Lüders' and quantum Jeffrey's rules as entropic projections

Ryszard Paweł Kostecki

Perimeter Institute for Theoretical Physics, 31 Caroline Street North, N2L 2Y5 Waterloo, Ontario, Canada[*]
Institute for Theoretical Physics, Department of Physics, University of Warsaw, Hoża 69, 00-681 Warszawa, Poland

ryszard.kostecki@fuw.edu.pl

August 15, 2014

Abstract

We prove that the standard quantum mechanical description of a quantum state change due to measurement, given by Lüders' rules, is a special case of the constrained maximisation of a quantum relative entropy functional. This result is a quantum analogue of the derivation of the Bayes–Laplace rule as a special case of the constrained maximisation of relative entropy. The proof is provided for the Umegaki relative entropy of density operators over a Hilbert space as well as for the Araki relative entropy of normal states over a W^*-algebra. We also introduce a quantum analogue of Jeffrey's rule, derive it in the same way as above, and discuss the meaning of these results for quantum bayesianism.

1 Introduction

An important part of a mathematical setting of probability theory can be considered as a special case of quantum theoretic kinematics. This can be seen most clearly when quantum theoretic kinematics is reformulated in algebraic terms, with quantum states defined as normal positive (or normalised) functionals over W^*-algebras (see e.g. [107, 79] for an overview). The Borel–Steinhaus–Kolmogorov measure theoretic approach to foundations of probability theory is then recovered precisely from commutative W^*-algebras and quantum states over them.[1] However, both quantum theory and probability theory are equipped with additional structures, which describe possible mappings of probabilities or quantum states. Can these prescriptions of information dynamics also be directly related to each other? In particular, there are various considerations [16, 17, 18, 106, 24, 109, 48, 49, 67, 117, 112, 19, 101, 63] of von Neumann's and Lüders' rules [120, 84] as noncommutative *analogues* of the Bayes–Laplace rule [11, 29, 31]. It is tempting to ask whether this analogy could be turned to something more definite, both conceptually and mathematically.

The main motivation for this paper is a series of results [96, 94, 95, 123, 33, 38, 126, 122, 23, 51, 43] showing that both Bayes–Laplace and Jeffrey's rules (as well as some other rules [39, 66, 43]) can be derived as special cases of the constrained minimisation of various information distances on probabilistic models (in Section 2 we review briefly some of these results). This has led us to conjecture [78] that Lüders' rules may be special cases of constrained maximisation of a quantum relative entropy. In Section 4 we prove that the weak Lüders rule of quantum state change due to "nonselective quantum measurement" is a special case of quantum entropic projection, provided by the minimisation of the Araki distance subject to a specific set of constraints. We also introduce a quantum analogue of Jeffrey's rule and derive it as another special case of constrained Araki distance minimisation. These are the two main results of this paper. In addition, we show that the strong Lüders rule of quantum state change due to "selective quantum measurement" can be obtained from these results as the limiting case of quantum Jeffrey's rule

---

[*]Current affiliation.

[1]The same holds for the Kappos' approach to foundations based on Caratheodory's measure theory on abstract boolean algebras, as well for Le Cam's approach based on abstract Banach lattices, if the latter is restricted to the 'coherent' models. In general, the semi-finite measures are recovered not from quantum states, but from semi-finite normal weights (see Section 3).

or by regularised Araki distance minimisation. In Section 5 we show that both weak Lüders' rule and strong von Neumann's rule (which is the same as strong Lüders' rule for pure states) are also quantum entropic projections for a different quantum distance functional. With an exception of a derivation of a quantum Jeffrey's rule, our results hold for arbitrary W^*-algebras, so they are applicable in quantum field theoretic and relativistic quantum information problems.

These results extend earlier considerations of Lüders' rules as *analogues* of the Bayes–Laplace rule with a novel mathematical and conceptual content: all these rules are special cases of the constrained relative entropic inference. In this sense, 'quantum bayesianism' can be considered a branch of 'quantum relative entropism'. A discussion of this issue is provided in Section 6.

