

Quantum theory as inductive inference¹

Ryszard Paweł Kostecki

Institute of Theoretical Physics, University of Warsaw, Hoża 69, 00-681 Warszawa, Poland

September 11, 2011

Abstract

We present the elements of a new approach to the foundations of quantum theory and probability theory which is based on the algebraic approach to integration, information geometry, and maximum relative entropy methods. It enables us to deal with conceptual and mathematical problems of quantum theory without any appeal to frameworks of Hilbert spaces and measure spaces.

1 Introduction

What is the relationship between information theory and quantum theory? The notion of information, quantified by entropy, is usually defined using the notion of probability [112]. But Ingarden and Urbanik [65, 66] showed that the notion of information is independent of the notion of probability, and that the latter can be defined by the former. Moreover, the mathematical setting of probability theory can be considered as a special case of the mathematical setting of quantum theory, with Hilbert spaces [50, 104, 108, 120, 106] or W^* -algebras [92, 57, 109] replacing the role played by the measure spaces $(\mathcal{X}, \mathcal{U}(\mathcal{X}), \mu)$ in the Borel–Kolmogorov approach. Hence, in order to consider information theory as more fundamental than quantum theory, it is necessary to develop it independently of the notion of probability, and to equip it with the dynamical structures sufficient to recover dynamics of quantum theoretic models.

In this paper we will present the elements of a new mathematical formulation of information kinematics and information dynamics, equipped with a new interpretation, which are aimed to solve the above task. We begin with replacing the ordinary Borel–Kolmogorov setting based on probability measures μ on commutative countably additive algebras $\mathcal{U}(\mathcal{X})$ of subsets of a given set \mathcal{X} by *information kinematics* based on spaces of finite positive integrals on abstract (but *integrable*) commutative or non-commutative algebras. This change has two reasons:

- (1) the conflict between the Bayes–Laplace [12, 37] and the Borel–Kolmogorov [13, 84] approaches to mathematical foundations of the probability theory can be resolved in terms of the Le Cam–Whittle [89, 90, 117, 118] approach, which is based on the Daniell–Stone [32, 33, 34, 114, 115] theory of integration on vector lattices, using the canonical association of a vector lattice to a given boolean algebra;
- (2) the algebraic approach to mathematical foundations of quantum theory [110, 111, 55], provided in terms of positive linear functionals on non-commutative C^* -algebras avoids several important problems of the Hilbert space based approach, including

¹Published in: Mohammad-Djafari A., Bercher J., Bessière P. (eds.), *Proceedings of the 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, AIP Conf. Proc. **1305** (2011), Springer, Berlin, pp. 24-35. The published version was abridged due to strict 8 pages article size limit. This text is an improved version of an original unabridged paper.

the problem of unitary inequivalence of different Hilbert space representations of a single non-commutative algebra in infinite-dimensional cases, which plays a crucial role in construction of mathematically strict models in relativistic quantum field theory and continuous quantum statistical mechanics [41, 14, 55, 10, 7].

In order to provide a refined analysis and characterisation of the information kinematic spaces of integrals, we equip them with the non-symmetric information distance function, known as *information deviation* (negative *relative entropy*). Relative entropy defines the *information geometry* of information kinematic spaces in terms of such objects as riemannian metrics, affine connections, and non-linear projections (on the convex subsets), which can be naturally derived from relative entropy under some mild conditions. We discuss the characterisation of a particular family (D_p) of relative entropy functionals. As a consequence of this characterisation, some preferred class of information geometries is selected. We introduce the *information dynamics* as a mapping on the space of finite integrals provided by the constrained maximisation of relative entropy.

When applied to the abstract commutative boolean algebras, the above setting generalises the standard approach to probability theory *and* statistical inference. When applied to the abstract non-commutative W^* -algebras, the above setting provides new mathematical framework for kinematics and dynamics of quantum theory. In both (commutative and non-commutative) regimes we are able to recover the standard approaches to foundations as special cases of our approach.

The new kinematical setting of quantum theory is defined as follows. Given W^* -algebra \mathcal{N} , the *quantum information model* $\mathcal{M}(\mathcal{N})$ is defined as (some) subset of the space \mathcal{N}_*^+ of normal positive finite linear functionals on \mathcal{N} . The space $\mathcal{M}(\mathcal{N})$, together with its non-linear quantum information geometry, is considered as a replacement of the linear Hilbert space \mathcal{H} in the role of the kinematic setting of a quantum theoretical model. The elements of $\mathcal{M}(\mathcal{N})$ will be called *quantum information states*. The family D_p of quantum relative entropies is closely related with non-commutative $L_p(\mathcal{N})$ spaces (because D_p are quantum Bregman entropies), so the preferred geometry of $\mathcal{M}(\mathcal{N})$ induced by D_p can be analysed by quantum information geometric representations of abstract $\mathcal{M}(\mathcal{N})$ spaces in non-commutative $L_p(\mathcal{N})$ spaces. This replaces the use of commutative $L_2(\mathcal{X}, \mathcal{U}(\mathcal{X}), \mu)$ spaces which represent the abstract Hilbert spaces \mathcal{H} . Among $L_p(\mathcal{N})$ spaces, of particular importance is the $L_2(\mathcal{N})$ space, which is unitary isomorphic to the Hilbert space \mathcal{H}_H of Haagerup's standard representation of \mathcal{N} , and unitary isomorphic to the Hilbert space \mathcal{H}_ω of the Gel'fand–Naimark–Segal representation associated with any faithful $\omega \in \mathcal{N}_*^+$. The quantum information geometric representations $\ell_{1/2} : \mathcal{M}(\mathcal{N}) \rightarrow L_2(\mathcal{N})$ serve as a key feature allowing the reconstruction of the ordinary Hilbert space based setting of quantum theory as a special case of our approach.

The new dynamical setting of quantum theory is defined as follows. The description of temporal behaviour and—more generally—information dynamics of quantum models in terms of unitary or completely positive mappings of operators over \mathcal{H} is replaced by the mapping on $\mathcal{M}(\mathcal{N})$ generated by the constrained maximisation of quantum relative entropy D_p . This map can be understood in terms of information geometry as a non-linear projection. Its well-defined constraints (guaranteeing the existence and uniqueness of the solution of the variational problem) are specified by non-empty closed convex subsets of non-commutative $L_p(\mathcal{N})$ spaces. These subsets define the domain of the entropic projection. If the temporal dependence of the constraints is reflected in the temporal dependence of the projection, then this projection generates a temporal trajectory on

$\mathcal{M}(\mathcal{N})$, representing the *temporal evolution* of quantum information states.

Defining quantum kinematics in terms of spaces $\mathcal{M}(\mathcal{N})$ of integrals over abstract (qualitative) non-commutative W^* -algebras \mathcal{N} instead of dealing with Hilbert space representation-dependent density operators and von Neumann algebra representations of abstract W^* -algebras provides novel perspective on the foundations of quantum theory. This approach directly bypasses the problems of representation-dependent separation of the Hilbert space vectors into amplitudes and phases that has troubled the attempts to consider the orthodox Hilbert space approach to quantum theory as an extension of probability theory. More generally, the problem of contextuality of probabilistic predictions derived from orthodox approach that are implied by the restriction of spectral representation to commutative subalgebras plays no role in our approach.

On the mathematical level, our approach follows the ideas of an algebraic approach to quantum theory developed by Segal, Haag and others [110, 111, 54] (for reviews see [41, 14, 55, 10, 7]). On the conceptual level, we follow the ideas of Jaynes [67, 68, 69, 70, 73, 74, 76] on the character of statistical mechanics and quantum theory as probabilistic statistical inference theory, and the ideas of Ingarden [59, 60, 61, 62, 63, 64] on the foundational role of the information theory and information geometry in statistical mechanics and quantum theory. In this sense, we aim at unification of the information-theoretic and algebraic perspectives on the foundations of statistical mechanics and quantum theory.

However, as opposed to existing ‘objective bayesian’ (e.g., [121, 122, 19, 18, 20]) and ‘subjective bayesian’ (e.g., [42, 22, 45, 46]) derivations and/or interpretations of elements of the formalism of quantum theory, our derivation is completely independent of the notions and semantics of probability theory and Hilbert space based quantum theory. This is achieved due to the kinematic properties of our approach, which arise as an extension of an algebraic approach to quantum theory (the latter treats non-commutative algebras as fundamental entities of quantum theory).

On the other hand, as opposed to the standard settings of quantum information theory and algebraic quantum theory, our approach can be directly used to construct dynamical (‘interacting’, predictive, experimentally verifiable) models of quantum field theory. This is achieved due to the dynamic properties of our approach which arise as an extension of the approaches of Jaynes [77, 93, 71, 72, 73, 75] and Zubarev [127, 129, 131, 130, 82, 128, 95] to non-equilibrium finite-dimensional quantum statistical mechanics (these approaches derive the time-dependent density operators from the constrained relative entropy maximisation with time-dependent constraints, see [52, 132, 53] for review).

This paper is only a brief introduction to the more detailed mathematical and conceptual discussion provided in the forthcoming series of papers [88]. Its aim is to present the basic principles of the new setting in the context of mathematical formalisms of the standard setting for probability theory and quantum theory.

2 Towards new foundations of probability and statistical inference theory

2.1 Beyond standard approaches

The Bayes–Laplace and the Borel–Kolmogorov approaches to foundations of probability theory can be viewed as extreme cases of an application of two competing principles: evaluational (kinematical) and relational (dynamical). The notion of probability in the Borel–Kolmogorov approach is specified by the *probability measure* $\mu : \mathcal{U}(\mathcal{X}) \rightarrow [0, 1]$ on the countably additive boolean algebra $\mathcal{U}(\mathcal{X})$ of (some) subsets of a given set \mathcal{X} . This

notion is essentially evaluational, and it is justified by an appeal to evaluational framework provided by the measure theory. Thus, it can deal with infinite sets. However, it lacks any *generic* notion of conditionalisation. On the other hand, the notion of probability in the Bayes–Laplace approach is specified by the *conditional probability map* $\mathcal{B} \ni A \mapsto p(A|I) \in [0, 1]$, which assigns a quantitative value to an element of a boolean algebra \mathcal{B} under condition that the element $I \in \mathcal{B}$ has boolean value ‘true’. This notion is essentially relational, and it is usually justified by an appeal to framework of probabilistic statistical inference provided by Bayes’ rule²

$$p(x|\theta) \mapsto p_{\text{new}}(x|\theta) := p(x|\theta) \frac{p(b|x \wedge \theta)}{p(b|\theta)}, \quad (1)$$

Thus, the Borel–Kolmogorov setting is of evaluational (kinematic) character and works well with countably additive algebras, while the Bayes–Laplace setting is of relational (dynamical) character and works well with finitely additive algebras. These two foundational approaches are independent of each other and in this sense are not in any conflict. However, they generate an important *decision problem*, because the Borel–Kolmogorov approach lacks any *generic* notion of conditionalisation of probabilistic inferences under some evidence, while the Bayes–Laplace approach lacks any *generic* extension to countably additive algebras. Thus, each of them is in some sense insufficient. We can resolve this problem by an appeal to two less known foundational approaches (due to Le Cam and Whittle), and by some additional observations.

