prop. 21].

Proposition 6.4 (Generative and sequential sums)**.** *Let π, f_k, F_k, p_k and ζ_k be given by Definition 5.4, and let $q_k := p_{k-1}\,;\, f_k \colon Y_{<k} \to Y_k$ be the predictives.*

(i) Generative. *The batch observation F_k is the iterated dependent sum of $f_1, \dots, f_k$ over (X, π), each reading X alone.*

(ii) Sequential. *$\pi\,;\, F_k = q_1 \mathbin{⨟} \cdots \mathbin{⨟} q_k$, and $\zeta_k = (q_1 \mathbin{⨟} \cdots \mathbin{⨟} q_k) \mathbin{⨟} p_k$. Hence $\pi \mathbin{⨟} F_k$ and $(q_1 \mathbin{⨟} \cdots \mathbin{⨟} q_k) \mathbin{⨟} p_k$ agree up to the symmetry of $X \otimes Y_{\le k}$, and $[p_k]_{\pi;F_k} = [F_k]^\dagger_\pi$ in $\mathrm{PS}(\mathscr{C})$.*

Proof. (i) By induction, $F_{k-1} \mathbin{⨟} f_k = \langle F_{k-1}, f_k\rangle = F_k$ when f_k reads only X. (ii) By Theorem 5.5(i), $\zeta_k = \pi\,;\, \langle F_k, \mathrm{id}_X\rangle = (\pi\,;\, F_k)\,;\, \langle \mathrm{id}, p_k\rangle$, and deleting X from $\zeta_{k-1}\,;\, (\mathrm{id} \otimes \langle f_k, \mathrm{id}_X\rangle)$ gives $\pi\,;\, F_k = (\pi\,;\, F_{k-1})\,;\, \langle \mathrm{id}, p_{k-1}\,;\, f_k\rangle$, which is $(\pi\,;\, F_{k-1}) \mathbin{⨟} q_k$. Induction and (2.3) give the sums, and Proposition 6.2(i) the class. □

Outlook

By Proposition 6.4, the disintegrations of Proposition 6.1 carry the run of Theorem 5.5 in $\mathrm{PS}(\mathscr{C})$, where each posterior is the next prior. The contextad of these disintegrations has discrete fibres; a contextad whose fibres also carry maps of data, as the foot cells of K carry channels, is left open.

Acknowledgements

Part of this work was presented at Categories in Tokyo on 3 November 2026.

References

[BH22] D. Braithwaite, J. Hedges. Dependent Bayesian lenses: categories of bidirectional Markov kernels with canonical Bayesian inversion. Preprint, arXiv:2209.14728, 2022.

[BHS23] D. Braithwaite, J. Hedges, T. St Clere Smithe. The compositional structure of Bayesian inference. In *MFCS 2023*, LIPIcs 272, 24:1–24:15, 2023.

[CDDG17] F. Clerc, V. Danos, F. Dahlqvist, I. Garnier. Pointless learning. In *FoSSaCS 2017*, Lecture Notes in Comput. Sci. 10203, Springer, 2017.

[CJ19] K. Cho, B. Jacobs. Disintegration and Bayesian inversion via string diagrams. *Math. Struct. Comput. Sci.* 29(7), 938–971, 2019.

[CM24] M. Capucci, D. J. Myers. Contextads as wreaths; Kleisli, Para, and Span constructions as wreath products. Preprint, arXiv:2410.21889v1, 2024. Numbers refer to v1.

[DLRS25] E. Di Lavore, M. Román, P. Sobociński. Partial Markov categories. Preprint, arXiv:2502.03477, 2025. Numbers refer to v5.

[FKMSW24] T. Fritz, A. Klingler, D. McNeely, A. Shah-Mohammed, Y. Wang. Hidden Markov models and the Bayes filter in categorical probability. Preprint, arXiv:2401.14669, 2024. Numbers refer to v3.

[Fré51] M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon, Sect. A* (3) 14, 53–77, 1951.

[Fri20] T. Fritz. A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics. *Adv. Math.* 370, 107239, 2020.

[Jac19] B. Jacobs. The mathematics of changing one's mind, via Jeffrey's or via Pearl's update rule. *J. Artif. Intell. Res.* 65, 783–806, 2019.

[JS91] A. Joyal, R. Street. The geometry of tensor calculus, I. *Adv. Math.* 88(1), 55–112, 1991.

[KW21] K. Kamiya, J. Welliaveetil. A category theory framework for Bayesian learning. Preprint, arXiv:2111.14293, 2021.

[Lê25] H. V. Lê. Batch learning equals online learning in Bayesian supervised learning. Preprint, arXiv:2510.16892, 2025.

[Sel11] P. Selinger. A survey of graphical languages for monoidal categories. In *New Structures for Physics*, Lecture Notes in Phys. 813, 289–355, Springer, 2011.

[Smi20] T. St Clere Smithe. Bayesian updates compose optically. Preprint, arXiv:2006.01631, 2020. Numbers refer to v2.

[Smi22] T. St Clere Smithe. Mathematical foundations for a compositional account of the Bayesian brain. PhD thesis, University of Oxford, arXiv:2212.12538, 2022.