History of the problem. An inference based on minimisation of quantum distance on quantum models was first proposed by Herbut [64]. He derived the weak Lüders rule from a constrained minimisation of norm distance in the Hilbert–Schmidt operator space. Unfortunately, his work has been left unnoticed by all works cited below. Several years later Marchand and collaborators [89, 14, 12, 13, 56, 54, 86, 87, 88] used Bures' distance [20] (which is metrical), and argued that its constrained minimisation should be considered as a description of the change of state of information *due to* "quantum measurement" described by a specific form of a coarse graining map [85] (conditioned upon a subset of an operator algebra, and predual to a specific form of noncommutative conditional expectation [55]). Independently of this body of work, in [22] it was proposed to use the constrained minimisation of the WGKL distance (3.34) of probabilities arising from traces of density operators to derive the post-measurement quantum state. The rules of inferential change of quantum states based on a constrained minimisation of other metrical distances on quantum models were later reconsidered by other authors [59, 60, 34, 105] and some derivations of the strong von Neumann rule were obtained (see Section 5). The reinterpretation of von Neumann's and Lüders' rules for "quantum measurement" as principles of inductive inference conditioned on specific information, and analogous to the Bayes–Laplace rule, was proposed at about the same time by Bub [16, 17, 18] (however, it can be claimed [112], that already von Neumann was aware of the possibility of such interpretation). These two lines of thought were (implicitly) joined in Hadjisavvas' [59, 60] postulate that a quantum state change due to acquisition of data (e.g. in a measurement) should be provided by means of constrained minimisation of the JMGK distance (5.18), as well as in Donald's [40, 41] postulate that a description of "quantum measurement", understood as an inductive inference, should be provided by means of constrained minimisation of the Araki distance.[2] Donald stressed also that this procedure «clearly allows for approximate measurement, and indeed they are required (all real measurements are approximate), if K is taken to have a non-empty interior» [41], where $K \subseteq \mathcal{S}(\mathcal{N})$ is the constraint set. However, he provided no derivation of any of Lüders' rules (nor any other "quantum measurement" rules) from this procedure. Independently of the above works, Warmuth [122] used constrained minimisation of Umegaki's distance to derive a generalisation of the Bayes–Laplace rule to the case of density operators, with conditional probabilities replaced by covariance matrices. This generalisation has not reproduced Lüders' rules.

Our derivation of Lüders' rule is not only the first such result obtained for Umegaki's and Araki's distances, but also first result of this type obtained for any nonsymmetric quantum information distance. All results for quantum Jeffrey's rule are new. See also a closely related paper [62], where the analogous results for weak and strong Lüders' rules are derived using another technique (based on differentiation, as opposed to generalised pythagorean theorem).[3]

---

[2]«Leaving aside any possible applications in the area of quantum communication theory, it seems to me that, almost regardless of the interpretation one places on quantum mechanics, this is an appropriate way of modelling the quantum measurement process» [41]. He states this refering to a constrained minimisation of a different distance functional, which however coincides with Araki's distance $D_1|_{\mathcal{N}_{*1}^+}$ at least for injective W^*-algebras $\mathcal{N}$.

[3]After finishing this paper, we were informed about reference [97], where it is shown that $\sigma = \sum_i P_i \rho_i P_i$,

2 Bayes–Laplace and Jeffrey's rules as entropic projections

The modern mathematical formulation of the Bayes–Laplace foundations for probability theory [11, 30] is based on finitely additive boolean algebras $\mathcal{A}$ and conditional probabilities, defined as maps $p(\cdot|\cdot) : \mathcal{A} \times \mathcal{A} \ni (x, y) \mapsto p(x|y) \in [0, 1]$. Its kinematics is given by the rules

$$p(x|y) + p(\neg x|y) = 1, \tag{2.1}$$

$$p(x|y \wedge z)p(y|z) = p(x \wedge y|z), \tag{2.2}$$

from which the **Bayes–Laplace theorem** [11, 29, 31] follows,

$$p(x|b \wedge \eta) = p(x|\eta)\frac{p(b|x \wedge \eta)}{p(b|\eta)}. \tag{2.3}$$