The approach of Whittle [117, 118] to the foundations of probability theory is based on two independent notions: probabilistic expectations *and* conditional expectations. Both are defined in terms of the Daniell–Stone integrals on the Daniell–Stone vector lattices. For a given set \mathcal{X} the *Daniell–Stone vector lattice* is defined as a lattice \mathfrak{A} with a unit \mathbb{I} such that:

- (i) \mathfrak{A} is a subset of a set of functions $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$,
- (ii) $f, g \in \mathfrak{A} \Rightarrow \sup\{f, g\}, \inf\{f, g\} \in \mathfrak{A}$,
- (iii) $0 \leq f \in \mathfrak{A} \Rightarrow \inf\{f, \mathbb{I}\} \in \mathfrak{A}$.

The Daniell–Stone *integral* is defined as a map $\omega : \mathfrak{A} \rightarrow \mathbb{R}$ that is positive, linear, and monotonically sequentially continuous (that is, $\omega(\inf(f_n)) = \inf(\omega(f_n))$ for every convergent monotone sequence $\{f_n\} \subseteq \mathfrak{A}$ with $\inf\{f_n\} \in \mathfrak{A}$). The *probabilistic expectation* is defined as a normalised Daniell–Stone integral. The *conditional expectation* is then defined [117] as a function $\mathcal{E}_\omega(\cdot|g) \in \mathfrak{A}$ such that

$$\omega((f - \mathcal{E}_\omega(f|g))h(g)) = 0 \quad \forall f \in \mathfrak{A} \quad \forall h : \mathbb{R} \rightarrow \mathbb{R}, \quad (2)$$

²Note that the Ramsey–de Finetti type [107, 35, 36] or Cox’s type [26, 27] derivations of the Bayes theorem (or, equivalently, of the algebraic rules of ‘probability calculus’) *assume* that the conditional probabilities $p(A|I)$ are to be used in order to draw inferences on the base of conditioned premises (evidence). Hence, they *assume* that *some* rule of probability updating has to be used, because only under this assumption it is possible to speak of conditioned elements of the algebra as ‘evidence’, or to speak of conditional probabilities as ‘inferences’. In consequence, the use of the notion of conditional probability amounts to use of *some* probability updating principle in the first place. It amounts to use of some particular algebraic rules of transformation of conditional probabilities only under *additional* assumptions, which might not be relevant in the general case.

where $h(g) \in \mathfrak{A}$. For any two probabilistic expectations ω and ϕ , the conditional expectation \mathcal{E}_ϕ defines uniquely the ‘updating’ (conditioning) of expectation ω by

$$\omega \mapsto \omega_{\text{new}} := \omega(\cdot | \mathcal{E}_\phi(\cdot | g)) := \omega \circ \mathcal{E}_\phi(\cdot | g). \quad (3)$$

Given expectations ω_1, ω_2 , we will call $\omega_3(\cdot | \cdot) := \omega_1 \circ \mathcal{E}_{\omega_2}(\cdot | \cdot)$ a *conditioned expectation*. The *probability* and *conditional probability* are defined, respectively, by $p(A) := \omega(\chi_A)$ and $p(A|B) := \omega(\chi_A | \chi_B)$, where $\chi_A(x)$ is a characteristic function on \mathcal{X} , while ω is a probabilistic expectation.

The reformulation of probability theory by Whittle refers to the underlying ‘sample space’ \mathcal{X} as a domain of definition of the Daniell–Stone lattice \mathfrak{A} . However, it is not necessary to assume this. The Daniell–Stone integration theory requires only to assume that there is given a *Riesz vector lattice*, which is defined by the conditions (ii) and (iii) of definition of Daniell–Stone vector lattice. In such approach, more sophisticated properties of integration theory can be guaranteed by imposing some additional conditions on the Riesz vector lattice. In consequence, the probability theory can be in principle developed for the Riesz vector lattices. This was done by Le Cam [89, 90], who defined the underlying object of a probability theory as a Riesz vector lattice \mathfrak{A} equipped with a Banach norm, which turns \mathfrak{A} into a Banach lattice.

The Daniell–Stone integration theory on Riesz vector lattices is then a promising candidate for the mathematical framework of information theory built without reference to the notion of probability. However, in order to keep backwards compatibility with the Bayes–Laplace and the Borel–Kolmogorov approaches, as well as in order to keep forwards compatibility with the W^* -algebraic reformulation of quantum theory, we need to relate Riesz vector lattices with commutative boolean algebras. This can be provided using the fact that to each boolean algebra \mathcal{B} (which might be finitely additive or countably additive) there corresponds a canonically associated Riesz vector $\mathfrak{A}(\mathcal{B})$, defined as a set of characteristic function of the open-and-compact subsets of the Stone spectrum of \mathcal{B} . The Stone spectrum of \mathcal{B} is defined as a set of all ring homomorphisms from \mathcal{B} to \mathbb{Z}_2 .

If the underlying boolean algebra \mathcal{B} is countably additive, Dedekind complete, and allows at least one strictly positive semi-finite measure, then it is possible to construct a range of commutative L_p spaces over $\mathfrak{A}(\mathcal{B})$, with $p \in [1, \infty]$, that do not depend on the choice of the Daniell–Stone integral on $\mathfrak{A}(\mathcal{B})$ [44]. We will call such boolean algebras *camDcb-algebras*, and denote the corresponding spaces by $L_p(\mathcal{B})$. Any given measure space $(\mathcal{X}, \mathfrak{U}(\mathcal{X}), \mu)$ determines an associated camDcb-algebra \mathfrak{U} by

$$\mathfrak{U} := \mathfrak{U}(\mathcal{X}) / \{A \in \mathfrak{U}(\mathcal{X}) \mid \mu(A) = 0\}. \quad (4)$$

On the other hand, accordingly to the Loomis–Sikorski representation theorem [91, 113], every camDcb-algebra \mathfrak{U} equipped with a measure $\tilde{\mu}$ generates a corresponding measure space $(\mathcal{X}, \mathfrak{U}(\mathcal{X}), \mu)$.

Given a camDcb-algebra \mathfrak{U} , we consider the spaces $\mathcal{M}(\mathfrak{U}) \subseteq L_1(\mathfrak{U})^+$ of finite positive Daniell–Stone integrals over \mathfrak{U} as a kinematic model for commutative information theory. This way we can unify both standard approaches to foundations of probability theory. The setting of Borel–Kolmogorov approach is recovered by restriction to the evaluational component, via expectation ω on the lattice of characteristic functions over camDcb-algebra \mathfrak{U} and the Loomis–Sikorski representation of \mathfrak{U} in terms of a measure space for a measure $\tilde{\mu}_\omega$ on \mathfrak{U} defined by this expectation. The setting of Bayes–Laplace approach is

recovered by restriction to the relational component, via conditional expectations on the lattice of characteristic functions over finitely additive boolean algebra \mathcal{B} .

In order to propose a dynamics of this theory, we will replace the relational component of Whittle's (and Le Cam's³) approach by minimisation of some more general functional on $\mathcal{M}(\mathcal{U})$. The reason for the replacement is an important disproportion between the kinematic and dynamic structures of Whittle's approach: while the probabilistic expectation is just a normalised Daniell–Stone integral, the definition of a conditional expectation seems to be an *ad hoc* postulate. If $\omega(f^2) < \infty$, what is equivalent with $f \in L_2(\mathfrak{A}, \omega)$, then (2) is equivalent with a variational definition [118]

$$\mathcal{E}_\omega(f|g) := \arg \inf_{f_e \in \mathfrak{A}} \omega((f - f_e(f, g))^2), \quad (5)$$

but there is still no justification given of a choice of such functional. Let us note that, due to duality between $L_1(\mathfrak{A})$ and $L_\infty(\mathfrak{A})$, the use of conditional expectation as a mapping on the space of integrable functions

$$L_\infty(\mathfrak{A}) \ni f \mapsto \mathcal{E}_\phi(f, g) \in L_\infty(\mathfrak{A})$$

is equivalent with the use of \mathcal{E}_ϕ as a mapping on the space of integrals

$$L_1(\mathfrak{A}) \ni \omega \mapsto \omega \circ \mathcal{E}_\phi(\cdot, g) \in L_1(\mathfrak{A}).$$

This means that conditional expectations can be considered as some particular mappings on information models $\mathcal{M}(\mathcal{U}) \subseteq L_1(\mathcal{U})^+$, which are (dually) defined by some particular variational principle (5). This leads us to ask about the possibility of introducing other more general dynamic (relational) principle on $\mathcal{M}(\mathcal{U})$, which would include the updating by conditional expectations as a special case.

2.2 Information geometry

Consider the notion of relative distance between the finite positive integrals, defined by the *deviation* (negative *relative entropy*) functionals,

$$D : L_1(\mathcal{U})^+ \times L_1(\mathcal{U})^+ \ni (\omega, \phi) \mapsto D(\omega, \phi) \in [0, +\infty],$$

such that $D(\omega, \phi) \geq 0$ and $D(\omega, \phi) = 0$ iff $\omega = \phi$. Deviation is one of two foundational structures of the information geometry theory [24, 2, 96, 4], allowing to introduce and analyse the geometric structures on information models $\mathcal{M}(\mathcal{U})$. The other is the structure of the differential manifold that can be imposed on $\mathcal{M}(\mathcal{U})$ by the local coordinate systems given by embeddings into appropriate Banach spaces (in the finite dimensional case these are \mathbb{R}^n [24, 2], while in the infinite dimensional case these are L_{Φ_1} Orlicz spaces [105]). We will use the framework of information geometry to define the dynamical principle of the information theory.

Eguchi showed [39, 40], for $\dim \mathcal{M}(\mathcal{U}) < \infty$, that any deviation that is symmetric in first derivatives and has a negative definite hessian determines uniquely (up to a scalar

³The relational component of Le Cam's approach [89, 90] is provided by *transition maps*, which in our case are equivalent with the positive norm-preserving linear maps $T^c : \mathcal{M}_1(\mathcal{U}_1) \rightarrow \mathcal{M}_2(\mathcal{U}_2)$. The mappings (3), generated by conditional expectations, are special case of these transition maps.

