Categories of Brègman operations and epistemic (co)monads

Ryszard Paweł Kostecki

*National Quantum Information Center & Institute of Informatics*
*Faculty of Mathematics, Physics, and Informatics, University of Gdańsk*
*Andersa 27, 81-824 Sopot, Poland*
*International Center for Theory of Quantum Technologies, University of Gdańsk*
*Wita Stwosza 63, 80-308 Gdańsk, Poland*
*kostecki@fuw.edu.pl*

March 13, 2021

Abstract

We construct a categorical framework for nonlinear postquantum inference, with embeddings of convex closed sets of suitable reflexive Banach spaces as objects and pullbacks of Brègman quasi-nonexpansive mappings (in particular, constrained maximisations of Brègman relative entropies) as morphisms. It provides a nonlinear convex analytic analogue of Chencov’s programme of geometric study of categories of linear positive maps between spaces of states, a working model of Mielnik’s nonlinear transmitters, and a setting for nonlinear resource theories (with monoids of Brègman quasi-nonexpansive maps as free operations, their asymptotic fixed point sets as free sets, and Brègman relative entropies as resource monotones). We construct a range of concrete examples for semi-finite JBW-algebras and any $W^*$-algebras. Due to relative entropy’s asymmetry, all constructions have left and right versions, with Legendre duality inducing categorical equivalence between their well-defined restrictions. Inner groupoids of these categories implement the notion of statistical equivalence. The hom-sets of a subcategory of morphisms given by entropic projections have the structure of partially ordered commutative monoids (so, they are resource theories in Fritz’s sense). Further restriction of objects to affine sets turns Brègman relative entropy into a functor. Finally, following Lawvere’s adjointness paradigm for deductive logic, but with a semantic twist representing Jaynes’ and Chencov’s views on statistical inference, we introduce a category-theoretic multi-(co)agent setting for inductive inference theories, implemented by families of monads and comonads. We show that the brègmanian approach provides some special cases of this setting.

1 Introduction

This paper stems from the questions: 1) How to construct a theory of nonlinear (post)quantum operations, valid also in continuously infinite dimensions, and exhibiting useful information-, resource-, and category-theoretic structures? 2) How to formalise the logics of inductive (predictive, statistical) inference theory, in (some) analogy to category-theretic formalisations of the logics of deductive inference?

Extension of the maximum entropy approach from model construction (objects) [73, 160, 100, 101, 102] to inductive inferences (morphisms) [155, 115, 35, 50, 90] allows to derive the Bayes–Laplace and Jeffrey’s rules [177, 174, 46, 67], Lüders’ and quantum Jeffrey’s rules [89], partial trace [141], and (preduals of) conditional expectations [114] as special cases of constrained maximisation of the Kullback–Leibler/Umegaki relative entropy (left or right $D_1$-projections). However, to establish a full-fledged relative entropic (post)bayesian setting, two related problems have to be solved: 1) choice of a sufficiently rich and well-behaved class of relative entropies; 2) identification of a family of morphisms, which (roughly) could be for entropic projections what CPTP maps are for (preduals of) conditional expectations. Brègman [35] introduced a class $-D_q$ of relative entropies satisfying a generalisation of a pythagorean equation with respect to left $D_q$-projections (independently, Chencov [50] discovered its right version for $D_1$). The works [3, 6, 44, 45, 18, 19, 150] (among others) established a successful Banach space generalisation of Brègman’s theory. Reich [147] introduced a class of left strongly $D_q$-quasi-nonexpansive maps, $LSQ(\Psi)$, which is closed under composition,
and (under some additional conditions [19, 148]) contains left $D_\Psi$-projections. Right versions of these results also hold [131, 132]. Right and left $D_\Psi$-projections are generalisations of Hilbert space metric projections, better behaved and weaker than Banach space metric projections [3, 4]. They are characterised by the corresponding pythagorean equations, while the criteria of convergence defining $\text{LSQ}(\Psi)$ and $\text{RSQ}(\Psi)$ are, in essence, topological versions of these equations. However, good behaviour of these maps requires some additional geometric properties of the underlying Banach space $X$. These properties do not hold neither for generic base norm spaces nor for preduals of JBW- and $W^*$-algebras. The missing link, provided by us in [113], is twofold: 1) introducing Brègman $\ell$-information $D_\ell,\Psi$ by a bijective pullback ($\ell$-embedding) of $D_\Psi$ from geometrically well behaved (e.g., reflexive) space $X$, constructed over not so well behaved (e.g., base norm) spaces (and doing the same with $\text{SOQ}(\Psi)$ maps); 2) providing rich family of models (i.e., triples $(X, \ell, \Psi)$), by: 2a) establishing characterisation of legedreness, and a sufficient condition for $\text{LSQ}$-adaptedness and $\text{RSQ}$-composability, of a family $\Psi_\varphi := \int_0^1 \varphi(t)\,dt$, where $\varphi$ is a gauge function of a nonlinear duality map $\mathcal{D}^G\Psi_\varphi = J_\varphi : X \to X^*$ [25, 11]; 2b) constructing a range of concrete models in nonassociative and noncommutative settings. The resulting families $\text{LSQ}(\ell, \Psi)$ and $\text{RSQ}(\ell, \Psi)$ provide nonlinear convex analytic analogues of linear CPTP maps, based on the geometry of generalised pythagorean equation, as opposed to tensor products. While $\ell$-embeddings of brègmanian structures solve a mathematical problem, they introduce a conceptual one: the basic objects (information state spaces) of a theory are $\ell$-closed $\ell$-convex sets, which, as opposed to compact convex sets, do not rely on base norm/order unit semantics (allowing for “information theory without probability” (c.f. [97]) on objects and hom-sets of categories which do not admit any (generalised) probabilistic structure). This relativisation of the type of optimal/ideal experimental data with respect to the choice of the system of inductive inference (c.f. [66]) requires to be coherently addressed. For this purpose, in §8 we propose a categorical approach to adjointness between theoretical model construction and predictive verification, modeled after Lawvere’s approach to categorical deductive logic [119], yet with a twist, taking into account Chencov’s and Jaynes’ approaches to mathematical and conceptual foundations of statistical inference. We show that, under some conditions, it forms a resource theory of intersubjective knowledge (with agency of free operations and coagency of selection of referential experimental designs).

2 Brègman projections and quasi-nonexpansive maps

In terminological (resp., mathematical) agreement with [175, 149, 170, 33] (resp., [71, 58]), we define: an information on a set $Z$ as a map $D : Z \times Z \to [0, \infty]$ such that $D(x,y) = 0 \iff x = y$; a relative entropy as $-D$. Given a function $f : Y \to [-\infty, \infty]$ on a real Banach space $Y$ with $\text{efd}(f) := \{x \in Y \mid f(x) \neq \infty\} \neq \emptyset$, $f^F$ will denote a Fenchel dual of $f$ with respect to a bilinear duality map $[(x,y) := y(x) \in \mathbb{R} \forall (x,y) \in Y \times Y^*, \text{where } Y^* \text{ denotes the Banach dual}}$.

In §2 we recall basic notions of convex nonlinear analysis and brègmanian theory in reflexive Banach spaces, discussed in details, with further references, in [31, 154, 113]. As for §3, the notions of Brègman $\ell$-information and $\ell$-projection were introduced (abstractly) in [112] (with $\ell$-embeddings generalising earlier ideas of [134, 105] and [142, 111, 179, 81, 80, 145, 104]), and are studied (concretely, with the corresponding $D_\Psi$-quasi-nonexpansive $\ell$-operations, including the examples of §4, as well as an extension to nonreflexive case, left and right $D_\Psi$-Chebyshev sets, continuity of $D_\Psi$-projections, limitations of Legendre duality, etc) in [113]. The rest of this paper is new, and can be seen as a category-theoretic counterpart to [113].

Our focus on categories of inductive inference morphisms, instead of axiomatisation of probability spaces, follows the insights: <<Many physicists take it for granted that their theories can be either refuted or verified by comparison with experimental data. In order to evaluate such data, however, one must employ statistical estimation and inference methods which, unfortunately, always involve an ad hoc proposition. (...) no verification is possible unless the relevant inference method is an integral part of the theory>> [162] (c.f. also [82]), <<the motion creates the form>> [137] (c.f. also [59]), and <<the main goal of statistician is to choose a priori reasonable families guaranteeing good rates of convergence of loss functions>> [52]. In addition, we see the passage from ‘linearity’ (in the sense of [135, 60, 83, 136]) to ‘nonlinearity’ (in our sense) along the lines of: <<the great watershed in optimization isn’t between linearity and nonlinearity, but convexity and nonconvexity>> [152] (c.f. also [122]).
of $Y$. In what follows, $X$ denotes a reflexive real Banach space, int denotes an interior in norm topology of $X$, and $\Psi : X \to [-\infty, \infty]$ is Legendre [151, 18] (so, its Gateaux derivative is a bijection, $D^G\Psi : \text{int}(\text{efd}(\Psi)) \to \text{int}(\text{efd}(\Psi^F))$, with $(D^G\Psi)^{-1} = D^G\Psi^F$. A map $D_\Psi : X \times X \to [0, \infty]$, $D_\Psi(z, w) := \Psi(z) - \Psi(w) - \langle z - w, D^G\Psi(w) \rangle = \Psi(z) + \Psi^F(D^G\Psi(w)) - \langle z, D^G\Psi(w) \rangle$ for $w \in \text{int}(\text{efd}(\Psi))$ and $\infty$ otherwise [42, 35, 109, 44], is an information [18], called Bregman information. For $y \in \text{int}(\text{efd}(\Psi))$, $C \subseteq X$, and $\partial \neq C \cap \text{int}(\text{efd}(\Psi))$, if the set $\arg\inf_{x \in C} \{D_\Psi(x, y)\}$ (resp., $\arg\inf_{x \in C} \{D_\Psi(y, x)\}$) is a singleton, then its element will be denoted $\overline{\Psi}^D_{C^*}(y)$ (resp., $\overline{\Psi}^D_{C}(y)$), and called left (resp., right) $D_\Psi$-projection of $y$ onto $C$. Both left and right $D_\Psi$-projections are idempotent. If $\partial \neq C \subseteq \text{int}(\text{efd}(\Psi))$ is convex and closed, then $\forall y \in \text{int}(\text{efd}(\Psi)) \exists \overline{\Psi}^D_{C}(y)$ [18]. Furthermore, $\overline{\Psi}^D_{K^*} = D^G\Psi \circ \overline{\Psi}^D_{D^G\Psi(K)} \circ D^G\Psi$, and $\overline{\Psi}^D_{D^G\Psi(M)} = D^G\Psi \circ \overline{\Psi}^D_{M} \circ D^G\Psi^F$ for nonempty, closed, convex sets $D^G\Psi(K)$ and $M$ [23, 22, 131, 129]. If $K$ is a closed affine subspace of $X$, then the left pythagorean equation, $D_\Psi(x, y) = D_\Psi(x, \overline{\Psi}^D_{K^*}(y)) + D_\Psi(\overline{\Psi}^D_{K}(y), y) \forall (x, y) \in K \times \text{int}(\text{efd}(\Psi))$, holds [35, 6, 5]. If $D^G\Psi(K)$ is a closed affine subspace of $X$, then the right pythagorean equation, $D_\Psi(x, y) = D_\Psi(x, \overline{\Psi}^D_{K^*}(x)) + D_\Psi(\overline{\Psi}^D_{K}(x), y) \forall (x, y) \in \text{int}(\text{efd}(\Psi)) \times K$, holds [50, 131].

Given $\partial \neq M \subseteq \text{int}(\text{efd}(\Psi))$ and a function $T : M \to \text{int}(\text{efd}(\Psi))$, $Fix(T) := \{x \in M \mid T(x) = x\} \neq \partial$ is called a set of fixed points, while $Fix(T)$, called a set of asymptotic fixed points consists of such $x \in M$ that there exists a sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq M$ weakly convergent to $x$ with $\lim_{n \to \infty} \|x_n - Tx_n\| = 0$. In general, $Fix(T) \subseteq Fix(T)$. $T : M \to \text{int}(\text{efd}(\Psi))$ is called: completely $D_\Psi$-nonexpansive (or CN($\Psi$)) [45] iff $D_\Psi(T(x), T(y)) \leq D_\Psi(x, y) \forall x, y \in M$; left strongly $D_\Psi$-quasi-nonexpansive (or LSQ($\Psi$)) [47, 147, 132] iff $D_\Psi(x, T(y)) \leq D_\Psi(x, y) \forall (x, y) \in X \times \text{int}(\text{efd}(\Psi))$; right strongly $D_\Psi$-quasi-nonexpansive (or RSQ($\Psi$)) [131] iff $D_\Psi(T(x), y) \leq D_\Psi(x, y) \forall (x, y) \in \text{int}(\text{efd}(\Psi)) \times M$ and $(p \in F(T))$. $\Psi$ is bounded, supercoercive, (uniformly Fréchet differentiable and totally convex) on bounded subsets of $X$, $\partial \neq K \subseteq \text{int}(\text{efd}(\Psi))$, $\{T_1, \ldots, T_n\}$ are LSQ($\Psi$) functions $K \to K$ such that $\hat{F} := \bigcap_{i=1}^{n} Fix(T_i) \neq \partial$ and $T := T_n \circ \cdots \circ T_1$, then $Fix(T) \subseteq \hat{F}$, and if $Fix(T) \neq \partial$ then $T$ is RSQ($\Psi$) [147, 154, 132]. Such $\Psi$ will be called LSQ-compositional. If, additionally, $\text{efd}(\Psi) = X$, then we will call it LSQ-adapted. If $\Psi : X \to \mathbb{R}$ (i.e., bounded, uniformly continuous, and totally convex) on bounded subsets of $X$, $\partial \neq K \subseteq X$, $\{T_1, \ldots, T_n\}$ are RSQ($\Psi$) functions $K \to K$ such that $\hat{F} := \bigcap_{i=1}^{n} Fix(T_i) \neq \partial$ and $T := T_n \circ \cdots \circ T_1$, then $Fix(T) \subseteq \hat{F}$, and if $Fix(T) \neq \partial$ then $T$ is RSQ($\Psi$) [132]. Such $\Psi$ will be called RSQ-compositional. If, additionally, $\Psi$ is totally convex on $X$, $\Psi^F$ is totally convex on $\text{int}(\text{efd}(\Psi^F))$, and $D^G\Psi$ is weakly sequentially continuous, then we will call $\Psi$ RSQ-adapted.