The 'marginal probability' $p(b|\eta)$, called also an 'evidence', is a normalising constant calculated from

$$p(b|\eta) = \sum_{i \in I} p(b|x_i \wedge \eta)p(x_i|\eta), \tag{2.4}$$

where I is a countable set, while the set $\{x_i \in \mathcal{A} \mid i \in I\}$ is exhaustive ($\bigvee_{i \in I} x_i = 1$) and its elements are mutually exclusive ($x_i \wedge x_j = 0$ for $i \neq j$). In the simple cases this set may consist of two elements: $\{x, \neg x\}$. The dynamics of this approach is given by the **Bayes–Laplace rule**

$$p(x|\eta) \mapsto p_{\mathrm{new}}(x|\eta) := p(x|\eta)\frac{p(b|x \wedge \eta)}{p(b|\eta)}. \tag{2.5}$$

The map (2.5) determines a rule of construction of the new conditional probability assignment associated to x under the constraint that certain additional statements ('facts', 'data', 'events') b are considered as ('appear as', 'are known as', 'are regarded as') true or false. So, $p_{\mathrm{new}}(x|\eta)$ in (2.5) is interpreted as a ('posterior') conditional probability assigned to x whenever the truth value of b is given ('known'). If attribution of a definite truth value to $b \in \mathcal{A}$ is interpreted as an 'acquisition of data/facts', then the Bayes–Laplace rule can be understood as a procedure of statistical inference that transforms 'prior' information states $p(x|\eta)$ about *all* 'hypotheses' $x \in \mathcal{A}$ into 'posterior' information states $p_{\mathrm{new}}(x|\eta)$, under constraints provided by the acquired data b *and* the assumed relationship between b and x which is encoded in the information states $p(b|x \wedge \eta)$. The probability $p(b|x \wedge \eta)$ is sometimes called the 'sampling probability' (because before the constraint b is applied, $p(b|x \wedge \eta)$ can represent a probability associated with possible values of constraints for a fixed hypothesis x) or the 'likelihood' (because after the constraint b is applied, $p(b|x \wedge \eta)$ is considered as a probability of b as a function over possible hypotheses x).

Jeffrey [70, 71, 72] proposed an alternative dynamical rule, now called **Jeffrey's rule**,

$$p(x|\eta) \mapsto p_{\mathrm{new}}(x|\eta) := \sum_{i=1}^{n} p(x|b_i)\lambda_i = \sum_{i=1}^{n} \frac{p(x \wedge b_i|\eta)}{p(b_i|\eta)}\lambda_i, \tag{2.6}$$

where $n \in \mathbb{N}$, $\{b_1, \ldots, b_n\}$ is a set of exhaustive and mutually exclusive elements of $\mathcal{A}$, and the constraints $\lambda_i = p_{\mathrm{new}}(b_i|\eta)$ $\forall i \in \{1, \ldots, n\}$ hold. The defining equality in (2.6) is equivalent to the condition

$$p_{\mathrm{new}}(x|b_i) = p(x|b_i) \quad \forall x \in \mathcal{A} \ \ \forall i \in \{1, \ldots, n\}. \tag{2.7}$$

If $n = 2$ with $b_2 = \neg b_1 =: \neg b$ and if $p_{\mathrm{new}}(\neg b|\eta) = 0$, then Jeffrey's rule (2.6) reduces to the Bayes–Laplace rule (2.5).

---

where P_i are rank 1 projectors, minimises the Umegaki distance $D_1|_{\mathfrak{B}(\mathcal{H})^+_{\star 1}}(\rho, \sigma)$. This is a special case of our result for the weak Lüders rule. The generalisation to our result is stated without proof in [25].