factor) the riemannian metric g and a pair (∇, ∇^*) of two affine connections on $\mathcal{M}(\mathcal{U})$. These objects are defined by the equations

$$\begin{aligned} g_\mu(u, v) &:= (\partial_u)_\mu(\partial_v)_\mu D(\nu, \mu)|_{\nu=\mu}, \\ g_\mu((\nabla_\mu)_u v, w) &:= -(\partial_u)_\mu(\partial_v)_\mu(\partial_w)_\nu D(\nu, \mu)|_{\nu=\mu}, \\ g_\mu(v, (\nabla_\mu^*)_u w) &:= -(\partial_u)_\nu(\partial_w)_\nu(\partial_v)_\mu D(\nu, \mu)|_{\nu=\mu}, \end{aligned}$$

where $(\partial_u)_\mu$ is a Gâteaux derivative at $\mu \in \mathcal{M}(\mathcal{U})$ in the direction $u \in \mathbf{T}\mathcal{M}(\mathcal{U})$ [40]. If the torsion and Riemann curvature tensors of both these connections are equal to zero, then $(\mathcal{M}(\mathcal{U}), g, \nabla, \nabla^*)$ is called a *dually flat manifold* and there exists a pair (ℓ, ℓ^*) of coordinate systems on $\mathcal{M}(\mathcal{U})$, called *dually flat coordinates*, such that ℓ consist of ∇ -geodesics, while ℓ^* consists of ∇^* -geodesics. They are called *orthogonal* at $q \in \mathcal{M}(\mathcal{U})$ iff

$$g_q((\partial_\ell)_q, (\partial_{\ell^*})_q) = 0.$$

This is equivalent to the condition

$$\llbracket \ell(q), \ell^*(q) \rrbracket = 0,$$

where $\llbracket \cdot, \cdot \rrbracket$ denotes the duality between vectors spaces that are codomains of the coordinate systems ℓ and ℓ^* (for $\dim \mathcal{M}(\mathcal{U}) < \infty$ this is just a self-duality of \mathbb{R}^n understood as vector and covector space, while for $\dim \mathcal{M}(\mathcal{U}) = \infty$ this is a Banach space duality between Orlicz spaces [48]). A point $p_Q \in \mathcal{Q} \subseteq \mathcal{M}(\mathcal{U})$ is called a ∇ -*projection* of $p \in \mathcal{M}(\mathcal{U})$ iff the ∇ -geodesic connecting p and p_Q is orthogonal to \mathcal{Q} with respect to g , while $\mathcal{Q} \subseteq \mathcal{M}(\mathcal{U})$ is called ∇ -*convex* iff for all $p_1, p_2 \in \mathcal{Q}$ there exists a unique ∇ -geodesic connecting p_1 and p_2 and entirely included in \mathcal{Q} . As shown by Amari [2] (and earlier results by Chentsov [23, 24] and Csiszár [29]), if \mathcal{Q} is a ∇^* -convex closed set then the unique ∇ -projection p_Q of $p \in \mathcal{M}(\mathcal{U})$ on \mathcal{Q} is equal to D -projection $\mathfrak{P}_Q^D(p)$ [2]

$$\mathcal{Q} \ni p_Q = \mathfrak{P}_Q^D(p) := \arg \inf_{q \in \mathcal{Q}} D(p, q). \quad (6)$$

Hence, the dual flatness of the geometry of $\mathcal{M}(\mathcal{U})$ is a necessary condition for the existence of the unique ∇ -projections on the ∇^* -convex subspaces provided by the minimum of deviation functional.

2.3 Some special information geometries

Let us consider now three important families of deviation functionals (which determine the corresponding families of Norden–Sen geometries).

1. The family of *Bregman deviations* [15] can be defined for $\dim \mathcal{M}(\mathcal{U}) < \infty$ by [11, 123]

$$D_\Psi(p_1, p_2) := \Psi(\ell_\Psi(p_1)) + \Psi^{\mathbf{L}}(\ell_\Psi^*(p_2) - \llbracket \ell_\Psi(p_1), \ell_\Psi^*(p_2) \rrbracket),$$

where $\Psi : \mathbb{R}^n \rightarrow]-\infty, +\infty]$ is a convex function, $\Psi^{\mathbf{L}} : \mathbb{R}^n \rightarrow]-\infty, +\infty]$ is its convex Legendre–Fenchel dual function, defined by

$$\Psi^{\mathbf{L}}(y) := \sup_{x \in \mathbb{R}^n} \{ \llbracket x, y \rrbracket - \Psi(x) \} \quad \forall y \in \mathbb{R}^n,$$

$\ell_\Psi : \mathcal{M}(\mathcal{U}) \rightarrow \mathbb{R}^n$, while its dual function $\ell_\Psi^* : \mathcal{M}(\mathcal{U}) \rightarrow \mathbb{R}^n$ is defined by

$$\ell_\Psi^*(q) = \text{grad} \Psi(\ell_\Psi(q)) \quad \forall q \in \mathcal{M}(\mathcal{U}).$$

Every Bregman deviation satisfies the *generalised cosine equation*

$$D_{\Psi}(p_1, p_2) + D_{\Psi}(p_2, p_3) - D_{\Psi}(p_1, p_3) = [\ell_{\Psi}(p_1) - \ell_{\Psi}(p_2), \ell_{\Psi}^*(p_3) - \ell_{\Psi}^*(p_2)].$$

Moreover, the Norden–Sen geometry of arbitrary Bregman deviation is dually flat, with dually flat coordinates $(\ell_{\Psi}, \ell_{\Psi}^*)$. Nagaoka and Amari [97] show that any Norden–Sen dually flat geometry with the pair (ℓ, ℓ^*) of dually flat coordinates determines a unique corresponding Bregman deviation.

2. The family of *Csiszár–Morimoto deviations* [28, 94, 1] is defined by

$$D_{\mathfrak{f}}(\mu, \nu) := \int \mu \mathfrak{f}\left(\frac{\nu}{\mu}\right), \quad (7)$$

where $\mathfrak{f} : \mathbb{R}^+ \rightarrow \mathbb{R}$ is a convex function with $\mathfrak{f}(1) = 0$. It was characterised by Csiszár [30] for $\dim \mathcal{U} < \infty$ by the condition of non-increasing under the information loss provided by partitioning of the space \mathcal{X} and invariance under permutations of partitions. This condition can be restated in representation-independent terms as

$$D(\mu, \nu) \geq D(\mu \circ T, \nu \circ T), \quad (8)$$

where $T : L_{\infty}(\mathcal{U}) \rightarrow L_{\infty}(\mathcal{U}')$ are *Markov maps*, defined as sequentially continuous functions that are linear and positive ($0 \leq x \leq \mathbb{I} \Rightarrow 0 \leq T(x) \leq \mathbb{I}$).⁴ The condition (8) is called *Markov monotonicity*.

3. The family of *Zhu–Rohwer deviations* [125, 126] is defined by

$$D_{\gamma}(\mu, \nu) := \begin{cases} \int \left(\frac{\mu}{1-\gamma} + \frac{\nu}{\gamma} - \frac{\mu^{\gamma} \nu^{1-\gamma}}{\gamma(1-\gamma)} \right) & : \gamma \in]0, 1[\\ \int \lim_{\gamma' \rightarrow \gamma} \left(\frac{\mu}{1-\gamma'} + \frac{\nu}{\gamma'} - \frac{\mu^{\gamma'} \nu^{1-\gamma'}}{\gamma'(1-\gamma')} \right) & : \gamma \in \{0, 1\}. \end{cases}$$

It reduces to

$$D_1(\mu, \nu) := \int \left(\mu - \nu + \mu \log \left(\frac{\mu}{\nu} \right) \right) = D_0(\nu, \mu)$$

for $\gamma \in \{0, 1\}$, and to the negative of the *Kullback–Leibler relative entropy*

$$-\mathbf{S}_{KL}(\mu, \nu) := D_1(\mu, \nu)|_{L_1(\mathcal{U})_1^+} = \int \mu \log \left(\frac{\mu}{\nu} \right) = D_0(\nu, \mu)|_{L_1(\mathcal{U})_1^+}$$

for $\gamma \in \{0, 1\}$ and $L_1(\mathcal{U})_1^+ := \{\omega \in L_1(\mathcal{U})^+ \mid \omega(\mathbb{I}) = 1\}$. The family D_{γ} belongs to the class of Bregman deviations. The corresponding dually flat coordinates are given by γ -embeddings

$$\ell_{\gamma} : \mathcal{M}(\mathcal{U}) \ni \mu \mapsto \frac{\mu^{\gamma}}{\gamma} \in L_{1/\gamma}(\mathcal{U}),$$

which generalise the Nagaoka–Amari [97, 2] γ -embeddings and the Zhu–Rohwer [126, 124] γ -embeddings. The Chentsov–Amari dual coordinate system for $\gamma \in]0, 1[$ is given by $(\ell_{\gamma})^* = \ell_{1-\gamma}$ (the case $\gamma \in \{0, 1\}$ requires separate treatment).

⁴If the sequential continuity is replaced by a stronger condition of norm continuity or weak topological continuity (which seems to be necessary prerequisite for potential characterisation of (7) by (8) in $\dim \mathcal{U} = \infty$ case), then the maps T are just the duals of Le Cam’s transition maps T^c in terms of the Steinhaus–Nikodým $L_1(\mathcal{U})^{\mathbf{B}} \cong L_{\infty}(\mathcal{U})$ Banach space duality.

Csiszár [31] showed that $D_1(\mu, \nu)|_{L_1(\mathcal{U})_1^+}$ are the unique deviations on $L_1(\mathcal{U})_1^+$ that are Csiszár–Morimoto deviations *and* Bregman deviations. Amari [3] showed that D_γ are the unique deviations on $L_1(\mathcal{U})^+$ that are satisfying this condition. Both proofs are provided for $\dim \mathcal{U} < \infty$ case.

2.4 Information dynamics

Using this characterisation of D_γ , we propose a new approach unifying probability theory and statistical inference theory in a single theory of inductive inference (information dynamics), independent of the notion of probability. Its evaluational-kinematic component was already specified as given by the spaces $\mathcal{M}(\mathcal{U})$ of finite positive integrals and their geometry.

Its relational component is given in the following way. Let P be a positive measure on $\mathcal{M}(\mathcal{U})$. Define a (D, P) -optimal estimate with respect to $\mathcal{Q} \subseteq \mathcal{M}(\mathcal{U})$ as

$$\mathfrak{P}^{D,P}(p) := \arg \inf_{q \in \mathcal{Q}} \int_{p \in \mathcal{M}(\mathcal{U})} P(p) D(p, q) \in \mathcal{Q},$$

and define a (D, P) -ideal estimate as a (D, P) -optimal estimate with respect to $\mathcal{M}(\mathcal{U})$. These notions provide the generalisation of the notion of D -projection (6). Under two assumptions:

- (1) the inference procedure should select (D, P) -optimal estimates,
- (2) D should be Markov monotone Bregman deviation

we can define the *information dynamics* as given by the *constrained maximum relative γ -entropy updating rule*,

$$\mathcal{M}(\mathcal{U}) \ni p \mapsto \mathfrak{P}_F^{D_\gamma, E}(p) := \arg \inf_{q \in \mathcal{M}(\mathcal{U})} \left\{ \int_{p' \in \mathcal{M}(\mathcal{U})} E(p', p) D_\gamma(p', q) + F(q) \right\} \in \mathcal{M}(\mathcal{U}), \quad (9)$$

where $E(\cdot, p)$ is a positive measure on $\mathcal{M}(\mathcal{U})$, while $F : \mathcal{Q} \rightarrow]-\infty, +\infty]$ is a constraining function. The map (9) provides a selection of D_γ -optimal estimates with respect to \mathcal{Q} that are weighted by E and satisfy the constraints F . If $F(q)$ and $\int_{p' \in \mathcal{M}(\mathcal{U})} E(p', p) D_\gamma(p', q)$ are lower semi-continuous and convex in q , and if the infimum in (9) is finite, then the updating procedure selects a unique $\mathfrak{P}_F^{E, D_\gamma}(p)$. In particular, $\int_{p' \in \mathcal{M}(\mathcal{U})} E(p', p) D_\gamma(p', q)$ is lower semi-continuous and convex, if $E(p', p) = dp' \delta(p - p')$, and in such case the existence and uniqueness of $\mathfrak{P}_F^{D_\gamma}(p)$ is guaranteed by the requirement that $\ell_{1-\gamma}(\mathcal{Q})$ is non-empty, closed and convex (see e.g. [79]). The measure $E(\cdot, p)$ will be called a *prior relative to p* .

We consider (9) as a fundamental principle of information dynamics (in commutative case). In our opinion, its status in information theory is analogous to the status of lagrangean variational principle in classical field theory.