The results of [19, 148, 131] imply [113]: (i) For any LSQ-adapted $\Psi$ and nonempty closed convex $K \subseteq \text{int}(\text{efd}(\Psi))$, $Fix(\overline{\Psi}^{D_\Psi}_{K^*}) = Fix(\overline{\Psi}^{D_\Psi}_{K}) = K$, hence $\overline{\Psi}^{D_\Psi}_{K}$ is LSQ($\Psi$); (ii) For any RSQ-adapted $\Psi$ and closed convex $\partial \neq M \subseteq \text{int}(\text{efd}(\Psi))$, $Fix(\overline{\Psi}^{D_\Psi}_{D^G\Psi^F(M)}) = Fix(\overline{\Psi}^{D_\Psi}_{D^G\Psi^F(M)}) = D^G\Psi^F(M)$, hence $\overline{\Psi}^{D_\Psi}_{D^G\Psi^F(M)}$ is RSQ($\Psi$).

3 $\ell$-operations and nonlinear resource theories of states

Given a set $Z \subseteq \text{int}(\text{efd}(\Psi))$, a set $U$, and a bijection $\ell : U \to Z$, we define the Bregman $\ell$-information on $U$ as $D_\Psi(\phi(\ell(\psi)), \psi) := D_\Psi(\ell(\phi), \psi)$ for $\phi, \psi \in U$. The properties of $D_\Psi$ can be naturally extended to the properties of $D_{\ell, \Psi}$, by turning $\ell$ into a homeomorphism. Given $\mathbb{C} \subseteq U$, if $\ell(\mathbb{C})$ is convex (resp., affine; closed), then $\mathbb{C}$ will be called $\ell$-convex (resp., $\ell$-affine; $\ell$-closed). So, the $\ell$-closure of $C \subseteq U$ is a closure of $C$ in the topology induced by $\ell$ from the norm topology of $X$. A left (resp., right) $D_{\ell, \Psi}$-projection is defined by $\overline{\Psi}^{D_\Psi}_{\ell(C)^*}(\psi) := \overline{\Psi}^{D_\Psi}_{\ell(C)}(\ell(\psi))$ (resp., $\overline{\Psi}^{D_\Psi}_{\ell(C)}(\psi) := \overline{\Psi}^{D_\Psi}_{\ell(C)}(\ell(\psi))$) for any $\ell$-closed $\ell$-convex (resp., $(D^G\Psi \circ \ell)$-closed $(D^G\Psi \circ \ell)$-convex) set $C$ and any
ψ ∈ U. For ∅ ≠ W ⊆ U and T : ℓ(W) → Z, T^ℓ : ℓ^{-1}∅ T ∘ ℓ : W → U will be called an ℓ-operation (or an ℓ-transmitter). The classes of D_ψ-quasi-nonexpansive maps on X determine the corresponding classes of ℓ-operations on U (i.e., (ℓ,Ψ)-transmitters), in particular: CN(ℓ,Ψ), LSQ(ℓ,Ψ), and RSQ(ℓ,Ψ). We will denote Fix(T^ℓ) := ℓ^{-1}(Fix(T)). Each y ∈ int(efd(Ψ^F)) defines an (ℓ,Ψ)-observable on U, given by y ∘ ℓ : U → R.3

Given a set U (of states), we define a resource theory of states (c.f., e.g., [94, 62, 54, 163]) as a triple (P, S, R), where P is a submonoid of a monoid End(U) of endomorphisms of U, ∅ ⊆ S ⊆ U satisfies P(S) ⊆ S, and R := {r : U → R^+ | r(φ) ≥ r(ψ) ∀φ ∈ U ∀ψ ∈ P}. The elements of P (resp., S; R) are called free operations (resp., free states; resource monotones). For example, in our setting, we have:

\[ (i \subset L)^R_{ℓ,Ψ} (T, S_T, \{D^R_{S,T}\}) \] (resp., (i \subset L)^R_{ℓ,Ψ} (T, S_T, \{D^R_{S,T}\})): if T is a submonoid of CN(ℓ,Ψ) with T(S_T) ⊆ S_T, and ∅ ≠ S_T ⊆ U is ℓ-closed ℓ-convex (resp., (Ω^GΨ o ℓ)-closed (Ω^GΨ o ℓ)-convex), then
\[ D^R_{S,T} := \inf_{φ ∈ S_T} \{D_ℓ(φ, ·)\} \] (resp., D^R_{S,T} := inf_{φ ∈ S_T} (D_ℓ(φ, ·), φ)) is a resource monotone; an interesting example is given by S_T = {φ ∈ U | ∀ψ ∈ U ℓT(ψ) = φ};

\[ (ii \subset L)^R_{ℓ,Ψ} (T, Fix(T), \bigcup_{φ ∈ T} \{D_ℓ(φ, ·)\}) \] (resp., (ii \subset L)^R_{ℓ,Ψ} (T, Fix(T), \bigcup_{φ ∈ T} (D_ℓ(φ, ·), φ))): if Ψ is LSQLQ-(resp., RSQ)-compositional, ∅ ⊆ K ⊆ U, T ⊆ LSQ(ℓ,Ψ) (resp., T ⊆ RSQ(ℓ,Ψ)) is a monoid such that T^ℓ : K → K ∀T^ℓ ∈ T, T^ℓ ≠ ∅ and Fix(T_1 ∩ ... ∩ T_n) ≠ ∅ ∀{T_1, ..., T_n} ⊆ T, then D_ℓ(ψ, ·) (resp., D_ℓ(·, φ)) is a resource monotone for any φ ∈ Fix(T).

In these examples D_ℓ(ψ) plays three different roles: it provides resource monotones, controls the behaviour of free operations, and participates in the construction of a space of free states. If Ψ is LSQLQ-adapted (resp., RSQ-adapted), then (iii \subset L)^R_{ℓ,Ψ} (resp., (iii \subset L)^R_{ℓ,Ψ}) is a special case of (ii \subset L)^R_{ℓ,Ψ} (resp., (ii \subset L)^R_{ℓ,Ψ})). Limited structural benefits of CN(Ψ) maps 4 allow to consider: Fix(T) as a generic notion of a free state space in Brègman resource theories; elements of LSQ(ℓ,Ψ) \ T and RSQ(ℓ,Ψ) \ T as the generic nonfree operations. The (linear) witnesses of S are defined as the elements of \{y ∈ int(efd(Ψ^F)) \ | [x, y] ≥ 0 for x ∈ S\}. Using [132], we observe that any set \{T_1, ..., T_m\}, m ∈ N, of resource theories of the type (ii \subset L)^R_{ℓ,Ψ} admits a convex envelope co{T_1, ..., T_m} := {T^ℓ ∈ RSQ(ℓ,Ψ) \ T : \sum_{i=1}^m w_i T_i, w_i ≥ 1, \{T_1, ..., T_m\} ⊆ T_1 ∪ ... ∪ T_m, (w_1, ..., w_n) ∈ [0,1]^n}, satisfying

D_ℓ(ψ, T^ℓ(ψ), φ) ≤ \sum_{i=1}^m w_i D_ℓ(ψ, T_1^ℓ(ψ), φ) ∀ψ, φ ∈ K. So, while co{T_1, ..., T_m} may not be a monoid

3If W is a convex set, then an ℓ-operation is belongs to Mielnik’s nonlinear transmitters [135, 136]. In our case, ℓ-convexity of W is more fundamental property than convexity, so T plays the role of a nonlinear transmitter, with (X, X^*) linearly representing (states/sources/resources, observables/sinks/witnesses), while T^ℓ is its nonlinear representation. The monoids of ℓ-operations can be seen as analogues of Mielnik’s semigroups of mobility [137], while (ℓ, )-observables provide an instance of Mielnik’s observables [136]. So, (ii \subset L)^R_{ℓ,Ψ} provide a weakened version of Mielnik’s nonlinear generalised quantum theory. (Mielnik identifies observables with any maps f : U → R, if U is convex. In our case one may consider such f as an ‘intensive’ observable, which is sampled in terms of ‘subjective’ ℓ-Ψ-observables; the same goes for nonlinear transmitters as well as U itself. §8.4.Ex.2 provides further development of this dialectic.)

4This generalization does not satisfy any of the properties that the classical nonexpansive operators do [154]. However, see §4.2.Ex.2(b) and its consequence in §5, providing a nontrivial intersection of (ℓ,Ψ)-transmitters T^ℓ with CTP maps (on preduals of any W^∗-algebras N). This shall be considered in parallel to a characterization of (preduals of) conditional expectations on (finite dimensional) N as right D_ℓ-projections [114], which, combined with §4.4.Ex.4, turns submonoids of (preduals of) conditional expectations into special cases of the type (ii \subset L)^R_{ℓ,Ψ} resource theory. Taking into account Chencov’s geometric approach [51] to Wald’s statistical decision theory [172, 173] (see §8), and Holevo’s approach [92, 93] to selection of POVM as a minimiser of a quantum statistical decision rule, one can view T^ℓ can as an analogue of a statistical decision rule.
(hence, not a resource theory of states), it provides a setting of a multi-resource theory of states, with its elements understood as (generally, nonfree) operations, decomposable (“tomographable”) into weighted mixtures of free operations from different individual resource theories. See [43] for further brégmanian resource theoretic discussion.

4 Examples of $D_{\ell, \Psi}$

**Ex.1.** (a) For any Banach space $X$, a duality map is defined as $j_{\varphi} : X \ni x \mapsto \{ y \in X^* \mid \varphi(x) = \langle x, y \rangle \}$, where $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$, called a gauge, is strictly increasing, continuous, with $\varphi(0) = 0$ and $\lim_{t \to \infty} \varphi(t) = \infty$. For any gauge $\varphi$, $\Psi_{\varphi} := \int_0^1 x \, d\varphi(t) : X \to \mathbb{R}^+$ is continuous, convex, and increasing [11, 180]. If $X$ is Gateaux differentiable, then $j_{\varphi} = \Psi_{\varphi}$ is Gateaux differentiable and strictly convex [11]. E.g., $\Psi_{\varphi} = \beta \cdot [1/\beta]$ (resp., $\| \cdot \|_X$) for $\varphi(t) = t^{1/\beta}$ (resp., $\frac{1}{\beta} t^{1/\beta-1}$) with $\beta \in [0, 1]$. In [113], using and extending the results of [180, 45, 181, 150], we prove: 1) $\Psi_{\varphi}$ is Legendre if $X$ is Gateaux differentiable and strictly convex; 2) $\Psi_{\varphi}$ is LSQ-adapted and RSQ-compositional (and $\mathcal{D}^G \Psi_{\varphi}$ is norm-to-norm continuous [150]) if $X$ is locally uniformly convex and uniformly Fréchet differentiable.

(b) For any base norm space $Y$, it is reflexive iff its base is weakly compact (see, e.g., [7]). In such case, some results of brégmanian theory apply directly, under weakening of $\mathcal{D}^G \Psi$ to right Gateaux derivative (so, Legendre duality is lost), and with $\ell = id_Y$ (more generally, $\ell$ be taken to be any automorphism of $Y$). This holds, in particular, for any finite dimensional $Y$, for type $I_2$ JBW-factors [168] (which are exactly the spin factors $\mathcal{H} \oplus \mathcal{R}$, where $\mathcal{H}$ is a Hilbert space with dim $\mathcal{H} \geq 2$ [159]), as well as for state spaces of orthomodular posets satisfying Jordan–Hahn decomposition property [76].

(c) Given any base norm space $Y$, if $U$ is a generating positive cone of $Y$ and $\ell$ is a map from $U$ (or $Y$) into a geometrically well-behaved Banach space $X$, then $D_{\Psi}$ determines $D_{\ell, \Psi}$ on $U$ (or $Y$) and thus also on the base of $Y$. Ex.2 and Ex.3 provide the special cases of this situation, with $X$ implementing Ex.1(a) (for any $W^*$-algebra $\mathcal{N}$, using the uniqueness of a polar decomposition, we extend the bijective embedding to the whole predual space $\mathcal{N}^*$, under replacement of $\| \cdot \|$ by re $\| \cdot \|$ in all formulas), calling for further investigation of convexity and differentiability of Banach spaces of integrals over general base norm spaces [165, 166, 167].

**Ex.2.** (a) If $A$ is a semi-finite JBW-algebra with a faithful normal semi-finite trace $\tau : A^+ \to [0, \infty]$, then the nonassociative $L_{1/\gamma}$ spaces, $\gamma \in [0, 1]$, defined by $(L_{1/\gamma}(A, \tau) := \mathcal{A}_{\tau}^{-1} L_{1/\gamma}, \| \cdot \|_{1/\gamma} := (\tau((\| \cdot \|_{1/\gamma})^\gamma))$, where $A_{\tau} := \{ x \in A^+ \mid \tau(x) < \infty \}$ [1, 2, 98], are uniformly convex and uniformly Fréchet differentiable for $\gamma \in [0, 1]$ [99]. Hence, for any gauge $\varphi$, $\Psi_{\varphi, \gamma} := \int_0^{1/\gamma} d\varphi(t)$ is LSQ-adapted and RSQ-compositional. By means of $\phi = \tau(h_{\phi} \circ \cdot)$, the nonassociative Mazur map $\ell_{\phi} : A_{\tau} \ni \phi \mapsto sign(h_{\phi}) h_{\phi}^\gamma \in L_{1/\gamma}(A, \tau)$ determines $D_{\gamma, \varphi} := D_{\ell_{\phi}}, \Psi_{\varphi, \gamma} : A_\tau \times A_\tau \to [0, \infty]$. Due to isometric isomorphism of $L_{1/\gamma}(A, \tau)$ for different $\tau$ [12], $D_{\gamma, \varphi}$ do not depend on $\tau$. For $\beta, \gamma \in [0, 1]$ and $\omega, \phi \in A_{\gamma}$, $\varphi(t) = t^{1/\beta-1}/\beta$ yields $D_{\gamma, \beta}(\omega, \phi) = (\tau(\omega))^{\gamma/\beta} + \frac{1}{\beta} (\tau(\omega))^{\gamma/\beta} - \frac{1}{\beta} (\tau(h_{\phi}))^{\gamma/\beta-1} (h_{\phi})^\gamma \phi(1/\gamma)$ for $\omega \ll \phi$ and $\infty$ otherwise. (b) The same (including the formula for $D_{\gamma, \beta}$) holds for any $W^*$-algebra $\mathcal{N}$ and $\gamma \in [0, 1]$, due to uniform convexity and uniform Fréchet differentiability of $(L_{1/\gamma}(\mathcal{N}), \| \cdot \|_{1/\gamma})$ spaces [164, 133, 110], under replacement of $\phi = \tau(h_{\phi} \circ \cdot)$ by $\phi = f \phi$, and with $\ell_{\phi} : \mathcal{N} \ni \phi \mapsto u_{\phi}(\phi) \in L_{1/\gamma}(\mathcal{N})$, where $\phi = \phi(1/\gamma) \phi(1-\gamma)$ is a polar decomposition, while $f$ and $\| \cdot \|$ are defined as in [75]. For $\varphi(t) = t^{(1-\gamma)/\gamma(1-\gamma)} := \varphi_0(t)$ we obtain $D_{\gamma}(\omega, \phi) = (\gamma(1-\gamma))^{-1} \int (\gamma \phi + (1-\gamma) \phi - \omega \phi(1/\gamma))$ whenever $\omega \ll \phi$ and $\infty$ otherwise, introduced in [104, 112], and unifying $D_{\gamma}$ of [127, 87]. All CPTP maps are completely $D_{\gamma}$-nonexpansive [104], so, the resource theories $(\ell_{\gamma, \tau}, \Psi_{\varphi, \gamma})$ are valid submonoids (hence, resource theories) of CPTP maps.