The first derivation of the Bayes–Laplace and Jeffrey's rules from constrained minimisation of the information distance functional (more precisely, the WGKL distance) was provided by Williams [123]. However, as he admitted, his derivation has a restricted validity, because it does not deal with conditional probabilities. This can be solved following van Fraassen's suggestion: «When two spaces are used (parameter or hypothesis space and observation or sample space) I shall think of these as subspaces of a larger one (possibly produced by a product construction), so that in a single context all propositions are represented by measurable sets in a single space» [118]. Following the results of Caticha and Giffin [23, 51], consider a constrained minimisation of the WGKL distance (3.34),

$$p(\chi, \theta) \mapsto p_{\text{new}}(\chi, \theta) := \underset{q(\chi, \theta) \in \mathcal{M}}{\arg \inf} \left\{ \int_{\mathcal{X}} \tilde{\mu}(\chi) q(\chi, \theta) \log \left(\frac{q(\chi, \theta)}{p(\chi, \theta)} \right) + F(q(\chi, \theta)) \right\}, \tag{2.8}$$

for $p, q \in \mathcal{M} := \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_1^+$, $\dim \mathcal{M} =: n < \infty$, with parametrisation $\theta : \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \to \Theta \subseteq \mathbb{R}^n$ allowing to consider a measure space $(\Theta, \mho_{\text{Borel}}(\Theta), d\theta)$ as well as a product measure space $(\mathcal{X} \times \Theta, \mho(\mathcal{X} \times \Theta), \tilde{\mu} \times d\theta)$, and with constraints given by

$$F(q(\chi, \theta)) = \lambda_1 \left(\int_{\mathcal{X}} \tilde{\mu}(\chi) \int_{\Theta} d\theta q(\chi, \theta) - 1 \right) + \lambda_2(\chi) \left(\int_{\Theta} d\theta q(\chi, \theta) - \boldsymbol{\delta}(\chi - b) \right), \tag{2.9}$$

where λ_1 and $\lambda_2(\chi)$ are Lagrange multipliers, and $\boldsymbol{\delta}(\chi - b)$ is Dirac's delta at $b \in \mathcal{X}$. The posterior probability selected as a unique solution of this variational problem is given by

$$p_{\text{new}}(\chi, \theta) = \frac{p(\chi, \theta) e^{\lambda_2(\chi)}}{\int_{\mathcal{X}} \tilde{\mu}(\chi) \int_{\Theta} d\theta p(\chi, \theta) e^{\lambda_2(\chi)}}, \tag{2.10}$$

where $\lambda_2(\chi)$ is determined via

$$\frac{\int_{\Theta} d\theta p(\chi, \theta) e^{\lambda_2(\chi)}}{\int_{\mathcal{X}} \tilde{\mu}(\chi) \int_{\Theta} d\theta p(\chi, \theta) e^{\lambda_2(\chi)}} = \boldsymbol{\delta}(\chi - b). \tag{2.11}$$

Hence,

$$p_{\text{new}}(\chi, \theta) = \frac{p(\chi, \theta) \boldsymbol{\delta}(\chi - b)}{\int_{\Theta} d\theta p(\chi, \theta)} = \frac{p(\chi, \theta) \boldsymbol{\delta}(\chi - b)}{p(\chi)} =: \boldsymbol{\delta}(\chi - b) p(\theta | \chi),$$

which leads to the Bayes–Laplace rule (2.5) on Θ,[4]

$$p(\theta) \mapsto p_{\text{new}}(\theta) = \int_{\mathcal{X}} \tilde{\mu}(\chi) \boldsymbol{\delta}(\chi - b) p(\theta | \chi) = p(\theta | b), \tag{2.12}$$

whenever μ is such that $\int_{\mathcal{X}} \tilde{\mu}(\chi) \boldsymbol{\delta}(\chi - b) h(\chi) = h(b)$ (for example, if $\tilde{\mu}(\chi) = d\chi$). If the second constraint in (2.9) is replaced by a more general form,

$$F(q(\chi, \theta)) = \lambda_1 \left(\int_{\mathcal{X}} \tilde{\mu}(\chi) \int_{\Theta} d\theta q(\chi, \theta) - 1 \right) + \lambda_2(\chi) \left(\int_{\Theta} d\theta q(\chi, \theta) - f(\chi) \right), \tag{2.13}$$