The temporal character of this information dynamics might enter through the dependence of constraints F on the external time parameter t . If $E(p', p) = dp' \delta(p - p')$, then p plays a role of the initial state of this dynamics. Under these conditions, if the constraints $F(q)$ have the form $F(q, t)$, where the parameter t is interpreted as ‘time’, if $F(q, t = t_0) = 0$, and if the solutions $\mathfrak{P}_F^{D_\gamma}(p)$ also depend on t , then the trajectory

$$p(t) := \mathfrak{P}_{F(t)}^{D_\gamma}(p_0) \in \mathcal{M}(\mathcal{U})$$

with $p_0 := \mathfrak{P}_{F(t=t_0)}^{D_\gamma}(p)$ can be considered as a temporal evolution generated by the constraint $F(q, t)$ of relative entropic information dynamics (9). We will call $p(t)$ *temporal information dynamics* or *entropic evolution of information*.

2.5 Reconstruction of the standard frameworks

Our approach provides a particular solution to the problem of the relationship between information, inference and probability. The probability theory is just a special case of the information theory and requires no additional justification. The Whittle approach is recovered by the fact [8] that the duals of conditional expectations are characterised as minimisers of expectations of Bregman deviations, the Borel–Kolmogorov approach is recovered by forgetting the relational component, and passing to probabilistic measures, while the Bayes–Laplace approach is recovered by normalisation and by the fact [21, 49] that the Bayes rule is just a special case of the entropic updating rule with Dirac’s delta constraints.

Note that the differences between γ -ideal estimates (defined as $(D_\gamma, dp'\delta(p' - p))$ -ideal estimates) appear only in the third order of the Taylor series expansion of D_γ . Up to the second order, all D_γ are determined by the same riemannian metric g , which reduces to the Fisher–Rao metric for $L_1(\mathcal{U})_1^+$. Hence, the inferences based on D_γ on $\mathcal{M}(\mathcal{U})$ agree with the inferences based on $D_{1/2}$, which in turn agree with the inferences based on the Hilbert space L_2 norm. However, beyond the second order, the use of the L_2 space and its norm is *not* justified. Moreover, the projections in Hilbert space are linear operators, hence they can deal only with linear problems (constraints), as opposed to non-linear γ -projections onto convex subspaces of $L_{1/\gamma}$ spaces (with L_1 spaces playing most important role). The constrained relative γ -entropy updating incorporates Hilbert space based methods of projection and estimation as a special case. Under restriction to $L_1(\mathcal{U})_1^+$, the only justified method of inference is given by constrained maximisation of the Kullback–Leibler relative entropy \mathbf{S}_{KL} .

3 Towards new foundations of quantum theory

In this section, we will present the elements of a new approach to foundations of quantum theory, which is a *direct* generalisation of our reformulation of foundations of probability theory and statistical inference theory. The steps of the construction of this approach are the following: (1) construction of quantum information kinematics based on integration theory on non-commutative algebras and quantum information geometry of the spaces of the positive finite integrals, (2) construction of quantum information dynamics based on maximisation of constrained quantum relative entropy, (3) reconstruction of the orthodox kinematical and dynamical Hilbert space based framework for quantum theory as a special case of a new framework, (4) construction of adequate interpretation that covers the conceptual issues of the quantum theory using the new setup. We will discuss here mainly the points (1) and (2).

3.1 Non-commutative integration theory

An algebraic approach to the mathematical framework of quantum theory [110, 111] replaces the consideration of abstract Hilbert space by the consideration of abstract C^* -algebra \mathcal{A} with unit \mathbb{I} , which is also a Banach space over \mathbb{C} , and is equipped with a map

$*$: $\mathcal{A} \rightarrow \mathcal{A}$ such that

$$(ab)^* = b^*a^*, \quad (a+b)^* = a^* + b^*, \quad (\lambda a)^* = \bar{\lambda}a^*, \quad (a^*)^* = a,$$

$$\|a^*\| = \|a\|, \quad \|ab\| \leq \|a\|\|b\|, \quad \|\mathbb{I}\| = 1, \quad \|a^*a\| = \|a\|^2.$$

An *algebraic state* is a functional $\omega : \mathcal{A} \rightarrow \mathbb{C}$ that is linear, positive and normalised. A basic example of a C^* -algebra is $\mathfrak{B}(\mathcal{H})$, while a basic example of an algebraic state on it is $\text{tr}(\rho \cdot) : \mathfrak{B}(\mathcal{H}) \rightarrow \mathbb{C}$, where ρ is a density operator. While every algebraic state ω on $\mathfrak{B}(\mathcal{H})$ can be represented in terms of $\text{tr}(\rho_\omega \cdot)$, it is no longer true for other C^* -algebras that naturally arise in relativistic quantum field theory and continuous quantum statistical mechanics (see e.g. [14, 55]).

For any pair of a C^* -algebra \mathcal{A} and linear, positive functional ω on \mathcal{A} , the Gel'fand–Naimark–Segal (GNS) theorem associates a unique Hilbert space \mathcal{H}_ω and a unique (up to unitary equivalence) representation $\pi_\omega : \mathcal{A} \rightarrow \mathfrak{B}(\mathcal{H}_\omega)$ such that there exists $\Omega_\omega \in \mathcal{H}_\omega$ that is cyclic for $\pi_\omega(\mathcal{A})$, $\|\Omega_\omega\|^2 = 1$, and $\omega(A) = \langle \Omega_\omega, \pi_\omega(A)\Omega_\omega \rangle_\omega$ for all $A \in \mathcal{A}$. This theorem allows to bypass the mathematical setting of Hilbert spaces, dealing directly with the representation-independent properties of quantum theoretic formalism in terms of C^* -algebras and linear positive functionals on them.

In order to construct the new formulation of quantum theory, we need to resign from Hilbert spaces as the kinematic setting *and* to replace it by the setting based on integration theory on non-commutative algebras. Let us start from identification of the correct notions of an ‘integrable non-commutative algebra’ and of an ‘integral’ on it. A W^* -algebra is defined as such C^* -algebra \mathcal{N} which is a Banach dual to some Banach space. This Banach space is called a *predual* of \mathcal{N} and is denoted by \mathcal{N}_* . The W^* -algebra generalises the notion of an ‘integrable’ commutative algebra to the non-commutative case. The space \mathcal{N}_* consists of all functionals on a W^* -algebra \mathcal{N} that are *normal*, that is, $\omega(\sup \mathfrak{F}) = \sup_{a \in \mathfrak{F}} \omega(a)$ for every directed filter $\mathfrak{F} \subseteq \mathcal{N}$ with the upper bound $\sup \mathfrak{F}$. The spaces $\mathcal{N}_*^+ := \{\omega \in \mathcal{N}_* \mid \omega(A^*A) \geq 0\}$ of *quantum information states* provide the non-commutative analogue of the spaces $\mathcal{M}(\mathcal{U})$ of finite Daniell–Stone integrals on $\mathfrak{A}(\mathcal{U})$. The spaces $\mathcal{N}_*^+ := \{\omega \in \mathcal{N}_*^+ \mid \omega(\mathbb{I}) = 1\}$ of normalised quantum information states provide the non-commutative analogue of the spaces $L_1(\mathcal{U})_1^+$.

This is actually much more than just analogy. Falcone and Takesaki [43] formulated the canonical theory of non-commutative $L_p(\mathcal{N})$. The elements of $L_p(\mathcal{N})$ can be represented in form $x\omega^{1/p}$, where $x \in \mathcal{N}$ and $\omega \in \mathcal{N}_*$. This theory provides a notion of the integral $\int : L_1(\mathcal{N}) \rightarrow \mathbb{C}$ such that, in particular, $\int x\omega = \omega(x) \in \mathbb{C}$. This integral is meaningful also for such formulas like

$$\int x_1\phi_1^{\alpha_1} \cdots x_n\phi_n^{\alpha_n}$$

for $\phi_1, \dots, \phi_n \in \mathcal{N}_*^+$ and $x_1, \dots, x_n \in \mathcal{N}$ and $\alpha_1, \dots, \alpha_n \in \{z \in \mathbb{C} \mid \text{re}(z) \geq 0\}$ with $\sum_{i=1}^n \alpha_i = 1$. The bilinear map

$$L_1(\mathcal{N}) \times \mathcal{N} \ni (x, T) \mapsto \llbracket x, T \rrbracket := \int xT \in \mathbb{C}$$

gives the pairing between \mathcal{N} and $L_1(\mathcal{N})$ which identifies $L_1(\mathcal{N})$ with \mathcal{N}_* . By means of definition, $L_\infty(\mathcal{N})$ is identified with \mathcal{N} . The duality between $L_p(\mathcal{N})$ and $L_q(\mathcal{N})$, for $1/p + 1/q = 1$, is given by

$$L_p(\mathcal{N}) \times L_q(\mathcal{N}) \ni (S, T) \mapsto \llbracket S, T \rrbracket = \int ST \in \mathbb{C}.$$

The analytic structure of $L_p(\mathcal{N})$ is defined by the Banach norm $\|T\|_p := (\int |T|^p)^{1/p}$.

From the perspective of non-commutative integration theory, evaluation of $\text{tr}(\rho \cdot)$ on $\mathfrak{B}(\mathcal{H})$ as well as the evaluation of probability measure $\text{tr}(\rho E(\cdot))$ on boolean algebra $\mathcal{U}(\mathcal{X})$ (where E is a semi-spectral measure, while \mathcal{X} is usually identified with the spectrum of some operator) can be replaced by acting with $\omega \in \mathcal{N}_*^+ \cong L_1(\mathcal{N})^+$ on W^* -algebra \mathcal{N} . Because semi-spectral measures E are no longer used in this setting in order to define the probabilistic evaluations, von Neumann's spectral theorem loses its importance. Hence, the associated problems of contextuality of definition of operators corresponding to probabilistic descriptions, particular quantitative experimental setups and results of their use (see e.g. [17]) disappear. The *only* meaning of elements of W^* -algebras in our approach is: they are the *statements* of an (intersubjectively shared) *abstract language* subjected to quantitative evaluation.

Now we need to equip this general setting with the information geometric and dynamical structures.

3.2 Quantum information geometry

Our novel result is a generalisation of quantum information geometric structures which is based the Falcone–Takesaki theory. In analogy with Nagaoka–Amari [97, 4] and Zhu–Rohwer [125, 126] approaches, we define the *quantum γ -embeddings* (γ -coordinates) by

$$\ell_\gamma : \mathcal{N}_*^+ \ni \omega \mapsto \ell_\gamma(\omega) := \omega^\gamma / \gamma \in L_{1/\gamma}(\mathcal{N}).$$

We define the *quantum γ -deviation* $D_\gamma : \mathcal{N}_*^+ \times \mathcal{N}_*^+ \ni (\omega, \phi) \mapsto D_\gamma(\omega, \phi) \in \mathbb{R}$ by

$$D_\gamma(\omega, \phi) := \int \left(\frac{\omega}{1-\gamma} + \frac{\phi}{\gamma} - \text{re}(\ell_\gamma(\omega)\ell_{1-\gamma}(\phi)) \right) = \int \left(\frac{\omega}{1-\gamma} + \frac{\phi}{\gamma} - \frac{\text{re}(\omega^\gamma \phi^{1-\gamma})}{\gamma(1-\gamma)} \right),$$

for $\gamma \in]0, 1[$, and by the limit $\gamma \rightarrow \gamma'$ under integral sign for $\gamma' \in \{0, 1\}$. The *quantum relative γ -entropy* is defined as $\mathbf{S}_\gamma(\omega, \phi) := -D_\gamma(\omega, \phi)$, and $-\mathbf{S}(\phi, \omega) := D_0(\omega, \phi) = D_1(\phi, \omega)$. This definition reduces to the Jenčová–Ojima [78, 79, 100] quantum γ -deviation for any particular choice of reference weight ψ (due to isometric isomorphism of the Falcone–Takesaki $L_p(\mathcal{N})$ spaces with the Kosaki–Terp $L_p(\mathcal{N}, \phi)$ spaces), to the Umegaki–Araki quantum relative entropy [116, 5, 6] in the Petz [102] reference-independent form $\lim_{t \rightarrow +0} \frac{i}{t} \phi([D\omega : D\phi]_t - \mathbb{I})$ for $\omega, \phi \in \mathcal{N}_{*1}^+$, to the Hasegawa quantum γ -deviation [58] $\text{tr}(\rho_\phi - \rho_\omega^\gamma \rho_\phi^{1-\gamma}) / (\gamma - \gamma^2)$ for $\omega = \text{tr}(\rho_\omega \cdot)$, $\phi = \text{tr}(\rho_\phi \cdot)$ and type I W^* -algebra \mathcal{N} , and to the Zhu–Rohwer γ -deviation $D_\gamma(\omega, \phi)$ for commutative \mathcal{N} . The symbol $[D\omega : D\phi]_t$ denotes Connes' cocycle [25], which is a non-commutative generalisation of the Radon–Nikodým derivative.