**Ex.3.** Given a semi-finite $W^*$-algebra $\mathcal{N}$ with a faithful normal semi-finite trace $\tau : \mathcal{N}^+ \to [0, \infty]$, let $\mathcal{M}(\mathcal{N}, \tau)$ denote the topological $*$-algebra of $\tau$-measurable operators affiliated with $\mathcal{N}$. For any Orlicz function $T$, a noncommutative Orlicz space $(L_T(\mathcal{N}, \tau) := \{ x \in \mathcal{M}(\mathcal{N}, \tau) \mid \exists \lambda >$
Legendre transform allows us to define the categories $C$, $\text{Hom}$, and $D$ of orthogonal projection operators $\{\in \Upsilon\}$ on $\text{L}(N, \tau)$, where $\Upsilon(Y) := \sup\{x|y| - Y(x) \mid x \geq 0\}$. So, $\Psi_{\tau}, \Upsilon := \int_{\tau} d\Psi(t) : \text{L}(N, \tau) \rightarrow \text{R}^+$ is LSQ-adapted and RSQ-configuration.

Introducing noncommutative Kaczmarz map $\tau : \mathbb{N}^+ \ni \phi \rightarrow \Upsilon^{-1}(\phi)$ in $\text{L}(N, \tau)^+$, we obtain the family $D_{\tau, \Upsilon} := D_{\tau, \Psi_{\tau}, \Upsilon}$. Due to [13], it is independent of $\tau$. For $N = L_\infty(K, \mu)$, $\tau := \int_{\mu} \tau(t) > 0 \tau(t) > 0 \Rightarrow (\Omega := \frac{d\tau}{d\mu})$, $|\phi|_1 = |\phi|_1 = 1$, $\nu(t) = t^{1/1-\beta}/\beta$, $\beta \in [0, 1]$, and $\Upsilon(\phi, \omega) := \int_{\mu} \tau(t) \Upsilon^{\tau}(\phi)$, this gives $D_{\tau, \beta}(\omega, \phi) = \frac{1}{\beta} - \frac{1}{\beta} \Upsilon(\omega, \phi)/\Upsilon(\varphi, \varphi)$.

**Example 4.** For a Hilbert space $H$, $\dim H = n < \infty$, Umegaki's information $D_1(\rho, \phi) := \log(h_{\rho}(\log h_{\rho} - \log h_{\rho} - h_{\rho} - h_{\rho}) = \frac{1}{\beta} - \frac{1}{\beta} \Upsilon(\omega, \phi)/\Upsilon(\varphi, \varphi) = 0$ [170] equals $D_{\text{cl, rep}, \Psi_{\text{cl}, \Omega}}(\rho, \phi)$, where $\tau = \text{tr}_H(h_{\rho}) \in \Omega(H)^+$, $\lambda$ is a nonincreasing rearrangement of eigenvalues, while $\Psi(x) := \sum_{i=1}^{\infty} (x_i \log(x_i) - x_i)$ for $x \geq 0$ and $\infty$ otherwise [17]. (This extends to a separable dim $H = \infty$ case via [29].) So, Lüders' and quantum Jeffrey's rules [89], partial trace [141], and (preduals of) conditional expectations [114], as special cases of $D_{\text{cl}, \Omega}$, belong to $D_{\text{cl}, \Omega}$.

## 5 Categories

In what follows, $Z = \text{int}(\text{efl}(\Psi))$. We define the category $\text{Cvx}(\ell, \Psi)$ (resp., $\text{Aff}(\ell, \Psi)$), with objects given by $\ell$-closed $\ell$-convex (resp., $\ell$-closed $\ell$-affine) subsets of $U$, and including $\emptyset$, morphisms given by left $D_{\text{cl}}$-projectors onto $\ell$-closed $\ell$-convex (resp., $\ell$-closed $\ell$-affine) subsets of these subspaces (i.e., $\text{Hom}_{\text{Cvx}}(\ell, \Psi)(C, C)$ consists of $D_{\text{cl}}^D \circ D_{\text{cl}}^{K}$ with $K$ varying over all $\ell$-closed $\ell$-convex subsets of $C$), including $\emptyset$ (resulting in empty arrows, e.g., $\emptyset \in \text{Hom}_{\text{Cvx}}(\ell, \Psi)(C_1, C_2)$), and composition given by $D_{\text{cl}}^D \circ D_{\text{cl}}^{K} := D_{\text{cl}}^{D \circ K}$. Legendre transform allows us to define the categories $\text{Cvx}(\ell, \Psi)$ (resp., $\text{Aff}(\ell, \Psi)$), with objects given by $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$-closed ($\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$)-convex (resp., ($\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$)-closed ($\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$)-affine) subsets of $U$, including $\emptyset$, morphisms given by right $D_{\Psi}$-projections onto $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$-closed $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$-convex (resp., $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$-closed $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell)$)-affine subsets of these subsets, including $\emptyset$, and composition given by $D_{\text{cl}}^{D \circ K} := D_{\text{cl}}^{D} \circ D_{\text{cl}}^{K} : D_{\text{cl}}^{D} \circ D_{\text{cl}}^{K} = D_{\text{cl}}^{D \circ K}$.

---

Let $I \subseteq \mathbb{R}$ be an interval. We call $\text{Y} : \mathbb{R} \rightarrow \mathbb{R}$ to be **Orlicz** iff it is convex, with $\Upsilon(0) = 0$, $\Upsilon \neq 0$, and $\Upsilon(-u) = \Upsilon(u)$. An Orlicz $\Upsilon$ belongs to $\Delta_2$ iff $\sup_{u \geq 0} \Upsilon^{\tau}(u) < \infty$; $\Delta_\infty$ iff $\lim_{u \rightarrow \infty} \Upsilon^{\tau}(u) < \infty$; $\Delta^\varphi$ iff $\lim_{u \rightarrow \infty} \Upsilon^{\tau}(u) < \infty$; $U_{\text{C}}$ iff $\forall u \in [0, 1]$ $\exists \theta(a) \in [0, 1]$ $\forall u \in I$ $f(\frac{u^{\theta(a)}}{f(u)}) \leq 1 - (1 - \delta(u))(f(u) + f(u))$; $U_{\text{C}}$ (resp., $U_{\text{C}}$) iff $\exists \theta(a) > 0$ such that $\forall u \in [0, 1]$ $\forall \in U_{\text{C}}([0, 1])$ (resp., $\forall u \in U_{\text{C}}([0, 1])$). $\text{C}(\text{I})$ iff $\text{Y}$ is strictly convex on $I$.

The composition rule $\circ$ for left $D_{\text{cl}}$-projections is well defined and stable also in the computational sense. Its quantitative evaluation can be performed by means of an algorithm given in [20] (valid for any countable family $\{K_i\}_{i \in \mathbb{N}}$ and any $\Psi$ that is totally convex on bounded sets, hence, in particular, for any LSQ-adapted $\Psi$), or by means of [21, 36, 65] (valid for $\text{dim } H < \infty$, a finite family $(K_i)_{i \in \{1, \ldots, n\}}$, and Legendre $\Psi$ satisfying some additional conditions). For $X$ given by the Hilbert space $H$ and $\text{C} \circ \Psi_{\text{cl}, \Omega}(\ell) \subseteq \frac{1}{\beta} - \frac{1}{\beta} \Upsilon(\omega, \phi)/\Upsilon(\varphi, \varphi)$, the former algorithm turns to Haagaeu's [88] algorithm, while the latter turns to Dysktra's algorithm [69, 32, 86] (valid also for $\text{dim } H = \infty$, and extendable to countable families $\{K_i\}_{i \in \mathbb{N}}$). Under further restriction of $\{K_i\}$ to a finite family of closed linear subspaces of $H$, $D_{\text{cl}}^{K_{i \in \mathbb{N}}} \circ D_{\text{cl}}^{K_{i \in \mathbb{N}}}$ turn into orthogonal projection operators $P_{K_i} : H \rightarrow K_i$, while Dysktra's algorithm turns into Halperin's theorem [85] on strong convergence of a cyclic repetition of $P_{K_1} \cdots P_{K_i} \cdots P_{K_{i \in \mathbb{N}}} = \lim_{i \rightarrow \infty} ||(P_{K_1} \cdots P_{K_i}) - P_{K_{i+1}}|| = 0$ $\forall \O \in H$. When only two projections are considered, corresponding to a composition $D_{\text{cl}}^{K_{i+1}} \circ D_{\text{cl}}^{K_i}$ for linear subspaces $K_1$ and $K_2$, this becomes the von Neumann–Kakutani theorem [171, 108]. All these algorithms provide evaluation of the (finite or countable) left $D_{\text{cl}}$-projection $D_{\text{cl}}^{K_{i \in \mathbb{N}}}(x)$ in terms of a norm convergence of a cyclic sequence of algorithmic steps to the unique limit point. The differences in definitions of those algorithms correspond to different ranges of generality. In particular, while the direct extension on the von Neumann–Kakutani algorithm to closed convex sets converges weakly to an element in the nonempty intersection of $K_1$ and $K_2$ [34] (Kaczmarz's algorithm [106] is a special case of this extension, obtained for hyperplanes and $\text{dim } H = \infty$), the limit point may be not equal to a projection onto $K_1 \cap K_2$ [57] and the norm convergence generally does not hold [95], although the latter holds always for $\text{dim } H < \infty$, and can be guaranteed under additional conditions for $\text{dim } H = \infty$ [84]. On the other hand, the direct extension of Halperin's theorem to linear projections, of norm equal to 1, onto subspaces of uniformly convex Banach space is norm convergent and returns a projection, of norm equal to 1, onto an intersection [41]. For noncyclic algorithms, see [37, 143, 9, 40, 96, 17, 20].
Following Jaynes [102, 103], we consider an empty (resp., identity) arrow as an inference corresponding to overdetermination (resp., underdetermination) of constraints. $K_1 \cap K_2 = K_2 \cap K_1$ implies commutativity of $\circ$. Under restriction of composition by the condition $K_2 \subseteq K_1$, the infinitary algorithmic aspect of computation of $\circ$ can be dropped, defining the convenient categories $\text{Cvx}(\ell, \Psi)$, $\text{Aff}(\ell, \Psi)$, $\text{Cvx}(\ell, \Psi)$, and $\text{Aff}(\ell, \Psi)$. On the other hand, by dropping $\ell$-embeddings everywhere (i.e., moving to $\text{D}_\Psi$-projections on $X$), we obtain the categories $\text{Cvx}(\Psi)$, $\text{Aff}(\Psi)$, $\text{Cvx}(\Psi)$, and $\text{Aff}(\Psi)$, as well as their $\leq$-subcategories.

Composability of $\text{LSQ}(\Psi)$ (resp., $\text{RSQ}(\Psi)$) maps allows to define the category $\text{LSQ}(\Psi)$ (resp., $\text{RSQ}(\Psi)$) of subsets of $\text{int}(\text{efd}(\Psi))$, with elements of $\text{LSQ}(\Psi)$ (resp., $\text{RSQ}(\Psi)$) as arrows between them, including empty set as object and empty arrows as morphisms (via $\ell$-embedding, this gives $\text{LSQ}(\ell, \Psi)$ (resp., $\text{RSQ}(\ell, \Psi)$)). In general, the composition in $\text{LSQ}(\Psi)$, $\text{LSQ}(\ell, \Psi)$, $\text{RSQ}(\Psi)$ and $\text{RSQ}(\ell, \Psi)$ is not commutative, and their objects are not convex in any sense. Restriction of $\text{LSQ}(\Psi)$ (resp., $\text{LSQ}(\ell, \Psi)$; $\text{RSQ}(\Psi)$; $\text{RSQ}(\ell, \Psi)$) to objects given by the closed convex (resp., $\ell$-closed $\ell$-convex; $\mathcal{D}^\Psi$-closed $\mathcal{D}^\Psi$-convex; $\mathcal{D}^\Psi \cup \ell$-closed $\mathcal{D}^\Psi \cup \ell$-convex) sets determines a subcategory $\text{LSQ}_{\text{cvx}}(\Psi)$ (resp., $\text{LSQ}_{\text{cvx}}(\ell, \Psi)$; $\text{RSQ}_{\text{cvx}}(\Psi)$; $\text{RSQ}_{\text{cvx}}(\ell, \Psi)$). Taking subsets of $X$ as objects and elements of $\text{CN}(\Psi)$ as arrows defines $\text{CN}(\Psi)$ (and $\text{CN}(\ell, \Psi)$, via $\ell$). If $X$ is a real Hilbert space and $\Psi = \frac{1}{2} |x|^2$, then $\text{CN}(\Psi)$ coincides with the category $\text{Hi1H}_{\mathbb{R}}(X)$ of real Hilbert subspaces of $X$ and completely $\left|\cdot\right|$-nonexpansive maps. From §4.Ex.2(b) it follows that $\text{CN}(\ell_i/\gamma, \Psi_{\varphi_i})$, $\gamma \in [0, 1]$, coincides with the category of all CPTP maps on $\mathcal{N}_\Psi$.