---

[4]More precisely, it leads to a generalisation of this rule to a domain of countably additive boolean algebras (representable by the Loomis–Sikorski theorem), which contains contains the finitary rule (2.5) as a special case. While there is no universal (generic) extension of the notion of conditional probability to a measure theoretic framework, $p(\theta | \chi)$ can be interpreted as a density of a conditional measure under some conditions. This is not problematic as long as one interprets the above derivation as a proof of backwards compatibility of entropic projections with the Bayes–Laplace framework (which is finitary anyway), and not as a method of extending this framework from finite to countably additive boolean algebras.

corresponding to a condition $q(\chi) = \int_\Theta d\theta q(\chi, \theta) = f(\chi)$ with a given probability density $f \in \mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, then the entropic projection (2.8) reproduces Jeffrey's rule (2.6) on Θ,

$$p_{\text{new}}(x, \theta) = \frac{p(\chi, \theta)}{p(\chi)} f(\chi) =: p(\chi|\theta) f(\chi) = p(\chi|\theta) p_{\text{new}}(\chi), \qquad (2.14)$$

$$p(\theta) \mapsto p_{\text{new}}(\theta) = \int_\mathcal{X} \tilde{\mu}(\chi) f(\chi) \frac{p(\chi, \theta)}{p(\chi)} = \int_\mathcal{X} \tilde{\mu}(\chi) p(\theta|\chi) f(\chi) = \int_\mathcal{X} \tilde{\mu}(\chi) p(\theta|\chi) p_{\text{new}}(\chi). \quad (2.15)$$

The Bayes–Laplace rule changes information states by means of constraints imposed on the level of propositions, while the entropic projections utilise constraints imposed on probabilities. Hence, in order to recover the former rule from the maximum entropy rule one needs a very strong constraint, which forces a unique reference of probability distribution to an underlying space $\mathcal{X}$ of propositions: $\int_\Theta d\theta p_{\text{new}}(\chi, \theta) = \delta(\chi - b)$. So, while (for example) the mean value constraints can be partially dismissed by the new knowledge that is incorporated by the sequential maximum relative entropy updating (see e.g. [113, 52]), Dirac's delta constraints always remain preserved by subsequent updatings.[5] From this perspective, Jeffrey's rule can be understood as arising due to weakening of constraints, which are allowed to carry some additional uncertainty.

More generally, Diaconis and Zabell [33] have shown that, for a suitable choice of constraints, Jeffrey's rule can be derived from a constrained minimisation of any Csiszár–Morimoto $\mathfrak{f}$-distance [28, 98, 2] for a strictly convex $\mathfrak{f} :]0, \infty[\to \mathbb{R}$. In particular, [43] derived Jeffrey's rule from constrained minimisation of $D_0|_{L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_1^+}$. In Theorem 4.29 of Section 4 we will derive a quantum analogue of the latter result.

3 Quantum distances and relative modular theory

In this section we present the mathematical terminology that is used throughout the paper. We also briefly introduce some more advanced notions and results from the theory of operator algebras that are required to set up the mathematical background and notation for the W^*-algebraic part of the results presented in Sections 4 and 5. See [79] for a detailed discussion of these structures and their properties.