The properties of D_γ are analysed in [85]. They are the same as the properties of Jenčová–Ojima deviation [79], but with a key difference that the definition and properties of D_γ explicitly do not depend on any reference weight ψ . For $\gamma \in]0, 1[$, D_γ is both Petz's \mathfrak{f} -deviation [101] (which is a non-commutative generalisation of the Csiszár–Morimoto deviation) and the non-commutative generalisation of Bregman's deviation (see [79, 103] for special cases, and [85] for a general definition). D_γ (resp., \mathbf{S}) is convex and lower semi-continuous on the space $\mathcal{N}_{*0}^+ \times \mathcal{N}_{*0}^+$ (resp., $\mathcal{N}_{*0}^+ \times \mathcal{N}_{*0}^+$) endowed with the product of norm topologies [79] (resp., the topology of pointwise convergence on \mathcal{N}_* [99]). Here $\mathcal{N}_{*0}^+ := \{\omega \in \mathcal{N}_*^+ \mid \omega(x^*x) = 0 \Rightarrow x = 0\}$ denotes the space of all *faithful* quantum

information states on \mathcal{N} . If $\mathcal{Q} \subseteq \mathcal{M}(\mathcal{N})$ is weakly closed and if $\ell_\gamma(\mathcal{Q})$ is convex, then there exists a unique D_γ -projection on \mathcal{Q} ,

$$\mathcal{M}(\mathcal{N}) \ni \omega \mapsto \mathfrak{P}_{\mathcal{Q}}^{D_{1-\gamma}}(\omega) = \arg \inf_{\phi \in \mathcal{Q}} D_{1-\gamma}(\omega, \phi) \in \mathcal{Q}.$$

Following Amari's and Csiszár's results in commutative case, *we conjecture that D_γ (resp., D_1) is a unique deviation on \mathcal{N}_*^+ (resp., \mathcal{N}_{*1}^+) that belongs to both Petz's and non-commutative Bregman's families of quantum deviations.*

In [80, 81] Jenčová has introduced the differential manifold structure on the space \mathcal{N}_{*0}^+ . Using this differential structure, one can consider the tangent spaces, riemannian metrics and affine connections on $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_{*0}^+$. For $\dim \mathcal{M}(\mathcal{N}) = \infty$, the construction of tangent space of $\mathcal{M}(\mathcal{N})$ is a delicate issue. As a result, the standard approach [48, 47] is to define differential geometric objects on $\mathcal{M}(\mathcal{N})$ using the family of vector γ -bundles \mathcal{F}^γ over $\mathcal{M}(\mathcal{N})$ instead of tangent bundle $\mathbf{TM}(\mathcal{N})$. Following Gibilisco–Pistone [48], Gibilisco–Isola [47] and Jenčová [79], we generalise the γ -representation of γ -bundles using the embeddings into the tangent spaces of unit spheres $S_{1/\gamma}(\mathcal{N})$ of non-commutative $L_{1/\gamma}(\mathcal{N})$ spaces,

$$\mathcal{F}_\omega^\gamma(\mathcal{M}(\mathcal{N})) \ni v \mapsto \ell_\gamma^{\#(\omega)}(v) := x\phi^\gamma \in T_{\omega^\gamma} S_{1/\gamma}(\mathcal{N}),$$

where

$$\begin{aligned} T_\omega S_{1/\gamma}(\mathcal{N}) &:= \{x\phi^\gamma \in L_{1/\gamma}(\mathcal{N}) \mid \operatorname{re} \int \omega^{1-\gamma} x\phi^\gamma = 0\} \\ &= \{x\phi^\gamma \in L_{1/\gamma}(\mathcal{N}) \mid \operatorname{re} \omega([D\omega : D\phi]_{i_\gamma}(x)) = 0\}. \end{aligned} \quad (10)$$

In the special cases, this definition reduces to those given in [79, 47], and [48]. It allows us to define a *quantum γ -metric* $g_\omega^\gamma : \mathcal{F}_\omega^\gamma(\mathcal{M}(\mathcal{N})) \times \mathcal{F}_\omega^\gamma(\mathcal{M}(\mathcal{N})) \rightarrow \mathbb{R}$ by

$$g_\omega^\gamma(u, v) := \operatorname{re} \int \ell_\gamma^{\#(\omega)}(u)(j_\gamma \circ \ell_\gamma)^{\#(\omega)}(v) = \operatorname{re} \int \omega^{1-\gamma} x\omega^\gamma y, \quad (11)$$

where $j_\gamma := \ell_{1-\gamma} \circ \ell_\gamma^{-1}$ and $x, y \in \mathcal{N}$. The quantum γ -metric (11) generalises the Wigner–Yanase–Dyson γ -metric [119, 58]. The generalisation of the Gibilisco–Isola family of γ -connections [47], based on the projection of trivial affine structure of non-commutative $L_{1/\gamma}(\mathcal{N})$ spaces onto unit spheres $S_{1/\gamma}(\mathcal{N})$ is straightforward, and will be discussed, together with other quantum information geometric structures, in [86].

3.3 Quantum information dynamics

Using Bauer's maximisation principle [9] and the results of Jenčová [79], we conclude that if $\mathcal{Q} \subseteq \mathcal{M}(\mathcal{N})$ is a nonempty set such that $\ell_\gamma(\mathcal{Q}) \subset L_{1/\gamma}(\mathcal{N})$ is weakly closed and convex, and $F : \mathcal{Q} \rightarrow]-\infty, +\infty]$ is a weakly- $*$ lower semi-continuous convex function, then there exists a unique

$$\mathfrak{P}_F^{D_\gamma}(\omega) := \arg \inf_{\phi \in \mathcal{M}(\mathcal{N})} \{D_{1-\gamma}(\omega, \phi) + F(\phi)\}$$

if this infimum is finite.

We define the *quantum constrained maximum relative γ -entropy updating rule* by

$$\mathcal{M}(\mathcal{N}) \ni \omega \mapsto \mathfrak{P}_F^{D_\gamma, E}(\omega) := \arg \inf_{\phi \in \mathcal{M}(\mathcal{N})} \left\{ \int_{\varphi \in \mathcal{M}(\mathcal{N})} E(\varphi, \omega) D_\gamma(\varphi, \phi) + F(\phi) \right\} \in \mathcal{M}(\mathcal{N}), \quad (12)$$

for the positive measure $E(\cdot, \omega)$ on $\mathcal{M}(\mathcal{N})$, and constraints given by convex weak-* lower semi-continuous $F : \mathcal{Q} \rightarrow]-\infty, +\infty]$ with convex closed $\ell_{1-\gamma}(\mathcal{Q})$, which guarantees the existence and uniqueness of the result of this updating when this infimum takes a finite value and if $E(\varphi, \omega) = d\varphi\delta(\varphi - \omega)$. This rule equips the evaluational theory of integrals on non-commutative algebras with the relational (dynamical) counterpart, which provides a selection of γ -optimal estimates that are weighted by E , relative to \mathcal{Q} , and constrained by F . This generalises and replaces the Hilbert space based approach. The convex closed subspaces of Hilbert spaces associated with the orthogonal linear projection operators are replaced in our approach by the convex closed subspaces of non-commutative $L_{1/\gamma}(\mathcal{N})$ spaces, which used as the codomain of quantum information geometric embeddings ℓ_γ of the functionals $\omega \in \mathcal{M}(\mathcal{N})$. If $E(\varphi, \omega) = d\varphi\delta(\varphi - \omega)$, the constraints $F(\phi)$ depend on (discrete or continuous) ‘time’ parameter t in a way that $F(\phi, t = t_0) = 0$ and the solutions of (12) reflect this dependence on t , then the resulting trajectory $\mathfrak{P}_{F(t)}^{D_\gamma}(\omega_0)$ on quantum information manifold, where $\omega_0 := \mathfrak{P}_{F(t=t_0)}^{D_\gamma}(\omega)$ can be understood as a non-linear *temporal dynamics of quantum states*.

Interpretation of spaces $\mathcal{M}(\mathcal{N})$ as a general framework for non-linear quantum theoretic models and interpretation of the updating rule (12) as a general description of their non-linear behaviour (quantum dynamics) opens new perspectives for the foundations and applications of quantum theory.

3.4 Reconstruction of the Hilbert space based framework

The reconstruction of kinematic setting of ordinary Hilbert space based quantum theory is provided by the quantum information geometric representation of $\mathcal{M}(\mathcal{N})$ in terms of the non-commutative $L_2(\mathcal{N})$ space,

$$\ell_{1/2} : \mathcal{M}(\mathcal{N}) \ni \phi \mapsto 2\phi^{1/2} \in L_2(\mathcal{N}).$$

The space $L_2(\mathcal{N})$ equipped with the inner product

$$L_2(\mathcal{N}) \times L_2(\mathcal{N}) \ni (S, T) \mapsto \langle S, T \rangle := \int T^* S \in \mathbb{C}$$

is a Hilbert space that is unitary isomorphic with the Haagerup representation Hilbert space \mathcal{H}_H [56], and is unitary isomorphic with the GNS Hilbert space \mathcal{H}_ϕ for any choice of $\phi \in \mathcal{N}_{*0}^+$. The corresponding 1/2-deviation takes the form

$$D_{1/2}(x, y) = \frac{1}{2} \|x - y\|_{\mathcal{H}_H}^2.$$

The $D_{1/2}$ -projection is a minimiser of the Hilbert space norm, and defines a unique corresponding orthogonal projection operator. Hence, it is possible to recover the Hilbert space geometry as a special case of quantum information geometry.

In [88] we will present the derivation of temporal behaviour of ordinary Hilbert space based quantum theoretic models from the relative entropic information dynamics on

$\mathcal{M}(\mathcal{N})$. The brief idea of this recovery is following. The entropic updating generates non-linear mappings of Hilbert space vectors or density operators (that is often called a “reduction due to measurement”). In particular case, it reduces to the von Neumann–Lüders rule. However, in addition, to each stage of the non-unitary entropic evolution, there is associated a unique unitary evolution, specified by the Tomita–Takesaki modular automorphism, which generates the unitary evolution of vectors that represent the quantum information states of $\mathcal{M}(\mathcal{N})$ within \mathcal{H}_H . Because the Tomita–Takesaki automorphism is completely determined by underlying quantum information state, the entropic updating leads also to recovery of perturbations of generators of unitary dynamics that are determined by the constraints of dynamics (that is often called a “perturbation due to interaction”).