In analogy to Chencov’s approach [49, 51, 139] (generalising Blackwell’s [28] statistical equivalence), we will call the subsets $M_1$ and $M_2$ of $\ell^{-1}(\text{int}(\text{efd}(\Psi)))$ to be left equivalent if $\exists T_1, T_2 \in \text{Arr}(\text{LSQ}(\ell, \Psi))$ such that $T_1(M_1) = M_2$ and $T_2(M_2) = M_1$. Hence, the families of left equivalent subsets of $U$ coincide with the groups $\text{LSQ}(\ell, \Psi)$. Let $\text{LSQ}(\Theta, \ell, \Psi)$ be a subcategory of $\text{LSQ}(\ell, \Psi)$ such that each of its objects is bijectively parametrised by a set $\Theta$. Given $M_1, M_2 \subseteq U$, and a set $\Theta$, assume that there exist bijections $\theta_1 : \Theta \to M_1$ and $\theta_2 : \Theta \to M_2$. Adapting linear positive constructions [123, 144], we define a left $(\epsilon, \text{D}_\ell, \Psi)$-deficiency of $M_2$ with respect to $M_1$ as existence of such $T \in \text{Hom}_{\text{LSQ}(\Theta, \ell, \Psi)}(M_1, \cdot)$ that $\sup_{\theta \in \Theta} D_{\ell, \Psi}(\theta_2(\theta), T(\theta_1(\theta))) = \epsilon$: a left $\text{D}_\ell, \Psi$-deficiency of $M_2$ with respect to $M_1$ as $\delta_{\ell, \Psi}(M_2, M_1) := \inf_{T \in \Theta} \sup_{\theta \in \Theta} D_{\ell, \Psi}(\theta_2(\theta), T(\theta_1(\theta)))$, where $H := \text{Hom}_{\text{LSQ}(\Theta, \ell, \Psi)}(M_1, \cdot)$; a mutual left $\text{D}_\ell, \Psi$-deficiency of $M_1$ and $M_2$ as $\delta_{\ell, \Psi}(M_1, M_2) := \max\{\delta_{\ell, \Psi}(M_2, M_1), \delta_{\ell, \Psi}(M_1, M_2)\}$ (by definition, it is symmetric). Given $M_1, M_2, M_3 \in \text{Ob}(\text{LSQ}(\Theta, \ell, \Psi))$, if $\text{Hom}_{\text{LSQ}(\Theta, \ell, \Psi)}(M_1, M_2) \neq \emptyset$ then $\delta_{\ell, \Psi}(M_3, M_2) \leq \delta_{\ell, \Psi}(M_3, M_1)$. If $M_1$ and $M_2$ are left equivalent, then $\delta_{\ell, \Psi}(M_1, M_2) = 0$. Hence, all objects of a single groupoid in $\text{LSQ}(\Theta, \ell, \Psi)$ have zero mutual left $\text{D}_\ell, \Psi$-deficiency, yet the latter is nonzero between any elements of two distinct groupoids. All of these constructions have their right versions.

The existence and uniqueness of $\bigcup_{D_\Psi} y$ does not require norm boundedness (and thus weak

---

8If $X$ is separable, then $\text{Aff}(\Psi)$ has objects given by the countable sets of polynomial equations as data types, and morphisms given by programs (algorithms) that translate them (their solutions). More generally, if $X$ is a separable Banach space, then every convex closed subset $C \subseteq X$ is the intersection of the countable number of its supporting closed half-spaces [26], i.e., it is a (countable) polyhedron, which is the set of solutions for a countable system of linear inequalities (see [30] for a discussion of the nonseparable case). Hence, also $\text{Cvx}(\Psi)$, at least in the separable case, can be represented as a category of specific data types and computations between them. The resource theory (iii)_{\ell, \Psi} from §3 can be recast as a subcategory $\text{Cvx}(\ell, \Psi)$ of $\text{Cvx}(\ell, \Psi)$, determined by the choice of its terminal object to be given by $K$ (so the left $\text{D}_\Psi$-projections onto subsets of $K$ are not considered). In such cases, the free sets of every object $A$ in $\text{Cvx}(\ell, \Psi)$ correspond to the set $\text{Hom}_{\text{Cvx}(\ell, \Psi)}(K, A)$, which can be seen as an analogue of the fact that $\text{Hom}_{\text{Ban}_\mathbb{R}}(\mathbb{R}, Z)$ is equal to the unit ball of a real Banach space $Z$ [55], where $Z$ is a terminal (and also initial, hence zero) object in the category $\text{Ban}_\mathbb{R}$ of real Banach spaces and completely $\left|\cdot\right|$-nonexpansive maps [157]. Each $K \in \text{Ob}(\text{Aff}(\Psi))$ with $\text{codim}(K) = 1$ determines a hyperplane in $X$, which can be seen as a resource witness.

9While the cyclic algorithms mentioned in §5 exhibit norm convergence in $X$, one still may need either to have a refined quantification of the exactness of intermediate steps, or to quantify the convergence of algorithms with worse convergence behaviours. In such cases left $(\epsilon, \text{D}_\Psi)$-deficiency can be used to quantify the approximate exactness of a $k$-th cycle of computation of a left $\text{D}_\Psi$-projection onto (finite or countable) intersection $M_2 = K_1 \cap \cdots \cap K_\ell$, or, more generally, any cyclic convergence algorithm, given $k \in \mathbb{N}$, with $T := S^k$, where $S \in \text{LSQ}(\Psi)$ with $S : M_1 \to \text{int}(\text{efd}(\Psi))$. This illustrates a key property of $\text{D}_\Psi$ that underlies the flexibility of its applications: it allows to quantify both algorithmic and structural aspects of the suitable category of spaces, serving as a control interface between arithmetic and geometric layers of a theory.
compactness) of $Q$, due to coercivity of $D_\Psi(\cdot, y)$ (c.f. Remark 2.13 in 4th ed. of [15] and Lemma 7.3.(v) in [18]). Nevertheless, we can consider a subcategory $\text{CmpCvx}(\Psi)$ of $\text{Cvx}(\Psi)$, consisting of norm bounded, norm closed, convex (equivalently: convex and weakly compact) subsets of $X$ as objects and left $D_\Psi$-projections onto their subobjects as arrows (including empty set and empty arrows). The corresponding $r_\ast$, $\ell_\ast$, and $\subseteq_\ast$ versions of this category are defined analogously as for $\text{Cvx}(\Psi)$. For every $K \in \text{Ob}(\text{CmpCvx}(\ell, \Psi))$ we can canonically associate an order unit Banach space $A(K)$ of all continuous real valued affine functions on $K$ [107], as well as a base norm space $(A(K))^\ast$, together with an affine homeomorphism of $K$ onto the base of $(A(K))^\ast$ (extending to a linear isomorphism of $\bigcup_{n=1}^\infty \text{nco}(K \cup -K)$ onto $(A(K))^\ast$) [70], as well as a canonical embedding of $A(K)$ into an order unit Banach space $(A(K))^\ast\ast$ [72] (the latter is equal to the space of all bounded real valued affine functions on $K$ with the supremum norm). Hence, each $K \in \text{Ob}(\text{CmpCvx}(\ell, \Psi))$ determines a convex operational model in the sense of [60] (which is a special case [83] of Mielnik’s theory of linear transmitters [135, 136]). In consequence, $\text{CmpCvx}(\ell, \Psi)$ provides a specific nonlinear analogue of the category of convex operational models and positive linear maps with positive duals considered in [16].

6 Functors

We assume $\text{int}(\text{efd}(\Psi)) = X$. From above definitions it follows that every $\ell$ determines a (family of) functor(s), acting by $K \to \ell(K)$ on objects and $T \to T^\ell$ on arrows, which, together with the functor $\ell^{-1}$, establishes the equivalences of corresponding categories. If $\Psi$ is LSQ-adapted (resp., RSQ-adapted), then an embedding functor $\iota_\ell^\Psi : \text{Cvx}(\Psi) \to \text{LSQ}_{\text{cvx}}(\Psi)$ (resp., $\iota_\ell^R : \text{Cvx}(\Psi) \to \text{RSQ}_{\text{cvx}}(\Psi)$) and an induced functor $\iota_{\ell, \Psi}^L := \ell^{-1} \circ \iota_\ell^L : \text{Cvx}(\ell, \Psi) \to \text{LSQ}_{\text{cvx}}(\ell, \Psi)$ (resp., $\iota_{\ell, \Psi}^R := \ell^{-1} \circ \iota_\ell^R : \text{Cvx}(\ell, \Psi) \to \text{RSQ}_{\text{cvx}}(\ell, \Psi)$) are well defined, due to Fix$(\overline{\mathcal{D}^D_{\Psi}} \circ \overline{\mathcal{D}^D_{\Psi}}) = \overline{\text{Fix}(\overline{\mathcal{D}^D_{\Psi}}) \cap \text{Fix}(\overline{\mathcal{D}^D_{\Psi}})} = Q_1 \cap Q_2$. Given any set $Y$, let $\text{Pow}(Y)$ denote the category of all subsets of $Y$ with functions between them as morphisms. Consider a map $\text{co}_{\ell, \Psi}^L : \text{Ob}(\text{Pow}(X)) \to \text{Ob}(\text{Pow}(X))$, assigning to each subset $Y$ of a Banach space $X$ the closure of a convex hull $\text{co}(Y)$ of $Y$ in the weak topology of $X$ (it coincides with the norm closure of $\text{co}(Y)$). Let $\text{co}_{\ell, \Psi}^R : \text{Arr}(\text{Pow}(X)) \to \text{Arr}(\text{Pow}(X))$ be a map that assigns to each function $f : Y_1 \to Y_2$ a map $\overline{\mathcal{P}^Q_{\mathcal{D}^D_{\Psi}}} : \text{co}_{\ell, \Psi}^L(Y_1) \to \text{co}_{\ell, \Psi}^L(Y_2)$, where $Q = \text{co}_{\ell, \Psi}^L(f(Y_1))$. Then $\text{co}_{\ell, \Psi}^L : \text{Pow}(X) \to \text{Cvx}(\Psi)$ is a functor. Let $\text{co}_{\ell, \Psi}^R$ be a functor assigning: to each $Y \in \text{Ob}(\text{Pow}(X))$ an image of $\mathcal{D}^G_{\Psi} F$ of the weak closure of the convex hull of $\mathcal{D}^G_{\Psi}(Y)$; to each $f : Y_1 \to Y_2$, a map $\overline{\mathcal{P}^Q_{\mathcal{D}^D_{\Psi}}} : \text{co}_{\ell, \Psi}^R(Y_1) \to \text{co}_{\ell, \Psi}^R(Y_2)$, where $Q = \text{co}_{\ell, \Psi}^R(f(Y_1))$. With a forgetful functor $\text{Fr}_{\text{Set}} : \text{Cvx}(\Psi) \to \text{Pow}(X)$ (resp., $\text{Cvx}(\ell, \Psi) \to \text{Pow}(X)$), defined by forgetting convex and topological structure, we obtain an adjunction $\text{co}_{\ell, \Psi}^L : \text{Pow}(X) \dashv \text{Fr}_{\text{Set}}$ (resp., $\text{co}_{\ell, \Psi}^R : \text{Pow}(X) \dashv \text{Fr}_{\text{Set}}$). If $\Psi$ is LSQ-adapted, then, a mapping $\text{Fix}_{\ell, \Psi}^L$, defined by identity on objects of $\text{LSQ}_{\text{cvx}}(\Psi)$ and assigning $T \to \overline{\mathcal{P}^Q_{\text{Fix}(T)}}$ to each $T \in \text{Arr}(\text{LSQ}_{\text{cvx}}(\Psi))$, is a functor $\text{LSQ}_{\text{cvx}}(\Psi) \to \text{Cvx}(\Psi)$, satisfying $\iota_{\ell, \Psi}^L \dashv \text{Fix}_{\ell, \Psi}^L$. By composition, we obtain $\iota_{\ell, \Psi}^L \circ \text{co}_{\ell, \Psi}^L : \text{Pow}(U) \to \text{Cvx}(\ell, \Psi)$, $\iota_{\ell, \Psi}^R : \text{Cvx}(\ell, \Psi) \to \text{LSQ}_{\text{cvx}}(\ell, \Psi)$, $\iota_{\ell, \Psi}^R : \text{Cvx}(\ell, \Psi) \to \text{Pow}(U)$, and the respective adjunctions. If $\Psi$ is RSQ-adapted, then a mapping $\text{Fix}_{\ell, \Psi}^R$, defined by identity on objects of $\text{RSQ}_{\text{cvx}}(\Psi)$ and assigning $T \to \overline{\mathcal{P}^Q_{\text{Fix}(T)}}$ to each $T \in \text{Arr}(\text{RSQ}_{\text{cvx}}(\Psi))$, is a functor $\text{RSQ}_{\text{cvx}}(\Psi) \to \text{rCvx}(\Psi)$, satisfying $\iota_{\ell, \Psi}^R \dashv \text{Fix}_{\ell, \Psi}^R$. By composition with $\ell$, we obtain the functors $\text{co}_{\ell, \Psi}^L : \text{Pow}(U) \to \text{Cvx}(\ell, \Psi)$, $\iota_{\ell, \Psi}^L : \text{Cvx}(\ell, \Psi) \to \text{LSQ}_{\text{cvx}}(\ell, \Psi)$, $\iota_{\ell, \Psi}^R : \text{Cvx}(\ell, \Psi) \to \text{Pow}(U)$, and the respective adjunctions. If $\Psi$ is such that both LSQ$(\Psi)$ and RSQ$(\Psi)$ are well defined, and assuming additionally [132] that $\mathcal{D}^G_{\Psi}$ and $\mathcal{D}^G_{\Psi} F$ are (bounded and uniformly continuous) on bounded sets of $\text{int}(\text{efd}(\Psi))$ and $\text{int}(\text{efd}(\Psi^F))$, respectively, the Legendre maps determine an equivalence of categories, given by a pair of functors: $(\cdot)^\Psi : \text{RSQ}(\Psi) \to \text{LSQ}(\Psi)$ and $(\cdot)^{\Psi F} : \text{LSQ}(\Psi) \to \text{RSQ}(\Psi)$, acting by $C \mapsto \mathcal{D}^G_{\Psi}(C)$ and $K \mapsto \mathcal{D}^G_{\Psi} F(K)$ on objects, and by
conjugations $T \mapsto T^\Psi$ and $T \mapsto T^{\Psi_F}$ on morphisms, respectively. The same definition of $(\cdot)^\Psi$ and $(\cdot)^{\Psi_F}$, without extra conditions on $\Psi$, gives an equivalence of $\text{1Cvx}(\Psi)$ and $\text{rCvx}(\Psi)$.