3.1 States and weights over W^*-algebras

A ***C^*-algebra*** is a Banach space $\mathcal{C}$ over $\mathbb{C}$ with unit $\mathbb{I}$ that is also an algebra over $\mathbb{C}$ and is equipped with an operation $* : \mathcal{C} \to \mathcal{C}$ satisfying $(xy)^* = y^* x^*$, $(x + y)^* = x^* + y^*$, $x^{**} = x$, $(\lambda x)^* = \lambda^* x^*$, and $\|x^* x\| = \|x\|^2$, where λ^* is a complex conjugation of $\lambda \in \mathbb{C}$. A ***W^*-algebra*** is defined as such C^*-algebra that has a Banach predual. If a predual of C^*-algebra exists then it is unique. Given a W^*-algebra $\mathcal{N}$, we will denote its predual by $\mathcal{N}_\star$. Moreover, $\mathcal{N}_\star^+ := \{\phi \in \mathcal{N}_\star \mid \phi(x^* x) \geq 0 \,\forall x \in \mathcal{N}\}$, $\mathcal{N}_{\star 0}^+ := \{\phi \in \mathcal{N}_\star^+ \mid \omega(x^* x) = 0 \Rightarrow x = 0 \;\forall x \in \mathcal{N}\}$, $\mathcal{N}_{\star 1}^+ := \{\phi \in \mathcal{N}_\star^+ \mid \|\phi\| = 1\}$, $\mathcal{N}^{\text{sa}} := \{x \in \mathcal{N} \mid x^* = x\}$, $\mathcal{N}^+ := \{x \in \mathcal{N} \mid \exists y \in \mathcal{N} \;\; x = y^* y\}$, $\text{Proj}(\mathcal{N}) := \{x \in \mathcal{N}^{\text{sa}} \mid xx = x\}$. For $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\mathcal{N}_\star = \mathfrak{G}_1(\mathcal{H}) := \{x \in \mathfrak{B}(\mathcal{H}) \mid \|x\|_{\mathfrak{G}_1(\mathcal{H})} := \text{tr}(\sqrt{x^* x}) < \infty\}$. If $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is a localisable measure space, then $L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is a commutative W^*-algebra, and $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is its predual. Every commutative W^*-algebra can be represented in this form. This way the theory of W^*-algebras generalises both the localisable measure theory and the theory of bounded operators over Hilbert spaces. We define a ***statistical model*** as a set $\mathcal{M}(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu}) \subseteq L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})^+$, where $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is a localisable measure space. The elements of $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_1^+$ are Radon–Nikodým quotients of probability measures (dominated by $\tilde{\mu}$) with respect to $\tilde{\mu}$, and are called ***probability densities***. If $\mathcal{N}$ is a W^*-algebra, then we

---

[5]In the above discussion Dirac's delta constraint is applied by integration over Θ space, while usually the mean value constraints are applied by integration over $\mathcal{X}$ space, but nevertheless this remark holds in general.

define a **quantum model** as a set $\mathcal{M}(\mathcal{N}) \subseteq \mathcal{N}_\star^+$. The elements of $\mathcal{N}_\star^+$ will be called **quantum states** or (just) **states**.

A **weight** on a W^*-algebra $\mathcal{N}$ is defined as a function $\omega : \mathcal{N}^+ \to [0, +\infty]$ such that $\omega(0) = 0$, $\omega(x + y) = \omega(x) + \omega(y)$, and $\lambda \geq 0 \Rightarrow \omega(\lambda x) = \lambda\omega(x)$, with the convention $0 \cdot (+\infty) = 0$. A weight is called: **faithful** iff $\omega(x) = 0 \Rightarrow x = 0$; **finite** iff $\omega(\mathbb{I}) < \infty$; **semi-finite** iff a left ideal in $\mathcal{N}$ given by

$$\mathfrak{n}_\phi := \{x \in \mathcal{N} \mid \phi(x^*x) < \infty\} \tag{3.1}$$

is weakly-$\star$ dense in $\mathcal{N}$; **trace** iff $\omega(xx^*) = \omega(x^*x) \; \forall x \in \mathcal{N}$; **normal** iff $\omega(\sup\{x_\iota\}) = \sup\{\omega(x_\iota)\}$ for any uniformly bounded increasing net $\{x_\iota\} \subseteq \mathcal{N}^+$. A space of all normal semi-finite weights on a W^*-algebra $\mathcal{N}$ is denoted $\mathcal{W}(\mathcal{N})$, while the subset of all faithful elements of $\mathcal{W}(\mathcal{N})$ is denoted $\mathcal{W}_0(\mathcal{N})$. Every state is a finite normal weight, and every faithful state is a finite faithful normal state, hence the diagram

$$
\begin{array}{ccc}
\mathcal{N}_{\star 0}^+ & \hookrightarrow & \mathcal{W}_0(\mathcal{N}) \\
\cap \downarrow & & \cap \downarrow \\
\mathcal{N}_\star^+ & \hookrightarrow & \mathcal{W}(\mathcal{N})
\end{array}
\tag{3.2}
$$

commutes. For $\psi \in \mathcal{W}(\mathcal{N})$,

$$\mathrm{supp}(\psi) = \mathbb{I} - \sup\{P \in \mathrm{Proj}(\mathcal{N}) \mid \psi(P) = 0\}. \tag{3.3}$$