4 Some remarks on interpretation

We will finish this paper with a few remarks on the interpretation that we associate with the above formalism. These remarks are not intended to be comprehensive. Their role is only to indicate some of the key ideas. The extensive discussion of the conceptual issues related with our approach will be provided in [88]. See also [87] for a further discussion of information semantics and intersubjective interpretation in a context of foundations of probability theory and statistical inference.

4.1 Inductive inference and information semantics

In general, by *inductive inference* we understand some form of inductive logical reasoning, as opposed to deductive logical reasoning. The latter specifies premises by the valuations of sentences in *truth* values, and provides an inference procedure which is considered to lead to *certain* conclusion on the base of given premises. The former specifies premises by the valuations of sentences in *possible (plausible)* values and provides an inference procedure which is considered to lead to *most possible (most plausible)* conclusions on the base of given premises. From the mathematical perspective, the difference between deductive and inductive inference lays not in the form of logical valuations (these can be the same in both methods), but in the procedure of specifying conclusions on the base of premises. The conclusions of multiple application of deductive inference to the sequence of sets of premises depend *in principle* on all elements of all these sets, while the conclusions of the multiple application of inductive inference to the sequence of sets of premises depend *in principle* only on some elements of some of these sets. By this reason, the premises of inductive inference are also called *evidence*. An example of inductive inference procedure is any statistical reasoning based on probabilities. The evidence (called also ‘constraints of inference’) can consist, for example, of particular quantities with units called ‘experimental data’ *together with* a particular choice of a method which incorporates these ‘data’ into statistical inference. Any choice of such method *defines* the actual meaning of the ‘data’, and is a crucial element of the inference procedure. A standard example of such method is to ignore everything what is known about a sequence of numbers from a series of repeated measurements associated with a single abstract quality (such as a “position”), leaving only the value of arithmetic average and the value of a fluctuation around this average as a subject of comparison (e.g., by identification) with the mean and variance parameters of the gaussian probabilistic model.

Let us now consider the mathematical framework briefly presented in previous sections. On the level of *information semantics*, the underlying algebra (\mathcal{U} or \mathcal{N}) represents an *abstract qualitative language* subjected to quantitative evaluation, the spaces ($\mathcal{M}(\mathcal{U})$ or $\mathcal{M}(\mathcal{N})$) of finite integrals and their geometry (kinematic models of information) represent *quantified knowledge*, while the entropic updating (dynamics of information models) represents *quantitative inductive inference*.

This means that the algebra \mathcal{N} is *not* considered as consisting of ‘observables’ or ‘elementary events’. Due to its kinematical character, the quantum information model $\mathcal{M}(\mathcal{N})$ and its particular geometry is considered as a model of *potential* knowledge. In this sense, it would be much more relevant to use the word ‘observable quantity’ in order to refer to elements of $\mathcal{M}(\mathcal{N})$ than to the elements of \mathcal{N} . (This is especially appealing when one notices that orthodox quantum theoretical approach identification of self-adjoint elements of non-commutative algebra with ‘observables’ is applied only to *concrete* algebras—such that act on the Hilbert spaces—and not to the abstract algebras. However, the passage from abstract to concrete algebra, $\pi_\omega : \mathcal{N} \rightarrow \mathfrak{B}(\mathcal{H}_\omega)$ involves the dependence on $\omega \in \mathcal{M}(\mathcal{N})$.) Suggestions of terminological shift of this type are present e.g. in [83, 98], but such change would create an unnecessary semantic confusion, hence we prefer to simply resign from using the term ‘observable’.

The quantitative information dynamics of the model $\mathcal{M}(\mathcal{N})$ is formed by the additional choices of D_γ , E and F . On the semantic level, the functions E and F specify the *evidence* subjected to the inductive inference rule provided by entropic information dynamics (12). The resulting projection $\mathfrak{P}_F^{D_\gamma, E}$ is an inference: specification of most plausible state of knowledge subjected to given evidence. For a ‘temporal history’ $F = F(t)$ and an ‘initial state’ ω specified by $E(\varphi, \omega) = d\varphi\delta(\varphi - \omega)$, the information dynamics (12) takes a form of temporal evolution of quantum states $\omega(t) := \mathfrak{P}_{F(t)}^{D_\gamma}(\omega_0)$. It models the changes of the *actual* knowledge determined by the changes of what is considered to be an *actual* evidence.

This way quantum theory becomes a theory of quantitative inductive inference (information dynamics) provided with respect to the qualitative abstract language of non-commutative algebras. The constraints of this inductive logic are imposed by the choice of particular $\mathcal{M}(\mathcal{N})$, D_γ , E and F .

4.2 Intersubjective interpretation and farewell to ontology

The above information *semantics* requires an additional *interpretation* which would determine the particular operational and conceptual meaning attributed to the terms ‘knowledge’ and ‘change of knowledge’. In particular, this interpretation should determine the choice of a particular information kinematics (that is, $\mathcal{M}(\mathcal{N})$ and its information geometry) and a particular information dynamics (D_γ , E , F) when applied to some particular experimental situations.

In order to deal with the conceptual problems of quantum theory and to bypass the conflict between ‘objective’ and ‘personalistic’ bayesian interpretations of probability and statistical inference theory, we propose new *intersubjective interpretation*. According to it, the knowledge used to define particular theoretical model should bijectively correspond to the knowledge that is sufficient and necessary in order to *intersubjectively reproduce* corresponding ‘experiment of a given type’ that is intended to be described by this theoretical model. An experiment of a given type consists of an experimental setup of a given

type and its particular use, which amounts to setting a particular configuration (of controlled actual inputs) and a particular registration scale of potential outcomes (allowed results of use). An ‘experiment of a given type’ is an idealistic abstraction. However, this abstraction needs not to be understood in ontological sense. We consider it as a purely epistemic entity. The crucial question is: how to verify whether some particular individual setup under consideration and some particular actions and observations associated with it can be considered as an *intersubjectively valid* instance of an experiment of a given type? The answer is: an agreement with some particular knowledge has to be positively verified in operational terms. Hence, this knowledge actually *defines* an intersubjective notion of an experiment of a given type. The intersubjective interpretation amounts to require the bijective agreement between the kinematics of theoretical model and the terms of experimental setup construction, as well as the bijective agreement between the dynamics of theoretical model and the terms of use of experimental setup. We postulate bijection and not identification because we allow complete separation between the theoretical abstract language used to intersubjectively define and communicate theoretical models, and the operational language used to intersubjectively define and communicate corresponding experiments. In consequence, the intersubjective interpretation does not define the absolute (passive, static) meaning of the notion of ‘knowledge’. It defines only the relational (active, dynamic) meaning of this notion, as a particular relationship between kinematics-and-dynamics of theoretical model and construction-and-use of experimental setup.⁵

This interpretation introduces a key property expected from any scientific theory directly into conceptual foundations of quantum theory: a requirement that theory should allow unambiguous intersubjective verification in an experiment of a particular type. We require that every quantum theoretic model has to be *defined* in terms of some particular experimental design that can be used to intersubjectively verify the predictions of this theoretical model.

Quite similar postulate was already proposed in the context of semi-spectral approach to foundations of quantum theory [51, 16, 17, 38]. However, the semi-spectral approach is involved in the mathematical frameworks of Hilbert spaces and measure spaces which play no foundational role in our approach. The basic elements of our new mathematical foundations for quantum theory are interpreted as follows. An algebra \mathcal{N} is understood as an abstract qualitative language used as a common reference in an intersubjective communication about the abstract (idealised, theoretical, intentional) qualities that are subjected to quantification (quantitative evaluation, integration) in the course of the use of experimental setups. The quantum information model $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_*^+$ and its geometry is understood as the carrier of quantitative intersubjective knowledge describing the particular experimental setup under consideration. The choices of evidence E and F provides the description of particular control settings and particular range of allowed response outcomes. Together with the choice of D_γ , these choices determine intersubjectively accepted range (scale) of potential of quantitative results of use (outcomes) of experimental setup that is modelled in terms of $\mathcal{M}(\mathcal{N})$. As a consequence, the temporal information dynamics $\mathfrak{P}_{F(t)}^{D_\gamma}(\omega_0)$ provides the time-dependent description of most

⁵By the same reason (intersubjective context dependent bijective correspondence as opposed to absolutist identification), intersubjective interpretation cannot be considered as operationalism. On the other hand, it provides no ontological claims. Thus, it bypasses the naïvetés of conflict between ‘realism’ and ‘operationalism’.

plausible response outcomes that can be inferred from the given evidence.

We do not assume any *a priori* division of the notion of experimental setup and its use into “measuring device” and “measured system”. The *mathematical notion of a “measured system”* can be introduced as a special property (decomposition into tensor product) imposed on the algebra, quantum state space and quantum dynamics, providing this way the backwards compatibility of mathematical formalisms and their predictive validity. This way we can reintroduce suitable idealistic concept of a “measured system” for *some* experimental situations that allow it. Yet, this concept obtains here purely epistemic meaning, and it has no foundational role. Moreover, it might be just non-applicable in construction of theoretical models that correspond to *some* experimental situations. From the perspective provided by our approach, the *idealistic concept of a “measured system”* understood as kantian *Ding an sich* is completely irrelevant both for foundations and applications of quantum theory. Thus, the ontological interpretations of quantum theoretic formalism, as well as the internal problems of these interpretations (that are usually introduced by the use of such operationally undefined idealistic terms as ‘matter’, ‘fields’, ‘particles’, ‘quantum objects’, ‘randomness’, ‘universe’, ‘nature’), are also completely irrelevant for the foundations of quantum theory.

Acknowledgments

I am grateful to Chris Isham and Stanisław L. Woronowicz for very valuable discussions, to Jerzy Lewandowski for the support of my research, to Ariel Caticha, Paolo Gibilisco, Frank Hellmann, Ewa Infeld, Wojtek Kamiński and Włodek Natorf for their comments and suggestions on the early version of this paper, and to Fotini Markopoulou, Olaf Dreyer, Jamie Vicary, Cecilia Flori, Ewa Infeld and Tomasz Kołodziejski for helping with accomodation during my scientific visits. This work was partially supported by the FNP *Mistrz* grant 2007, ESF *Quantum Geometry and Quantum Gravity* short visit grant 1955, and ESF *Quantum Geometry and Quantum Gravity* exchange grant 2706.