7 Natural transformations and Hom-monoids

Let $[0, \infty]$ denote a category consisting of one object $\bullet$, with morphisms given by the elements of the set $\mathbb{R}^+ \cup \{\infty\}$, and their composition defined by addition [121]. Let $2$ denote the category consisting of two objects, one arrow between them, and the identity arrow on each of the objects. The category $[0, \infty]^2$ has morphisms of $[0, \infty]$ as objects, commutative squares in $[0, \infty]$ as morphisms, and commutative compositions of these squares as compositions. Let $K_1, K_2, K_3, K, L \in \text{Ob}(\text{1Aff}_{\text{lAff}}(\Psi))$, $K \subseteq K_2$ and $L \subseteq K_3$. For each $\phi \in Q$, left pythagorean equation implies the commutativity of the diagram (1). This defines a contravariant functor $D_\phi(\phi, \cdot) : \text{1Aff}_{\text{lAff}}(\Psi) \to [0, \infty]^2$, which naturally extends to a functor $D_\phi(\phi, \cdot) : \text{1Aff}^\subseteq(\Psi) \downarrow Q \to [0, \infty]^2$, where $\text{1Aff}^\subseteq(\Psi) \downarrow Q$ denotes a slice category of $\text{1Aff}^\subseteq(\Psi)$ over $Q$. For any two categories $\mathcal{C}$ and $\mathcal{D}$, cartesian closedness of the category $\text{Cat}$ of all small categories (with natural transformations as morphisms) implies that any functor $\mathcal{C} \to \mathcal{D}$ corresponds to a natural transformation in $\mathcal{D}$. Hence, $Q$ parametrises the family of natural transformations $D_\phi(\phi, \cdot)$ in the category of functors $\text{1Aff}^\subseteq(\Psi) \downarrow Q \to [0, \infty]$. Dependence of $D_\phi(\phi, \cdot)$ on $Q$ can be factored out by reducing considerations to singletons $Q = \{\phi\}$ (understood as 0-dimensional closed affine spaces). In (some) analogy to [14, 79], this allows us to state a problem of characterisation of $D_\phi$ as a natural transformation $D_\phi(\phi, \cdot)$.

\[ \begin{array}{cccc} x & \downarrow D_\phi(\phi, x) & \downarrow D_\phi(\phi, x) & \downarrow D_\phi(\phi, x) \\ \leftarrow D_{\phi, \psi}^D & \leftarrow D_{\phi, \psi}^D & \leftarrow D_{\phi, \psi}^D & \leftarrow D_{\phi, \psi}^D \\ \end{array} \]

Given any $Q \in \text{Ob}(\text{1Cvx}(\Psi))$, $\text{Hom}_{\text{1Cvx}(\Psi)}(\cdot, Q)$ can be equipped with the structure of a commutative partially ordered monoid [74], with $\overrightarrow{\mathcal{D}_Q} \circ \overrightarrow{\mathcal{D}_Q} := \overrightarrow{\mathcal{D}_Q \cup \mathcal{Q}_Q}$, $\overrightarrow{\mathcal{D}_Q} \leq \overrightarrow{\mathcal{D}_Q} := \overrightarrow{Q_1 \cup \mathcal{Q}_Q}$, and a distinguished zero object, given by $\overrightarrow{\mathcal{D}_Q}$. (Examples of computation of $\circ$ given in §5 apply here as well.) Hence, each $\text{Hom}_{\text{1Cvx}(\Psi)}(\cdot, Q)$ forms a resource theory in the sense of [78] (which generalises, in particular, the approaches of [126] and [64]). Viewing the order of extended positive reals as a feature distinct from their composition by addition turns $[0, \infty]$ into a commutative partially ordered monoid (with $x + \infty = \infty = \infty + x \forall x \neq \infty$). Thus, each functor $D_\phi(\phi, \cdot)$ can be seen as a morphism $\text{Hom}_{\text{1Aff}^\subseteq(\Psi)}(\cdot, Q) \to [0, \infty]$ inside the category of commutative partially ordered monoids.

(By Legendre duality, right pythagorean equation, and $\ell^{-1}$, the above applies also to categories of $\overrightarrow{\mathcal{D}_Q}$, $\overrightarrow{\mathcal{D}_Q}$, and $\overrightarrow{\mathcal{D}_Q}$.)

This equivalence may seem trivial, as built into the definition of $\text{rCvx}(\Psi)$. Yet, we see it is as a top of an iceberg: there exist right $D_\phi$-projections which are not Legendre transforms of the left $D_\phi$-projections [22], the equivalence between $\text{LSQ}(\Psi)$ and $\text{RSQ}(\Psi)$ classes holds only under special conditions [132], and there is an important difference between availability of LSQ- vs RSQ-adaptedness in models. Furthermore, while $\overrightarrow{\mathcal{D}_Q}$ correspond to Sanov-type theorems [155, 27, 126], $\overrightarrow{\mathcal{D}_Q}$ correspond to minimum contrast (e.g., maximum likelihood) estimation [50, 51, 71, 8]. In general, the dichotomy between $\overrightarrow{\mathcal{D}_Q}$ and $\overrightarrow{\mathcal{D}_Q}$ can be seen as $D_\phi$-version of a left/right split of a characteristic property $(y - P_Cx, x - P_Cx) \leq 0 \forall (x, y) \in \mathcal{H} \times C$ [10] of metric $(= D_{\phi, \psi})$ projections $P_C$ onto convex closed subsets $C$ in Hilbert space $\mathcal{H}$ under a passage from $\mathcal{H}$ to Banach spaces (left characterising metric projections [63, 153, 128], right characterising completely $\ell$-nonexpansive sunny retractions [39, 146]). This leads us to conjecture that the Legendre transform in Brézisian setting, under a suitable choice of categories (e.g., left and right $D_\phi$-Chebyshëv sets with some additional properties, guaranteeing the composability of respective $D_\phi$-projections), is an adjunction, with the above equivalence as a special case. Could it be approached via a nucleus of profunctor, as in [176]?
8 Epistemic (co)monads and epistemic resource theories

Lawvere [119] proposed to consider deductive theories of mathematical structures as categories, with their models given by functors. If $\mathbf{C}$ and $\mathbf{D}$ are categories, while $F : \mathbf{C} \to \mathbf{D}$ and $G : \mathbf{D} \to \mathbf{C}$ are functors, such that $F \dashv G$, then one can view [120] (c.f. [117, 91]): $\mathbf{C}$ as a category of (type theoretic) axiomatisations, with objects given by logical formulas and morphisms given by proofs (deductions), $\mathbf{D}$ as a category of (geometric) structures modeling these axioms, $F$ as the semantics (meaning) of $\mathbf{C}$ in $\mathbf{D}$, and $G$ as the syntax (formalisation) of $\mathbf{D}$ in $\mathbf{C}$. Interpreting syntax as a minimal axiomatisation, $F$ can be viewed as the most efficient solution to the problem posed by $G$, while $G$ can be seen as posing the most difficult problem that $F$ solves. On the other hand, Lawvere [118], Chencov [49], and Morse and Sacksteder [140] introduced the category of statistical inferences, with sets of probability densities (probabilistic models) as objects and positive norm-preserving linear maps as arrows. Chencov’s approach (viewing the objects as «figures» [48, 50] and their morphisms (statistical decision rules) as «movements» [49, 51], with statistical equivalence understood as inner groupoids) was focused at relationships between categorical and geometric structures of statistical models and inferences. In his view, the choice of a particular class of morphisms requires justification (he referred to Wald’s [172, 173] decision theory), providing a selection of the preferred class of maps with respect to a presumed criteria of optimality (given by the Bayes risk). Parallely, Jaynes [102, 103] stressed that: 1) probabilities are states of knowledge, which is conditioned upon in the criteria of intersubjective experimental reproducibility (thus, not completely subjective/personal); 2) the mathematical structure of a theory of inductive inference should be derived from (determined by) the criteria (requirements) guaranteeing optimality with respect to a particular logic of experimental designs/types of testable data (c.f. [169]): for each specific method of inductive inference, there are different experimental designs that can be optimally analysed with it (e.g., $\chi^2$ test makes no sense for a small sample size, the Bayes–Laplace rule is inapplicable to data given by arithmetic means identifiable with average values, etc).

Our conclusion from these insights, taking into account the large body of evidence on double-sidedness of relationships between ‘experimental facts’ and ‘intersubjective beliefs’ [68, 158, 77], is to: 1) consider pairs of: 1a) inductive inference categories, with geometric structures encoding/determining specific prescriptions of optimal/ideal models and inferences, 1b) experimental design categories, seen as logical (type theoretic), and encoding admissible/ideal types of experimental data and their (experimental) transformations; 2) use adjointness, with syntax given by predictive verification (involving frequentist asymptotics and quantitative control of convergence of algorithmic evaluation) and semantics given by model construction (involving infinitary geometric idealisations of finite data sets).

Any category $\mathbf{C}$ with object $X \in \text{Ob}(\mathbf{C})$ interpreted as a type of knowledge and morphism $f \in \text{Arr}(\mathbf{C})$ interpreted as its transformation will be called an epistemic universe. Consider two epistemic universes: $\text{ExpDes}$ of experimental designs (with objects given, e.g., by the sets of experimental configuration settings, morphisms given by the sets of parameters of the experimental operations that transform between these settings, and composition of morphisms $h = g \circ f$ representing experimental identification of ‘performing operation $h$’ with ‘sequential performing of operations $f$ and $g$’) and $\text{IndInf}$ of theoretical designs (with quantified knowledge/information state spaces as objects and inductive inferences/information processings as morphisms). A functor $I : \text{ExpDes} \to \text{IndInf}$ will be called a model construction (or interpretation) while a functor $P : \text{IndInf} \to \text{ExpDes}$ will be called a predictive verification. In scientific inductive inference, as opposed to mathematical deductive inference, the codomain of semantics is given by the category of inferences (and thus the syntax is provided by predictive verification), so “$\forall$ data $\exists$ inference that models it” (or: “whatever is measurable, it has to be made thinkable”). On the other hand, the formula “$\forall$ inferences $\exists$ data that models it” (or: “whatever is thinkable, it has to be made measurable”) is characteristic to magical thinking. In consequence, a predictive verification $P$ will
be called scientific (resp., magic) iff \( P \) is right (resp., left) adjoint to \( I \). Thus, in scientific inductive inference, \( \text{IndInf} \) plays a role of a geometric category (answering to a question «Who\text{ information?»} [102]), while \( \text{ExpDes} \) plays a role of a type theoretic category (answering to «Information about \text{what?»} [24]). However: 1) given the fixed choice of two categories, there can be various adjoint pairs of functors between them; 2) the experimental design (‘facts’) of some family of agents can be a theoretical design (‘beliefs’) for some other family of agents, and so on. These issues can be (partially) addressed by moving to (co)monads. A choice of a monad (dually: a comonad) on epistemic universe \( C \) determines the class of epistemic universes \( D \) and corresponding adjoint pairs \( I \vdash P \) that make \( C, D \) to be \( \text{ExpDes} \) (resp., \( \text{IndInf} \)) (dually: \( \text{IndInf} \) (resp., \( \text{ExpDes} \))). We will call them epistemic (co)\( \text{monads}. \) \(^{12}\)

Given a choice of a category \( \text{IndInf} \) of inductive inferences, an agent (resp., coagent) is identified with a monad \( J \) (resp., a comonad \( E \)) on \( \text{IndInf} \), encoding the range of available/allowed individual actions/free operations (resp., individually accepted/constructed ‘facts’). A pair \( (E, J) \) of an agent \( J \) and coagent \( E \) on \( \text{IndInf} \) will be called a subject (or a user). We define: an epistemic inference theory as a triple \((\text{IndInf}, E, J)\); a multi-(co)agent epistemic inference theory as \( U := (\text{IndInf}, \{E_i \mid i \in I\}, \{J_j \mid j \in J\}) \) (so, \( U \) becomes multi-user iff there is a fixed bijection \( I \cong J \)). Given a choice of a particular (nonunique) adjoint pair \( I \vdash P \) representing the epistemic comonad \( E \), the epistemic monad \( J = (J, \mu, \eta) \) can be functorially mapped along \( P \), resulting in a monad \( \tilde{J} = (\tilde{J}, \tilde{\mu}, \tilde{\eta}) \) over \( \text{ExpDes} \). Provided there exists a natural transformation \( \alpha : JP \Rightarrow PJ \) such that \( \alpha \circ \eta P = P \eta \) and \( \alpha \circ \mu P = P \mu \circ \alpha \circ \mu J \circ \alpha \circ J \eta \) [161]. In this context, a toy model of a “collective construction of a system of (scientific) facts” (in the sense of [158, 77]) is: given \( U \), the admitted range of possible experimental design categories is limited by the requirement that a single category \( \text{ExpDes} \) has to admit a collection of adjoint pairs \( I_i \vdash P_i \) \( \forall i \in I \), implementing the whole corresponding family \( \{E_i \mid i \in I\} \) of comonads of \( U \). Given subjects \( (E_i, J_i) \) on \( C_i, i \in \{1, 2\} \), and \( p, q \in \{\text{lax}, \text{colax}\} \), we define a \((p, q)\)-strategy a pair \((F_E, \alpha_E, (F_j, \alpha_j))\) of \( p \) morphism \((F_E, \alpha_E) : (C_1, E) \rightarrow (C_2, E)\) and \( q \) morphism \((F_j, \alpha_j) : (C_1, J) \rightarrow (C_2, J)\). Intersubjectivity amounts to relating different subjects in a given theory \( U \). Categorifying Chencov’s groupoids of statistical equivalence, we define intersubjective commensurability of \((\text{lax, lax})\)-strategies as an inner groupoid in 2-category \( \text{InterSubj}_{\text{lax, lax}} \) of subjects of \( U \) as 0-cells, pairs of \((\text{lax, lax})\)-strategies as 1-cells, and pairs of natural transformations \((\kappa_E, \kappa_J) : (F_E, F_j) \rightarrow (\tilde{F}_E, \tilde{F}_j)\), such that \((\tilde{\alpha}_E, \tilde{\alpha}_J) \circ (E_2, J_2)(\kappa_E, \kappa_J) = (\kappa_E, \kappa_J)(E_1, J_1) \circ (\alpha_E, \alpha_J)\) as 2-cells \(((\tilde{F}_E, \tilde{\alpha}_E, (F_j, \alpha_j)) \Rightarrow ((\tilde{G}_E, \tilde{\alpha}_E, (G_j, \alpha_j)))\). The corresponding \((\text{lax}, \text{colax})\)-, \((\text{colax}, \text{lax})\)-, and \((\text{colax}, \text{colax})\)- intersubjective categories and their inner commensurabilities (as well as further special cases, given by specialisation of natural transformations \( \alpha \) to be weak or strong) are defined analogously.