References

- [1] Ali S.M., Silvey S.D., 1966, *A general class of coefficients of divergence of one distribution from another*, J. Roy. Stat. Soc. B **28**, 131.
- [2] Amari S.-i., 1985, *Differential-geometrical methods in statistics*, Lecture Notes in Statistics **28**, Springer, Berlin.
- [3] Amari S.-i., 2009, *α -divergence is unique, belonging to both f -divergence and Bregman divergence classes*, IEEE Trans. Inf. Theor. **55**, 4925.
- [4] Amari S.-i., Nagaoka H., 1993, *Joho kika no hoho*, Iwanami Shoten, Tōkyō (engl. transl. rev. ed.: 2000, *Methods of information geometry*, American Mathematical Society, Providence).
- [5] Araki H., 1976, *Relative entropy for states of von Neumann algebras I*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **11**, 809. Available at: www.ems-ph.org/journals/show_pdf.php?issn=0034-5318&vol=11&iss=3&rank=9.
- [6] Araki H., 1977, *Relative entropy for states of von Neumann algebras II*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **13**, 173. Available at: www.ems-ph.org/journals/show_pdf.php?issn=0034-5318&vol=13&iss=1&rank=8.
- [7] Araki H., 1993, *Ryoshiha no suri*, Iwanami Shoten, Tōkyō (engl. transl. 1999, *Mathematical theory of quantum fields*, Int. Ser. Monogr. Phys. **101**, Oxford University Press, Oxford).
- [8] Banerjee A., Guo X., Wang H., 2005, *On the optimality of conditional expectation as a Bregman predictor*, IEEE Trans. Inf. Theor. **51**, 2664.
- [9] Bauer H., 1958, *Minimalstellen von Funktionen und Extremalpunkte*, Arch. Math. **9**, 389.
- [10] Baumgärtel H., 1995, *Operatoralgebraic methods in quantum field theory*, Akademie Verlag, Berlin.
- [11] Bauschke H.H., Borwein J.M., 1997, *Legendre functions and the method of random Bregman projections*, J. Conv. Anal. **4**, 27. Available at: www.emis.de/journals/JCA/vol.4_no.1/j86.ps.gz.
- [12] Bayes T., 1763, *An essay towards solving a problem in the doctrine of chances*, Phil. Trans. Roy. Soc. London **53**, 370 (reprinted in: 1958, *Biometrika* **45**, 293).
- [13] Borel É., 1909, *Les probabilités dénomerables et leurs applications arithmétiques*, Rend. Circ. Math. Palermo **27**, 247.
- [14] Bratelli O., Robinson D.W., 1979, 1981, *Operator algebras and quantum statistical mechanics*, Vol.1-2, Springer, Berlin.
- [15] Bregman L.M., 1967, *Relaksatsionnyĭ metod nakhozheniya obshcheĭ točki vypuklykh mnozhestv i ego primeneniya dlia zadach vypuklogo programmirovaniya*, Zh. vychestl. matem. matem. fiz. **7**, 620. (engl. transl.: 1967, *The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming*, USSR Comput. Math. Math. Phys. **7**, 200.).

- [16] Busch P., Grabowski M., Lahti P.J., 1989, *Some remarks on effects, operations, and unsharp measurements*, Found. Phys. Lett. **2**, 331.
- [17] Busch P., Grabowski M., Lahti P.J., 1995, *Operational quantum physics*, LNPM **31**, Springer, Berlin. (1997, 2nd corr. print.).
- [18] Caticha A., 1998, *Consistency, amplitudes and probabilities in quantum theory*, Phys. Rev. A **57**, 1572. [arXiv:quant-ph/9804012](https://arxiv.org/abs/quant-ph/9804012).
- [19] Caticha A., 1998, *Consistency and linearity in quantum theory*, Phys. Lett. A **244**, 13. [arXiv:quant-ph/9803086](https://arxiv.org/abs/quant-ph/9803086).
- [20] Caticha A., 2000, *Insufficient reason and entropy in quantum theory*, Found. Phys. **30**, 227. [arXiv:quant-ph/9810074](https://arxiv.org/abs/quant-ph/9810074).
- [21] Caticha A., Giffin A., 2006, *Updating probabilities*, in: Mohammad-Djafari A. (ed.), *Bayesian inference and maximum entropy methods in science and engineering*, AIP Conf. Proc. **872**, 31. [arXiv:physics/0608185](https://arxiv.org/abs/physics/0608185).
- [22] Caves C.M., Fuchs C.A., Schack R., 2001, *Quantum probabilities as bayesian probabilities*, Phys. Rev. A **65**, 022305. [arXiv:quant-ph/0106133](https://arxiv.org/abs/quant-ph/0106133).
- [23] Chentsov N.N., 1968, *Nesimmetrichnoye rasstoyanie mezhdu raspredeleniyami veroyatnostey, entropiya i teoriya Pifagora*, Mat. Zametki **4**, 323. (engl. transl. 1968, *Nonsymmetrical distance between probability distributions, entropy and the theorem of Pythagoras*, Math. Notes **4**, 686.).
- [24] Chentsov N.N., 1972, *Statisticheskie reshayushchie pravila i optimal'nye vyvody*, Nauka, Moskva (engl. transl.: 1982, *Statistical decision rules and optimal inference*, American Mathematical Society, Providence).
- [25] Connes A., 1973, *Une classification des facteurs de type III*, Ann. Sci. École Norm. Sup. 4 ème sér. **6**, 133. [numdam:ASENS-1973_4_6_2_133_0](https://arxiv.org/abs/math/9804012).
- [26] Cox R.T., 1946, *Probability, frequency and reasonable expectation*, Am. J. Phys. **14**, 1.
- [27] Cox R.T., 1961, *The algebra of probable inference*, John Hopkins University Press, Baltimore.
- [28] Csizsár I., 1963, *Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **8**, 85.
- [29] Csizsár I., 1975, *I-divergence geometry of probability distributions and minimization problems*, Ann. Prob. **3**, 146.
- [30] Csizsár I., 1978, *Information measures: a critical survey*, in: Kozesnik J. (ed.), *Transactions of the seventh Prague conference on information theory, statistical decision functions and random processes*, p.73.
- [31] Csizsár I., 1991, *Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems*, Ann. Stat. **19**, 2032.
- [32] Daniell P.J., 1918, *A general form of the integral*, Ann. Math. **19**, 274.
- [33] Daniell P.J., 1919, *Integrals in an infinite number of dimensions*, Ann. Math. **21**, 30.
- [34] Daniell P.J., 1920, *Further properties of the general integral*, Ann. Math **21**, 203.
- [35] de Finetti B., 1931, *Sul significato suggestivo della probabilità*, Mathematicae **17**, 298.
- [36] de Finetti B., 1970, *Teoria delle probabilità*, Einaudi, Torino (engl. transl.: 1974, *Theory of probability*, Vol.1-2, Wiley, New York).
- [37] de Laplace P.-S., 1812, *Théorie analytique des probabilités*, Courcier, Paris.
- [38] de Muynck W.M., 2002, *Foundations of quantum mechanics: an empiricist approach*, Kluwer, Dordrecht.
- [39] Eguchi S., 1983, *Second order efficiency of minimum contrast estimators in a curved exponential family*, Ann. Statist. **11**, 793. [euclid:1176346246](https://arxiv.org/abs/math/9804012).
- [40] Eguchi S., 1985, *A differential geometric approach to statistical inference on the basis of contrast functionals*, Hiroshima Math. J. **15**, 341. [euclid:1206130775](https://arxiv.org/abs/math/9804012).
- [41] Emch G.G., 1972, *Algebraic methods in statistical mechanics and quantum field theory*, Wiley, New York.
- [42] Emch G.G., 1998, *On the need to adapt de Finetti's probability interpretation to QM*, Banach Center Publ. **43**, 157.
- [43] Falcone T., Takesaki M., 2001, *The non-commutative flow of weights on a von Neumann algebra*, J. Funct. Anal. **182**, 170. Available at: www.math.ucla.edu/~mt/papers/QFlow-Final.tex.pdf.
- [44] Fremlin D.H., 2000, 2001, 2002, 2003, *Measure theory*, Vol.1-4, Torres Fremlin, Colchester.
- [45] Fuchs C.A., 2003, *Quantum mechanics as quantum information, mostly*, J. Mod. Opt. **50**, 987. [arXiv:quant-ph/0205039](https://arxiv.org/abs/quant-ph/0205039).
- [46] Fuchs C.A., 2010, *Qbism, the perimeter of quantum bayesianism*, [arXiv:1003.5209](https://arxiv.org/abs/1003.5209).
- [47] 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. Topics **2**, 169. Available at: www.mat.uniroma2.it/~isola/research/preprints/Gils01.pdf.
- [48] Gibilisco P., Pistone G., 1998, *Connections on non-parametric statistical manifolds by Orlicz space geometry*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **1**, 325.
- [49] Giffin A., 2009, *Maximum entropy: the universal method for inference*, Ph.D. thesis, State University of New York, Albany. [arXiv:0901.2987](https://arxiv.org/abs/0901.2987).
- [50] Gleason A., 1957, *Measures on the closed subspaces of a Hilbert space*, J. Math. Mech. **6**, 885.
- [51] Grabowski M., 1989, *What is an observable?*, Fund. Phys. **19**, 923.
- [52] Grandy W.T., Jr., 1987, 1988, *Foundations of statistical mechanics*, Vol.1-2, Reidel, Dordrecht.
- [53] Grandy W.T., Jr., 2008, *Entropy and the time evolution of macroscopic systems*, Oxford University Press, Oxford.
- [54] Haag R., 1959, *Discussion des "axiomes" et des propriétés asymptotiques d'une théorie des champs locale avec particules composées*, in: ? (ed.), *Les problèmes mathématique de la théorie quantique des champs (Lille 1957)*, Centr. Natl. Rech. Sci., Paris..
- [55] Haag R., 1992, *Local quantum physics*, Springer, Berlin. (1996, 2nd rev. enlarg. ed.).
- [56] Haagerup U., 1975, *The standard form of von Neumann algebras*, Math. Scand. **37**, 271.
- [57] Hamhalter J., 2003, *Quantum measure theory*, Kluwer, Dordrecht.
- [58] Hasegawa H., 1993, *α -divergence of the non-commutative information geometry*, Rep. Math. Phys. **33**, 87.
- [59] Ingarden R.S., 1963, *Information theory and variational principles in statistical theories*, Bull. Acad. Polon. Sci. Sér. Math. Astr. Phys. **11**, 541.
- [60] Ingarden R.S., 1976, *Quantum information theory*, Rep. Math. Phys. **10**, 43.
- [61] Ingarden R.S., 1978, *Information thermodynamics and differential geometry*, Memoirs Sagami Inst. Techn. **12**, 83.
- [62] Ingarden R.S., Janyszek H., 1982, *On the local riemannian structure of the state space of classical information thermodynamics*, Tensor (N.S.) **39**, 279.

- [63] Ingarden R.S., Janyszek H., Kossakowski A., Kawaguchi T., 1982, *Information geometry of quantum statistical systems*, Tensor (N.S.) **37**, 105.
- [64] Ingarden R.S., Kossakowski A., Ohya M., 1997, *Information dynamics and open systems – classical and quantum approach*, Kluwer, Dordrecht.
- [65] Ingarden R.S., Urbanik K., 1961, *Information as a fundamental notion of statistical physics*, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astronom. Phys. **9**, 313.
- [66] Ingarden R.S., Urbanik K., 1962, *Information without probability*, Colloq. Math. **9**, 131.
- [67] Jaynes E.T., 1957, *Information theory and statistical mechanics*, Phys. Rev. **106**, 620.
- [68] Jaynes E.T., 1957, *Information theory and statistical mechanics II*, Phys. Rev. **108**, 171.
- [69] Jaynes E.T., 1967, *Foundations of probability theory and statistical mechanics*, in: Bunge M. (ed.), *Delaware seminars in the foundations of science*, Vol.1, Springer, Berlin, p.77.
- [70] Jaynes E.T., 1979, *Where do we stand on maximum entropy?*, in: Levine R.D., Tribus M. (eds.), *The maximum entropy formalism*, MIT Press, Cambridge, p.15.
- [71] Jaynes E.T., 1985, *Generalized scattering*, in: Smith C.R., Grandy W.T., Jr. (eds.), *Maximum-entropy and bayesian methods in inverse problems*, Reidel, Dordrecht, p.377.
- [72] Jaynes E.T., 1985, *Macroscopic prediction*, in: Haken H. (ed.), *Complex systems – operational approaches in neurobiology, physics and computers*, Springer, Berlin, p.254.
- [73] Jaynes E.T., 1986, *Predictive statistical mechanics*, in: Moore G.T., Scully M.O. (eds.), *Frontiers of nonequilibrium statistical mechanics*, Plenum, New York, p.33.
- [74] Jaynes E.T., 1990, *Probability in quantum theory*, in: Żurek W.H. (ed.), *Complexity, entropy and the physics of information*, Addison Wesley, Reading.
- [75] Jaynes E.T., 1993, *Inferential scattering*, unpublished preprint, available at bayes.wustl.edu/etj/articles/cinfscat.pdf.
- [76] Jaynes E.T., 2003, *Probability theory: the logic of science*, Cambridge University Press, Cambridge.
- [77] Jaynes E.T., Scalapino D.J., 1963, *Irreversible statistical mechanics*, unpublished preprint, available at: bayes.wustl.edu/etj/articles/irreversible.stat.mac.pdf.
- [78] Jenčová A., 2003, *Affine connections, duality and divergences for a von Neumann algebra*, [arXiv:math-ph/0311004](https://arxiv.org/abs/math-ph/0311004).
- [79] Jenčová A., 2005, *Quantum information geometry and non-commutative L_p spaces*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **8**, 215. Available at: www.mat.savba.sk/~jencova/lpspaces.pdf.
- [80] Jenčová A., 2006, *A construction of a nonparametric quantum information manifold*, J. Funct. Anal. **239**, 1. [arXiv:math-ph/0511065](https://arxiv.org/abs/math-ph/0511065).
- [81] Jenčová A., 2010, *On quantum information manifolds*, in: Gibilisco P. et al (eds.), *Algebraic and geometric methods in statistics*, Cambridge University Press, Cambridge.
- [82] Kalashnikov V.P., Zubarev D.N., 1972, *On the extremal properties of the nonequilibrium statistical operator*, Physica **59**, 314.
- [83] Klauder J.R., 1997, *Coherent states in action*, in: Lim S.C. et al (eds.), *Frontiers in quantum physics*, Springer, Berlin, p.151. [arXiv:quant-ph/9710029](https://arxiv.org/abs/quant-ph/9710029).
- [84] Kolmogorov A.N., 1933, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Ergebnisse der Mathematik und Ihrer Grenzgebiete Bd. **2**, Springer, Berlin. (engl. transl. 1950, *Foundations of the theory of probability*, Chelsea, New York).
- [85] Kostecki R.P., 2011, *The general form of γ -family of quantum relative entropies*, Open Sys. Inf. Dyn. **18**, 191. [arXiv:1106.2225](https://arxiv.org/abs/1106.2225).
- [86] Kostecki R.P., 2011, *On general form of quantum information geometry*, to be submitted.
- [87] Kostecki R.P., 2011, *On principles of inductive inference*, submitted to: *Proceedings of 31st International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 10-15 July 2011, Waterloo*.
- [88] Kostecki R.P., 2011, *Towards new foundations of quantum theory I, II, III*, in preparation.
- [89] Le Cam L., 1964, *Sufficiency and approximate sufficiency*, Ann. Math. Statist. **35**, 1419.
- [90] Le Cam L., 1986, *Asymptotic methods in statistical decision theory*, Springer, Berlin.
- [91] Loomis L.H., 1947, *On the representation of σ -complete boolean algebras*, Bull. Amer. Math. Soc. **53**, 757.
- [92] Maeda S., 1990, *Probability measures on projections in von Neumann algebras*, Rev. Math. Phys. **1**, 235..
- [93] Mitchell W.C., 1967, *Statistical mechanics of thermally driven systems*, Ph.D. thesis, Washington University, St.Louis.
- [94] Morimoto T., 1963, *Markov processes and the H-theorem*, J. Phys. Soc. Jap. **12**, 328.
- [95] Morozov V.G., Röpke G., 1998, *Zubarev's method of a nonequilibrium statistical operator and some challenges in the theory of irreversible processes*, Cond. Matt. Phys. **16**, 673.
- [96] Morozova E.A., Chentsov N.N., 1991, *Estestvennaya geometriya semeïstv veroyatnostnykh zakonov*, Itogi Nauki i Tekhniki, Ser. Sovr. Probl. Mat. Fundam. Napravlenn. **83**, 133. (*The natural geometry of families of probabilistic laws*) Available at: www.mathnet.ru/links/0b7482826af98f74f707088ff33082f6/intf210.pdf.
- [97] Nagaoka H., Amari S.-i., 1982, *Differential geometry of smooth families of probability distributions*, Technical report METR **82-7**, University of Tōkyō, Tōkyō.
- [98] Neumaier A., 2003, *Ensembles and experiments in classical and quantum physics*, Int. J. Mod. Phys. B **17**, 2937. [arXiv:quant-ph/0303047](https://arxiv.org/abs/quant-ph/0303047).
- [99] Ohya M., Petz D., 1993, *Quantum entropy and its use*, Springer, Berlin.
- [100] Ojima I., 2004, *Temperature as order parameter of broken scale invariance*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **40**, 731. [arXiv:math-ph/0311025](https://arxiv.org/abs/math-ph/0311025).
- [101] Petz D., 1985, *Quasi-entropies for states of a von Neumann algebra*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **21**, 787. Available at: www.math.bme.hu/~petz/pdf/26quasi.pdf.
- [102] Petz D., 1986, *Properties of the relative entropy of states of von Neumann algebra*, Acta Math. Hungar. **47**, 65.
- [103] Petz D., 2007, *Bregman divergence as relative operator entropy*, Acta Math. Hungar. **116**, 127. Available at: www.renyi.hu/~petz/pdf/112bregman.pdf.
- [104] Piron C., 1976, *Foundations of quantum physics*, Benjamin, New York.
- [105] Pistone G., Sempi C., 1995, *An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one*, Ann. Statist. **23**, 1543.
- [106] Pitowsky I., 2006, *Quantum mechanics as a theory of probability*, in: Demopoulos W., Pitowsky I. (eds.), *Physical*

- theory and its interpretation: essays in honor of Jeffrey Bub*, Springer, Berlin, p.213. [arXiv:quant-ph/0510095](https://arxiv.org/abs/quant-ph/0510095).
- [107] Ramsey F.P., 1931, *The foundations of mathematics and other logical essays*, Routledge and Kegan, London.
- [108] Randall C.H., Foulis D.J., 1983, *Properties and operational propositions in quantum mechanics*, Found. Phys. **13**, 843.
- [109] Rédei M., Summers S.J., 2007, *Quantum probability theory*, Stud. Hist. Phil. Mod. Phys. **38**, 390. [arXiv:quant-ph/0601158](https://arxiv.org/abs/quant-ph/0601158).
- [110] Segal I.E., 1947, *Postulates for general quantum mechanics*, Ann. Math. **48**, 930.
- [111] Segal I.E., 1953, *A non-commutative extension of abstract integration*, Ann. Math. **57**, 401.
- [112] Shannon C., 1948, *A mathematical theory of communication*, Bell Syst. Tech. J. **27**, 379, 623.
- [113] Sikorski R., 1948, *On the representation of boolean algebras as fields of sets*, Fund. Math. **35**, 247.
- [114] Stone M.H., 1948, *Notes on integration I, II, III*, Proc. Natl. Acad. Sci. U.S.A. **34**, 336, 447, 483.
- [115] Stone M.H., 1949, *Notes on integration IV*, Proc. Natl. Acad. Sci. U.S.A. **35**, 50.
- [116] Umegaki H., 1962, *Conditional expectations in an operator algebra IV (entropy and information)*, Kōdai Math. Sem. Rep. **14**, 59.
- [117] Whittle P., 1970, *Probability*, Penguin, Harmondsworth.
- [118] Whittle P., 2000, *Probability via expectation*, (4th ed.), Springer, Berlin.
- [119] Wigner E., Yanase M., 1963, *Information content of distributions*, Proc. Nat. Acad. Sci. U.S.A. **49**, 910.
- [120] Wilce A., 2000, *Test spaces and orthoalgebras*, in: Coecke B., Moore D., Wilce A. (eds.), *Current research in operational quantum logic*, Kluwer, Dordrecht, p.?.
- [121] Youssef S., 1991, *A reformulation of quantum mechanics*, Mod. Phys. Lett. A **6**, 225.
- [122] Youssef S., 1994, *Quantum mechanics as complex probability theory*, Mod. Phys. Lett. A **9**, 2571.
- [123] Zhang J., 2004, *Divergence function, duality, and convex analysis*, Neural Comp. **16**, 159. Available at: www.lsa.umich.edu/psych/junz/Neural%20Comp%202004.pdf.
- [124] Zhu H., 1998, *Generalized Lebesgue spaces and application to statistics*, Santa Fe Inst. Tech. Rep. **98-06-44**, Santa Fe. Available at: www.santafe.edu/media/workingpapers/98-06-044.ps.
- [125] Zhu H., Rohwer R., 1997, *Measurements of generalisation based on information geometry*, in: Ellacott S.W. et al. (eds.), *Mathematics of neural networks: models, algorithms and applications*, Kluwer, Dordrecht, p.394. Available at: eprints.aston.ac.uk/514/1/NCRG_95_012.pdf.
- [126] Zhu H., 1998, *Information geometry, bayesian inference, ideal estimates and error decomposition*, Santa Fe Inst. Tech. Rep. **98-06-45**, Santa Fe. Available at: omega.albany.edu:8008/ignorance/zhu98.pdf.
- [127] Zubarev D.N., 1961, [*Statisticheskii operator dla neravnovesnykh sistem*], Dokl. Akad. Nauk SSSR **140**, 92. (engl. transl. 1962, *The statistical operator for nonequilibrium systems*, Sov. Phys. Dokl. **6**, 776).
- [128] Zubarev D.N., 1994, *Nonequilibrium statistical operator as a generalization of Gibbs distribution for nonequilibrium case*, Cond. Matt. Phys. **4**, 7.
- [129] Zubarev D.N., Kalashnikov V.P., 1969, *Ekstremalnye svoistva neravnovestnogo statisticheskogo operatora*, Teor. Mat. Fiz. **1**, 137. (*Extremal properties of the nonequilibrium statistical operator*).
- [130] Zubarev D.N., Kalashnikov V.P., 1970, *Derivation of the nonequilibrium statistical operator from the extremum of the information entropy*, Physica **46**, 550.
- [131] Zubarev D.N., Kalashnikov V.P., 1970, *Postroyenne statisticheskikh operatorov dla neravnovesnykh processov*, Teor. Mat. Fiz. **3**, 126. (*Construction of statistical operators for nonequilibrium processes*).
- [132] Zubarev D.N., Morozov V.G., Röpke G., 1996, *Statistical mechanics of nonequilibrium processes*, Vol.1-2, Akademie Verlag, Berlin. (2nd rev. enlarg. ed., 2002, *Statisticheskaya mekhanika neravnovesnykh processov*, Vol.1-2, Fizmatlit, Moskva.).