\(^{11}\) «Now, if it comes to making truth, magic can do it far more quickly and brilliantly than science. Magic is an experiment in omnipotence; it thinks to create facts by invoking them, as Absolute Will thinks to create truths by assuming them; so after all we cannot be surprised that Faust finds magic the best key to the universe» [156]. An adjoint triple \( P_m \vdash I \vdash P \), determines a pair \( I \circ P_m \vdash I \circ P \), of monad and comonad on \( \text{IndInf} \) and a dual pair \( P_m \circ I \vdash P_\circ I \circ \text{comonad and monad on ExpDes}, \) allowing for further interpretation along these lines.

\(^{12}\) So, an epistemic comonad on \( C \) limits the possible universes of intersubjective experimental knowledge (together with the corresponding model construction and predictive verification criteria) that are allowed to be built upon \( C \) understood as \( \text{IndInf} \). Dually, an epistemic monad on \( C \) limits the possible theoretical design categories (”optimal models and inferences”), and their relationship with \( C \) understood as \( \text{ExpDes} \). This leads to the concept of epistemic strategies for a given epistemic universe \( C \), understood as either choosing the specifically crafted monad and comonad (if they are not already given) or utilising the range of available adjunctions equivalent to the given monad and comonad. For example, aiming at maximisation of syntactic power of \( C \) as \( \text{ExpDes} \), given a fixed monad on it, one would use the largest possible \( i.e., \) the Eilenberg–Moore category. Dually, aiming at minimisation of semantic power of \( C \) as \( \text{IndInf} \), given a fixed comonad on it, one would use coKleisli category.

\(^{13}\) A map \((P, \alpha)\) is called a lax morphism (and: strict iff \( \alpha \) is an identity; weak iff \( \alpha \) is an isomorphism), while the inversion of direction of \( \alpha \) defines a colax morphism [161, 124]. Lax (resp., colax) morphism induces a functor between corresponding Eilenberg–More (resp., Kleisli) categories, so the choice among them encodes the choice of an epistemic strategy. Dually, given a representation of an epistemic monad, one can subject an epistemic comonad to a (co)lax morphism along this representation, resulting in a “doubly epistemic” comonad, encoding (some information about this) in what sense \( \text{IndInf} \), now viewed as an experimental design category, was a theoretical design for even more deeper layer of experimental design.
Every monad \((T, \mu, \eta)\) on a category \(\mathcal{C}\) gives rise to a monoid \(M_T := (\text{Nat}(\text{id}_\mathcal{C}, T), \mu(\cdot \circ \cdot), \eta)\). Hence, if \(\text{IndInf}\) has a terminal object \(1\), then, given an agent \(J\) on \(\text{IndInf}\), one can consider the objects of \(\text{IndInf}\) as resource spaces, with: free resources given by the objects in \(\{\sigma_1(1) \in \text{Ob}(\text{IndInf}) \mid \sigma \in \text{Nat}(\text{id}_\text{IndInf}, J)\}\), free operations given by \(M_J\), operations given by all natural transformations from \(\text{id}_\text{IndInf}\) to any agent/monad on \(\text{IndInf}\), and resource monotones given by the maps \(r : \text{Ob}(\text{IndInf}) \to [0, \infty]\) such that \(r \circ \sigma_A(A) \leq \sigma_A(A) \forall \sigma \in \text{Nat}(\text{id}_\text{IndInf}, J) \forall A \in \text{Ob}(\text{IndInf})\). Thus, in presence of \(1\) and of at least one nontrivial resource monotone, every (multi-agent) epistemic inference theory becomes a (multi-agent) resource theory. As opposed to set-theoretic case of §3, the collection of all operations may be not a monoid itself (lacking a corresponding agent). Hence, although inspired by [62, 61] and [53], the above setting does not reduce to theirs. On the other hand, the monoidal category \((\mathcal{C}^\text{op}, \circ, \text{id}_\mathcal{C})\) is not symmetric, so the above setting cannot be recast in terms of [56].

\textbf{Ex.1.} From §6 we obtain an epistemic inference theory \((\text{1Cvx}(\ell, \Psi), \text{coFr}_\Psi \circ \text{Fr}_\Psi, \text{Fix}^\ell_{\Psi, \Psi} \circ \text{Fr}_\Psi)\), with \(\Psi\) and \(\varphi\) varying as in §4.Ex.3, is a multi-agent epistemic inference theory. Kaczmarz map \(L_{\gamma_1}(N, \tau) \ni x = u_x|x| \mapsto u_x\gamma_1^{-1}(T_1(|x|)) \in L_{\gamma_2}(N, \tau)\) is a homomorphism [113], setting up categorical equivalences between \(\text{1Cvx}(\ell, \Psi, \varphi)\) for varying \(\Psi\) and fixed \(\varphi\) implying strict intersubjective commensurability of corresponding monads/agents on \(\text{Pow}(N)\). Each agent corresponds to a family of resource theories of states of type \((\text{iii}^\text{fr}_\ell, \Psi, \varphi)\), parametrised by \(\ell\)-closed \(\ell\)-convex sets of free states. On the other hand, \(\text{Pow}(N)\) has a terminal object, allowing to ask: what are the nontrivial resource monotones turning this example in a multi-agent resource theory?

\textbf{Acknowledgments}

I thank Lídia del Rio, Tobias Fritz, Karol Horodecki, and Anna Jenčová for discussions. This research was supported by 2015/18/E/ST2/00327 grant of National Science Centre, as well as by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. Part of this research was conducted during my visit at Department of Mathematical Informatics, Graduate School of Information Science, Nagoya University, on academic leave from University of Gdansk. I wish to express my sincere gratitude to Francesco Buscemi for an invitation, discussions, and kind hospitality.

\footnote{In particular, if \(\text{IndInf}\) is a poset \(P\), understood as a category (as in [130]), then monads \(J\) correspond bijectively to Moore closures on \(P\) (c.f., e.g., [138]), which are not the same as submonoids of endomorphisms End\((P)\). Nevertheless, we have a backwards compatibility with the embeddings of [62, 61]: an interpretation will be called embedding if it is full and faithful (meaning: theory should be capable of interpreting consistently all admitted experimental designs, but not necessarily vice versa). An embedding \(F : \mathcal{C} \to \mathcal{D}\) will be called: extensive if \(F(\mathcal{C})\) is a subcategory of \(\mathcal{D}\); intensive if there exists a functor \(G : \mathcal{D} \to \mathcal{C}\) such that \(F \circ G\) with the unit of adjunction being a natural isomorphism. Hence, an intensive embedding can be seen as a translation from more coarse-grained/concrete to more refined/abstract description, and defines a comonad \(E\) on \(\mathcal{D}\).}

\footnote{Analogous statements hold for corresponding monads/agents on \(\text{Pow}(A)\) (resp., \(\text{Pow}(N)\)), constructed according to §4.Ex.2(a) (resp., §4.Ex.2(b)), via nonassociative (resp., noncommutative) Mazur map \(L_q(A, \tau) \ni x \mapsto \text{sgn}(x)|x|^{p/q} \in L_q(A, \tau)\) (resp., \(L_q(N) \ni x = u_x|x| \mapsto u_x|x|^{p/q} \in L_q(N)\)) as homeomorphism [113] (resp., [111, 145]).}
References

(The following bijective Latin transliteration of Russian Cyrillic script is used: ъ = c, є = ch, Ѳ = zh, Ѡ = sh, ѽ = ь, Ѵ = y, ѵ = i, ѝ = yu, Ѷ = ya, Ѹ = e, ѹ = t, Ѻ = t', Ѵ = õ, Ѳ = l, ѽ = kh, and analogously for capitalised letters, with an exception of X = H at the beginnings of words.)

[1] Abdullaev R.Z., 1983, Neasosociativnye prostranstva $L_p$, Izv. Akad. Nauk UzSSR, Ser. fiz.-mat. 1983;6, 3–5. www.fuw.edu.pl/~kostecki/scans/abdullaev1983.pdf. ↑ 5.

[2] Abdullaev R.Z., 1984, Prostranstva $L_p$ dlya polukonechnykh $JBW$-alebr, Ph.D. Thesis, Institut Matematiki, Akademija Nauk Uzbekskoi SSR, Tashkent, www.fuw.edu.pl/~kostecki/scans/abdullaev1984.pdf. ↑ 5.

[3] Al’ber Ya.I., 1993, Generalized projection operators in Banach spaces: properties and applications, Funct. Diff. Equat. 1, 1–21. arXiv:funct-an/9311002. ↑ 1, 2.

[4] Al’ber Ya.I., 1996, Metric and generalized projection operators in Banach spaces: properties and applications, in: Kartsatos A.G. (ed.), Theory and applications of nonlinear operators of accretive and monotone type, Dekker, New York, pp.15–50. arXiv:funct-an/9311001. ↑ 2.

[5] Al’ber Ya.I., 2007, Young–Fenchel transformation and some new characteristics of Banach spaces, in: Jarosz K. (ed.), Function Spaces: Fifth Conference on Function Spaces, May 16–20, 2006, Southern Illinois University, Edwardsville, Illinois, Contemp. Math. 435, American Mathematical Society, Providence, pp.1–17. ↑ 3.

[6] Al’ber Ya.I., Butnariu D., 1997, Convergence of Bregman projection methods for solving consistent convex feasibility problems in reflexive Banach spaces, J. Optim. Theor. Appl. 92, 33–61. ↑ 1, 3.

[7] Alfsen E.M., Shultz F.W., 2003, Geometry of state spaces of operator algebras, Birkhäuser, Basel. ↑ 5.

[8] Amari S.-i., Nagaoka H., 1993, J¯ oh¯ o kika no h¯ oh¯ o, Iwanami Shoten, Tōkyō (Engl. transl. rev. ed.: 2000, Methods of information geometry, American Mathematical Society, Providence). ↑ 9.

[9] Amemiya I., Andô T., 1965, Convergence of random products of contractions in Hilbert space, Acta Sci. Math. Szeged 26, 239–244. ↑ 6.

[10] Aronszajn N., 1950, Introduction to the theory of Hilbert spaces, The Research Foundation of Oklahoma Agricultural and Mechanical College, Stillwater. ↑ 9.

[11] Asplund E., 1967, Positivity of duality mappings, Bull. Amer. Math. Soc. 73, 200–203. euclid.bams/1183528777. ↑ 2, 5.

[12] Ayupov Sh.A., Abdullaev R.Z., 1989, On isometries of non associative $L_p$ spaces, in: Accardi L., von Waldenfels W. (eds.), Quantum probability and applications IV. Proceedings of the Year of Quantum Probability, held at the University of Rome II, Italy, 1987, LNM 1396, Springer, Berlin, pp.99–106. ↑ 5.

[13] Ayupov Sh.A., Chilin V.I., Abdullaev R.Z., 2012, Orlicz spaces associated with a semi-finite von Neumann algebra, Comment. Math. Univ. Caroliniae 53, 519–533. arXiv:1108.3267. ↑ 6.

[14] Baez J.C., Fritz T., 2014, A bayesian characterization of relative entropy, Theor. Appl. Cat. 29, 421–456. www.tac.mta.ca/tac/volumes/29/16/29-16.pdf. ↑ 9.

[15] Barbu V., Precupanu T., 1978, Convexity and optimisation in Banach spaces, Editura Academiei/Sijthoff & Noordhoff, Bucureşti/Alphen aan den Rijn (4th rev. ed., 2012, Springer, Berlin). ↑ 8.

[16] Barnum H., Duncan R., Wilce A., 2013, Symmetry, compact closure and dagger compactness for categories of convex operational models, J. Phil. Log. 42, 501–523. arXiv:1004.2920. ↑ 8.

[17] Bauschke H.H., Borwein J.M., 1997, Legendre functions and the method of random Bregman projections, J. Conv. Anal. 4, 27–67. people.ok.ubc.ca/bauschke/Research/07.pdf. ↑ 6.

[18] Bauschke H.H., Borwein J.M., Combettes P.L., 2001, Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, Commun. Contemp. Math. 3, 615–647. people.ok.ubc.ca/bauschke/Research/18.pdf. ↑ 1, 3, 5, 8.
[19] Bauschke H.H., Borwein J.M., Combettes P.L., 2003, *Bregman monotone optimization algorithms*, Soc. Industr. Appl. Math. J. Contr. Optim. 42, 596–636. people.ok.ubc.ca/bauschke/Research/28.pdf. ↑ 1, 2, 3.

[20] Bauschke H.H., Combettes P.L., 2003, *Construction of best Bregman approximations in reflexive Banach spaces*, Proc. Amer. Math. Soc. 131, 3757–3766. people.ok.ubc.ca/bauschke/Research/27.pdf. ↑ 6.

[21] Bauschke H.H., Lewis A.S., 2000, *Dykstra's algorithm with Bregman projections: a convergence proof*, Optimization 48, 409–427. people.ok.ubc.ca/bauschke/Research/14.pdf. ↑ 6.

[22] Bauschke H.H., Macklem M.S., Wang X., 2011, *Chebyshev sets, Klee Sets, and Chebyshev centers with respect to Bregman distances: recent results and open problems*, in: Bauschke H.H., Burachik R.S., Combettes P.L., Elser V., Luke D.R., Wolkowicz H. (eds.), *Fixed-point algorithms for inverse problems in science and engineering*, Springer, Berlin, pp.1–21. arXiv:1003.3127. ↑ 3, 9.

[23] Bauschke H.H., Wang X., Ye J., Yuan X., 2009, *Bregman distances and Chebyshev sets*, J. Approx. Theory 159, 3–25. arXiv:0712.4030. ↑ 3.

[24] Bell J.S., 1990, *Against ‘measurement’*, Phys. World 3, 33–40 (also in: Miller A.I. (ed.), *Sixty-two years of uncertainty: historical, philosophical, and physical inquiries into the foundations of quantum mechanics*, Springer, Berlin, pp.17–31). www.tau.ac.il/~quantum/Vaidman/IQM/BellAM.pdf. ↑ 11.

[25] Beurling A., Livingston A.E., 1962, *A theorem on duality mappings in Banach spaces*, Ark. Mat. 4, 405–411. euclid:afm/1485893389. ↑ 2, 5.

[26] Bishop E., Phelps R.R., 1963, *The support functionals of a convex set*, in: Klee V., *Convexity*, American Mathematical Society, Providence, pp.27–35. ↑ 7.

[27] Bjelaković I., Deuschel J.-D., Krüger T., Seiler R., Siegmund-Schultze R., Szkoła A., 2005, *A quantum version of Sanov's theorem*, Commun. Math. Phys. 260, 659–671. arXiv:quant-ph/0412157. ↑ 9.

[28] Blackwell D.A., 1951, *Comparison of experiments*, in: Neyman J. (ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley, pp.93–102. euclid:bsmsp/1200500222. ↑ 7.

[29] Borwein J.M., Read J., Lewis A.S., Zhu Q.J., 1999, *Convex spectral functions of compact operators*, Int. J. Nonlin. Conv. Anal. 1, 17–35. homepages.wmich.edu/~zhu/papers/brlz.ps. ↑ 6.

[30] Borwein J.M., Vanderwerff J.D., 2004, *Constructible convex sets*, Set-Val. Anal. 12, 61–77. wayback.cecm.sfu.ca/Preprints03/2003-202.pdf. ↑ 7.

[31] Borwein J.M., Vanderwerff J.D., 2010, *Convex functions: constructions, characterizations and counterexamples*, Cambridge University Press, Cambridge. ↑ 2.

[32] Boyle P., Dykstra R.L., 1986, *A method for finding projections onto the intersection of convex sets in Hilbert spaces*, in: Dykstra R.L., Robertson T., Wright F.T. (eds.), *Advances in order restricted statistical inference. Proceedings of the Symposium on order restricted statistical inference held in Iowa City, Iowa, September 11–13, 1985*, LNS 37, Springer, Berlin, pp.28–47. ↑ 6.

[33] Bratteli O., Robinson D.W., 1979, 1981, *Operator algebras and quantum statistical mechanics*, Vol.1-2, Springer, Berlin (2nd rev. ed., 1987, 1997). ↑ 2.

[34] Brégman L.M., 1965, *Nakhozhdenie obšee tochki vypuklykh mnozhestv metodom posledovatel'no go proektirovaniya*, Dokl. Akad. Nauk SSSR 162, 487–490. mathnet.ru:dan31130 (Engl. transl. 1965, *The method of successive projection for finding a common point of convex sets*, Soviet Math. Dokl. 6, 688–692). ↑ 6.

[35] Bréguon L.M., 1967, *Relaksacionnyi metod nakhozhdeniya obše ı tochki vypuklykh mnozhestv i ego primenenie dlya resheniya zadach vypuklogo programmirovaniya*, Zh. vychestel. matem. matem. fiz. 7, 620–631. mathnet.ru:zvmmf7353 (Engl. transl.: 1967, *The relaxation method for finding common points of convex sets and its application to the solution of problems in convex
Nonexpansive projections and resolvents of accretive operators
[41] Bruck R.E. Jr., Reich S., 1977, A mathematical theory of resources
[56] Coecke B., Fritz T., Spekkens R.W., 2016, Quantum resource theories
[55] Cigler J., Losert V., Michor P.W., 1979, Banach modules and functors on categories of Banach spaces
[54] Chitambar E., Gour G., 2019, Agents, subsystems, and the conservation of information
[53] Chiribella G., 2018, Minimizing integrals in certain classes of monotone programming, USSR Comput. Math. Math. Phys. 7, 200–217).
[36] Brègman L.M., Censor Y., Reich S., 1999, Dykstra’s algorithm as the nonlinear extension of Bregman’s optimization method, J. Convex Anal. 6, 319–333. www.emis.de/journals/JCA/vol.6_no.2/j184.pdf. ↑ 6.
[37] Browder F.E., 1958, On some approximation methods for solutions of the Dirichlet problem for linear elliptic equations of arbitrary order, J. Math. Mech. 7, 69–80. ↑ 6.
[38] Browder F.E., 1966, Fixed point theorems for nonlinear semicontractive mappings in Banach spaces, Arch. Rat. Mech. Anal. 21, 259–269. ↑ 5.
[39] Bruck R.E. Jr., 1973, Nonexpansive projections on subsets of Banach spaces, Pacific J. Math. 47, 341–355. euclid:pjm/1102945870. ↑ 9.
[40] Bruck R.E. Jr., 1982, Random products of contractions in metric and Banach spaces, J. Math. Anal. Appl. 88, 319–332. ↑ 6.
[41] Bruck R.E. Jr., Reich S., 1977, Nonexpansive projections and resolvents of accretive operators in Banach spaces, Houston J. Math. 3, 459–470. ↑ 6.
[42] Brunk H.D., Ewing G.M., Utz W.R., 1957, Minimizing integrals in certain classes of monotone functions, Pacific J. Math. 7, 833–847. euclid:pjm/1103043663. ↑ 3.
[43] Buscemi F., Kostecki R.P., 2021, in preparation. ↑ 5.
[44] Butnariu D., Iusem A.N., 2000, Finite dimensional optimization, Kluwer, Dordrecht. ↑ 1, 3, 5.
[45] Butnariu D., Iusem A.N., 1997, Local moduli of convexity and their application to finding almost common fixed points of measurable families of operators, in: Censor Y., Reich S. (eds.), Recent developments in optimization theory and nonlinear analysis. AMS/IMU special session on optimization and nonlinear analysis, May 24–26, 1996, Jerusalem, Israel, Contemp. Math. 204, American Mathematical Society, Providence, pp.61–91. ↑ 1, 3.
arXiv:1409.5531. ↑ 12.

250, 59–86. arXiv:1409.5531. ↑ 12.

[57] Combettes P.L., 1993, *Signal recovery by best feasible approximation*, IEEE Trans. Image Process. 2, 269–271. ↑ 6.

[58] Csiszár I., 1995, *Generalized projections for non-negative functions*, Acta Math. Hung. 68, 161–185. ↑ 2.

[59] Davies E.B., 1974, *Symmetries of compact convex sets*, Quart. J. Math. Oxford 25, 323–328. ↑ 2.

[60] Davies E.B., Lewis J.T., 1970, *An operational approach to quantum probability*, Commun. Math. Phys. 17, 239–260. euclid:cmp/1103842336. ↑ 2, 8.

[61] del Rio L., Krämer L., 2017, *Operational locality in global theories*, Phil. Trans. Roy. Soc. A 376, 20170321. arXiv:1701.03280. ↑ 12.

[62] del Rio L., Krämer L., Renner R., 2015, *Resource theories of knowledge*, arXiv:1511.08818. ↑ 2, 12.

[63] Deutsch F., 1965, *Some applications of functional analysis to approximation theory*, Ph.D. Thesis, Brown University, Providence. ↑ 9.

[64] Devetak I., Harrow A.W., Winter A.J., 2008, *A resource framework for quantum Shannon theory*, IEEE Trans. Inform. Theor. 54, 4587–4618. arXiv:quant-ph/0512015. ↑ 9.

[65] Dhillon I.S., Tropp J.A., 2007, *Matrix nearness problems with Bregman divergences*, Soc. Industr. Appl. Math. J. Matrix Anal. Appl. 29, 1120–1146. authors.library.caltech.edu/9428/1/DHIsiamjmaa07.pdf. ↑ 6.

[66] Domotor Z., 1985, *Probability kinematics, conditionals, and entropy principles*, Synthese 63, 75–114. ↑ 2.

[67] Dykstra R.L., 1983, *An algorithm for restricted least squares regression*, J. Amer. Statist. Assoc. 78, 837–842. cda.psych.uiuc.edu/matlab_programming_class_2012/dykstra.pdf. ↑ 6.

[68] Edwards D.A., 1964, *On the homeomorphic affine embedding of a locally compact cone into a Banach dual space endowed with the vague topology*, Proc. London Math. Soc. Ser. 3 14, 399–414. ↑ 8.

[69] Eguchi S., 1983, *Second order efficiency of minimum contrast estimators in a curved exponential family*, Ann. Statist. 11, 793–803. euclid:aos/1176346246. ↑ 2, 9.

[70] Elsasser W.M., 1937, *On quantum measurements and the role of the uncertainty relations in statistical mechanics*, Phys. Rev. 52, 987–999. ↑ 1.

[71] Fakhruddin S.M., 1986, *Absolute flatness and amalgams in pomonoids*, Semigr. Forum 33, 15–22. www.digizeitschriften.de/dms/img/?PID=GDZPPN00125796X. ↑ 9.

[72] Falcone A.J., Takesaki M., 2001, *The non-commutative flow of weights on a von Neumann algebra*, J. Funct. Anal. 182, 170–206. www.math.ucla.edu/~mt/papers/QFlow-Final.tex.pdf. ↑ 5.

[73] Fischer H.R., Rüttiman G.T., 1978, *The geometry of the state space*, in: Marlow A.R. (ed.), *Mathematical foundations of quantum theory*, Academic Press, New York, pp.153–176. ↑ 5.

[74] Fleck L., 1935, *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*, Schwabe, Basel. (Engl. transl. 1979, *Genesis and development of the scientific fact*, Chicago University Press, Chicago). ↑ 10, 11.

[75] Fritz T., 2017, *Resource convertibility and ordered commutative monoids*, Math. Struct. Comp. Sci. 27, 850–938. arXiv:1504.03661. ↑ 9.

[76] Gagné N., Panangaden P., 2018, *A categorical characterization of relative entropy on standard
Borel spaces, Electr. Not. Theor. Comp. Sci. A 336, 135–153. arXiv:1703.08853. ↑ 9.

[80] Gibilisco P., Isola T., 1999, Connections on statistical manifolds of density operators by geometry of non-commutative $L^p$ spaces, Infin. Dim. Anal. Quant. Prob. Relat. Top. 2, 169–178. www.mat.uniroma2.it/~isola/research/preprints/Gils01.pdf. ↑ 2.

[81] Gibilisco P., Pistone G., 1998, Connections on non-parametric statistical manifolds by Orlicz space geometry, Infin. Dim. Anal. Quant. Prob. Relat. Top. 1, 325–347. art.torvergata.it/retrieve/handle/2108/49737/18230/IDAQP1998.pdf. ↑ 2.

[82] Giles R., 1970, Foundations of quantum mechanics, J. Math. Phys. 11, 2139–2160. ↑ 2.

[83] Gudder S.P., 1973, Convex structures and operational quantum mechanics, Commun. Math. Phys. 29, 249–264. euclid:cmp/1103858551. ↑ 2, 8.

[84] Gurin L.G., Polyak B.T., Raïk È.V., 1967, The product of projection operators, Zh. vychisl. matem. matem. fiz. 7, 1211–1228. matnet.ru:ppi892. ↑ 8.

[85] Halperin I., 1962, Information-theoretical aspects of quantum measurement, USSR Comput. Math. Math. Phys. 676, 1–24. ↑ 6.

[86] Han S.-P., 1988, A successive projection method, J. Math. Anal. Appl. 131, 383–391. ↑ 1.

[87] Hasegawa H., 1993, Information without probability, Math. Program. 77, 335–355. ↑ 6.

[88] Haugazeau Y., 1968, Sur les inéquations variationnelles et la minimisation de fonctionnelles convexes, Ph.D. Thesis, Université de Paris, Paris. ↑ 6.

[89] Hellmann F., Kamiński W., Kostecki R.P., 2016, Analog teorii statisticheskikh решений в некоммутативной теории вероятностей, Труды Московского математического общества за год 1972. Volume 26, American Mathematical Society, Providence, pp.133–149). ↑ 4.

[90] Hofmann M., 1995, On the interpretation of type theory in locally cartesian closed categories, in: Pacholski L., Tiuryn J. (eds.), Computer science logic: 8th workshop, CSL’94. Kazimierz, in: Pacholski L., Tiuryn J. (eds.), Springer, Berlin, pp.427–441. ↑ 6.

[91] Hofmann M., 1995, On the interpretation of type theory in locally cartesian closed categories, in: Pacholski L., Tiuryn J. (eds.), Computer science logic: 8th workshop, CSL’94. Kazimierz, in: Pacholski L., Tiuryn J. (eds.), Springer, Berlin, pp.427–441. ↑ 6.

[92] Hofmann M., 1995, On the interpretation of type theory in locally cartesian closed categories, in: Pacholski L., Tiuryn J. (eds.), Computer science logic: 8th workshop, CSL’94. Kazimierz, in: Pacholski L., Tiuryn J. (eds.), Springer, Berlin, pp.427–441. ↑ 6.

[93] Hasegawa H., 1993, Information without probability, Math. Program. 77, 335–355. ↑ 6.

[94] Iochum B., 1986, Cônes autopolaires et algèbres de Jordan, Springer, Berlin. ↑ 5.

[95] Iochum B., 1984, Cônes autopolaires et algèbres de Jordan, Springer, Berlin. ↑ 5.

[96] Iochum B., 1984, Cônes autopolaires et algèbres de Jordan, Springer, Berlin. ↑ 5.

[97] Ingarden R.S., Urbanik K., 1962, Information theory and statistical mechanics. II, Foundations of quantum mechanics, Colloq. Math. 9, 131–150. matwbn.icm.edu.pl/ksiazki/cm/cm9/cm9121.pdf. ↑ 2.

[98] Iochum B., 1984, Cônes autopolaires et algèbres de Jordan, Springer, Berlin. ↑ 5.

[99] Iochum B., 1984, Cônes autopolaires et algèbres de Jordan, Springer, Berlin. ↑ 5.

[100] Jaynes E.T., 1957, Information theory and statistical mechanics, Phys. Rev. 106, 620–630. bayes.wustl.edu/etj/articles/theory.1.pdf. ↑ 1.

[101] Jaynes E.T., 1957, Information theory and statistical mechanics. II, Phys. Rev. 108, 171–190. bayes.wustl.edu/etj/articles/theory.2.pdf. ↑ 1.

[102] Jaynes E.T., 1979, Where do we stand on maximum entropy?, in: Levine R.D., Tribus M. (eds.), The maximum entropy formalism: a conference held at the Mas-
Quantum information geometry and non-commutative $L_p$ spaces, Inf. Dim. Anal. Quant. Prob. Relat. Top. 8, 215–233. www.mat.savba.sk/~jencova/pdf/lpspaces.pdf. ↑ 2, 5.

Kaczmarz S., 1933, O homeomorfii pewnych przestrzeni. — The homeomorphy of certain spaces, Bull. Internat. Acad. Polon. Sci. Lett., Class. Sci. Math. Natur.: Sér. A, Sci. Math. 1933:2, 145–148. www.fuw.edu.pl/~kostecki/scans/kaczmarz1933.pdf. ↑ 2.

Kaczmarz S., 1937, Przybliżone rozwiązywanie układów równań liniowych. — Approximated Solution of Systems of Linear Equations, Int. J. Contr. 57, 1269–1271. ↑ 6.

Jaynes E.T., 2003, Probability theory: the logic of science, Cambridge University Press, Cambridge. ↑ 1, 7, 10, 11.

Kaczm, S., 1933, O homeomorfii pewnych przestrzeni. — The homeomorphy of certain spaces, Bull. Internat. Acad. Polon. Sci. Lett., Class. Sci. Math. Natur.: Sér. A, Sci. Math. 1933:2, 145–148. www.fuw.edu.pl/~kostecki/scans/kaczmarz1933.pdf. ↑ 2.

Kaczm, S., 1937, Przybliżone rozwiązywanie układów równań liniowych. — Approximate solution of systems of linear equations, Int. J. Contr. 57, 1269–1271. ↑ 6.

Kadison R.V., 1951, A representation theory for commutative topological algebra, Mem. Amer. Math. Soc. 7, American Mathematical Society, Providence. ↑ 8.

Kakutani S., 1940, Nakano shi no danwa nitsuite, Zenkoku Shijō Sūgaku Danwakai 192, 42–44. www.fuw.edu.pl/~kostecki/scans/kakutani1940.pdf. ↑ 6.

Kuwiel K.C., 1997, Proximal minimization methods with generalized Bregman functions, Soc. Industr. Appl. Math. J. Contr. Optim. 35, 1142–1168. ↑ 3.

Kosaki H., 1984, Applications of the complex interpolation method to a von Neumann algebra: non-commutative $L^p$-spaces, J. Funct. Anal. 56, 29–78. ↑ 5.

Kosaki H., 1984, Applications of uniform convexity of noncommutative $L^p$-spaces, Trans. Amer. Math. Soc. 283, 265–282. www.ams.org/journals/tran/1984-283-01/S0002-9947-1984-0735421-6. ↑ 2, 12.

Kostecki R.P., 2011, The general form of $\gamma$-family of quantum relative entropies, Open Sys. Inf. Dyn. 18, 191–221. arXiv:1106.2225. ↑ 2, 5.

Kostecki R.P., 2017, Postquantum Bregman relative entropies, arXiv:1710.01837 (version 3 in preparation). ↑ 2, 3, 5, 6, 12.

Kostecki R.P., Munk M.I.K., 2021, Conditional expectations as right entropic projections, in preparation. ↑ 1, 4, 6.

Kullback S., 1959, Information theory and statistics, Wiley, New York (2nd rev. ed. 1968, Dover, New York). ↑ 1.

Kunze W., 1990, Noncommutative Orlicz spaces and generalized Arens algebras, Math. Nachr. 147, 123–138. ↑ 6.

Lambek J., Scott P.J., 1986, Introduction to higher order categorical logic, Cambridge Studies in Advanced Mathematics 3, Cambridge University Press, Cambridge. ↑ 10.

Lawvere F.W., 1962, The category of probabilistic mappings with applications to stochastic processes, statistics, and pattern recognition, unpublished preprint. www.fuw.edu.pl/~kostecki/scans/lawvere1962.pdf. ↑ 10.

Lawvere F.W., 1963, Functorial semantics of algebraic theories, Ph.D. thesis, Columbia University, New York (reprint in: Lawvere F.W., 2004, Functorial semantics of algebraic theories and Some algebraic problems in the context of functorial semantics of algebraic theories, Repr. Theor. Appl. Cat. 5, 1–121. www.tac.mta.ca/tac/reprints/articles/5/tr5.pdf). ↑ 2, 10.

Lawvere F.W., 1969, Adjointness in foundations, Dialectica 23, 281–296 (reprint 2004, Repr. Theor. Appl. Cat. 16, 1–16. www.tac.mta.ca/tac/reprints/articles/16/tr16a.pdf). ↑ 10.

Lawvere F.W., 1973, Metric spaces, generalized logic, and closed categories, Rend. Sem. Mat. Fis. Milano 43, 135–166 (reprint 2002, Repr. Theor. Appl. Cat. 1, 1–37. www.tac.mta.ca/tac/reprints/articles/1/tr1.pdf). ↑ 9.
[144] Raginsky M., 2011, Shannon meets Blackwell and Le Cam: channels, codes, and statistical experiments, in: Kuleshov A.P, Blinovskii V.M., Ephremides A. (eds.), 2011 IEEE International Symposium on Information Theory Proceedings (ISIT 2011), IEEE, Piscataway, pp.1220–1224. maxim.ece.illinois.edu/pubs/raginsky_ISIT11.pdf. ↑ 7.

[145] Raynaud Y., 2002, On ultrapowers of non commutative $L_p$ spaces, J. Oper. Th. 48, 41–68. ↑ 2, 12.

[146] Reich S., 1973, Asymptotic behavior of contractions in Banach spaces, J. Math. Anal. Appl. 44, 57–70. ↑ 9.

[147] Reich S., 1996, A weak convergence theorem for the alternating method with Bregman distances, in: Kartsatos A.G. (ed.), Theory and applications of nonlinear operators of accretive and monotone type, Dekker, New York, pp.313–318. ↑ 1, 3.

[148] Reich S., Sabach S., 2011, Existence and approximation of fixed points of Bregman firmly nonexpansive mappings in reflexive Banach spaces, in: Bauschke H.H., Burachik R.S., Combettes P.L., Elser V., Luke D.R., Wolkowicz H. (eds.), Fixed-point algorithms for inverse problems in science and engineering, Springer, Berlin, pp.299–314. ssbach.net.technion.ac.il/files/2015/12/RS2010-3.pdf. ↑ 2, 3.

[149] Rényi A., 1961, On measures of entropy and information, in: Neyman J. (ed.), Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, Vol.1, University of California Press, Berkeley, pp.547–561. euclid:bsmsp/1200512181. ↑ 2.

[150] Resmerita E., 2004, On total convexity, Bregman projections and stability in Banach spaces, J. Math. Anal. Appl. 294, 251–263. ↑ 1, 5.

[151] Rockafellar R.T., 1967, Conjugates and Legendre transforms of convex functions, Canad. J. Math. 19, 200–205. sites.math.washington.edu/~rtr/papers/rtr014-LegendreTransform.pdf. ↑ 3.

[152] Rockafellar R.T., 1993, Lagrange multipliers and optimality, Soc. Indust. Appl. Math. Rev. 35, 183–238. pages.cs.wisc.edu/~ferris/cs730/sirev35_183.pdf. ↑ 2.

[153] Rubinshteĭn G.Sh., 1965, Ob odnoi ēkstremal’noi zadache v lineĭnom normirovannom prostranstve, Sibir. mat. zhurn. 6, 711–714. mathnet.ru:smj5169. ↑ 9.

[154] Sabach S., 2012, Iterative methods for solving optimization problems, Ph.D. Thesis, Technion – Israel Institute of Technology, Haifa. ssbach.net.technion.ac.il/files/2015/12/PhD-Dissertation.pdf. ↑ 2, 3, 4.

[155] Sanov I.N., 1957, O veroyatnosti bol’šikh otklonenii sluchainykh velichin, Matem. Sb. N.S. 42, 11–42. mathnet.ru:msb5043 (Engl. transl. 1961, On the probability of large deviations of random variables, Sel. Transl. Math. Statist. Probab. 1, 213–244). ↑ 1, 9.

[156] Santayana G., 1915, Goethe and German egotism, New Republ. 1, 15–16. newrepublic.com/article/114485/george-santayana-goethe-and-german-egotism-january-2-1915. ↑ 11.

[157] Shvarc A.S., 1963, Funktory v kategoriakh banakhovykh prostranstv, Dokl. Akad. Nauk SSSR 149, 44–47. mathnet.ru:dan27662. ↑ 7.

[158] Spengler O., 1918, 1923, Der Untergang des Abendlandes. Umriffe einer Morphologie der Weltgeschichte, Vol.1-2, Braumüller, Wien (2nd rev. ed. of Vol.1: 1922, Beck, München; Engl. transl.: 1926, The decline of the west, Knopf, New York). ↑ 10, 11.

[159] Stormer E., 1966, Jordan algebras of type I, Acta Math. 115, 165–184. ↑ 5.

[160] Stratonovich R.L., 1955, Entropiya v kvantovoi statistike, Zh. Éksp. Teor. Fiz. 28, 547–558 (Engl. transl. 1955, Entropy in quantum statistics, Soviet Phys. J. Exp. Theor. Phys. 1, 426–434. jetp.ac.ru/cgi-bin/dn/e_001_03_0426.pdf). ↑ 1.

[161] Street R., 1972, The formal theory of monads, J. Pure Appl. Alg. 2, 149–168. ↑ 11.

[162] Sýkora S., 1974, Quantum theory and the bayesian inference problems, J. Stat. Phys. 11, 17–27. www.ebyte.it/stan/1974_Sykora_QuantumBayesianInference.pdf. ↑ 2.

[163] Takagi R., Regula B., 2019, General resource theories in quantum mechanics and beyond: op-
erational characterization via discrimination tasks, Phys. Rev. X 9, 031053. arXiv:1901.08127.

[164] Terp M., 1981, \(L^p\)-spaces associated with von Neumann algebras, Københavns Univ. Math. Inst. Rapp. No. 3a+3b, Matematisk Institut, Københavns Universitet, København. www.fuw.edu.pl/~kostecki/scans/terp1981.pdf. ↑ 5.

[165] Tikhonov O.E., 1990, Banakhovy prostranstva, asociirovannye s prostranstvom sostoyanii, i funkciya informacii, Konstr. teor. funk. funk. anal. 7, 67–90. mathnet.ru:kuktf80. ↑ 5.

[166] Tikhonov O.E., 1992, Spektral’naya teoriya dlya prostranstv s bazovoi normoi, Konstr. teor. funk. funk. anal. 8, 76–91. mathnet.ru:kuktf90. ↑ 5.

[167] Tikhonov O.E., 1993, Trace inequalities for spaces in spectral duality, Studia Math. 104, 99–110. matwbn.icm.edu.pl/ksiazki/sm/sm104/sm10416.pdf. ↑ 5.

[168] Topping D.M., 1965, Jordan algebras of self-adjoint operators, Mem. Amer. Math. Soc. 53, American Mathematical Society, Providence. ↑ 5.

[169] Tribus M., 1969, Rational descriptions, decisions and designs, Pergamon, New York. ↑ 10.

[170] Umegaki H., 1961, On information in operator algebras, Proc. Jap. Acad. 37, 459–461. euclid:pja/1195523632. ↑ 2, 6.

[171] von Neumann J., 1933, Functional operators. Volume II: The geometry of orthogonal spaces, mimeographed lecture notes, Princeton University, Princeton (repr. 1950, Princeton University Press, Princeton). ↑ 6.

[172] Wald A., 1939, Contributions to the theory of statistical estimation and testing hypothesis, Ann. Math. Statist. 10, 299–326. ↑ 4, 10.

[173] Wald A., 1950, Statistical decision functions, Wiley, New York. ↑ 4, 10.

[174] Warmuth M.K., 2005, A Bayes rule for density matrices, in: Weiss Y., Schölkopf B., Platt J. (eds.), Advances in neural information processing systems 18, MIT Press, pp.1457–1464. papers.nips.cc/paper/2005/file/4191f5f6c1576762869ac49281130c9-Paper.pdf. ↑ 1.

[175] Wiener N., 1948, Cybernetics or control and communication in the animal and the machine, MIT Press, Cambridge (2nd rev. ed. 1961). ↑ 2.

[176] Willerton S., 2015, The Legendre-Fenchel transform from a category theoretic perspective, arXiv:1501.03791. ↑ 9.

[177] Williams P.M., 1980, Bayesian conditionalisation and the principle of minimum information, Brit. J. Phil. Sci. 31, 131–144. ↑ 1.

[178] Xu H.-K., Kim T.-H., Yin X.M., 2014, Weak continuity of the normalized duality map, J. Nonlin. Conv. Anal. 15, 595–604. ↑ 5.

[179] Zhu H., Rohwer R., 1997, Measurements of generalisation based on information geometry, in: Ellacott S.W., Mason J.C., Anderson I.J. (eds.), Mathematics of neural networks: models, algorithms and applications, Kluwer, Dordrecht, pp.394–398. eprints.aston.ac.uk/514/1/NCRG_95_012.pdf. ↑ 2.

[180] Zălinescu C., 1983, On uniformly convex functions, J. Math. Anal. Appl. 95, 344–374. ↑ 5.

[181] Zălinescu C., 2002, Convex analysis in general vector spaces, World Scientific, Singapore. ↑ 5.