For $\omega, \phi \in \mathcal{N}_\star^+$ we will write $\omega \ll \phi$ iff $\mathrm{supp}(\omega) \leq \mathrm{supp}(\phi)$.[6] An element $\omega \in \mathcal{N}^{\mathbf{B}+}$ is faithful iff $\mathrm{supp}(\omega) = \mathbb{I}$. If ϕ is a normal weight on a W^*-algebra $\mathcal{N}$ (which includes $\omega \in \mathcal{N}_\star^+$ as a special case), then the restriction of ϕ to a **reduced** W^*-algebra,

$$\mathcal{N}_{\mathrm{supp}(\phi)} := \{x \in \mathcal{N} \mid \mathrm{supp}(\phi)x = x = x\,\mathrm{supp}(\phi)\} = \bigcup_{x \in \mathcal{N}} \{\mathrm{supp}(\phi)x\,\mathrm{supp}(\phi)\}, \tag{3.4}$$

is a faithful normal weight (respectively, an element of $(\mathcal{N}_{\mathrm{supp}(\phi)})_{\star 0}^+$). If ϕ is semi-finite, then $\phi|_{\mathcal{N}_{\mathrm{supp}(\phi)}} \in \mathcal{W}_0(\mathcal{N}_{\mathrm{supp}(\phi)})$. Hence, given $\psi \in \mathcal{W}(\mathcal{N})$ and $P \in \mathrm{Proj}(\mathcal{N})$, $P = \mathrm{supp}(\psi)$ iff $\psi|_{\mathcal{N}_P} \in \mathcal{W}_0(\mathcal{N}_P)$ and $\psi(P) = \psi(PxP) \; \forall x \in \mathcal{N}^+$. In particular, for $\omega, \phi \in \mathcal{N}_\star^+$ and $\omega \ll \phi$, we have $\omega|_{\mathcal{N}_{\mathrm{supp}(\phi)}} \in \mathcal{W}_0(\mathcal{N}_{\mathrm{supp}(\phi)})$.

A **representation** of a C^*-algebra $\mathcal{C}$ is defined as a pair $(\mathcal{H}, \pi)$ of a Hilbert space $\mathcal{H}$ and a $*$-homomorphism $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$. A representation $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ is called: **nondegenerate** iff $\{\pi(x)\xi \mid (x, \xi) \in \mathcal{C} \times \mathcal{H}\}$ is dense in $\mathcal{H}$; **normal** iff it is continuous with respect to the weak-$\star$ topologies of $\mathcal{C}$ and $\mathfrak{B}(\mathcal{H})$; **faithful** iff $\ker(\pi) = \{0\}$. An element $\xi \in \mathcal{H}$ is called **cyclic** for a C^*-algebra $\mathcal{C} \subseteq \mathfrak{B}(\mathcal{H})$ iff $\mathcal{C}\xi := \bigcup_{x \in \mathcal{C}}\{x\xi\}$ is norm dense in $\mathfrak{B}(\mathcal{H})$. A representation $\pi : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ of a C^*-algebra $\mathcal{C}$ is called **cyclic** iff there exists $\Omega \in \mathcal{H}$ that is cyclic for $\pi(\mathcal{C})$. According to the Gel'fand–Naĭmark–Segal theorem [50, 111] for every pair $(\mathcal{C}, \omega)$ of a C^*-algebra $\mathcal{C}$ and $\omega \in \mathcal{C}^{\mathbf{B}+}$ there exists a triple $(\mathcal{H}_\omega, \pi_\omega, \Omega_\omega)$ of a Hilbert space $\mathcal{H}_\omega$ and a cyclic representation $\pi_\omega : \mathcal{C} \to \mathfrak{B}(\mathcal{H})$ with a cyclic vector $\Omega_\omega \in \mathcal{H}_\omega$, and this triple is unique up to unitary equivalence. It is constructed as follows. For a C^*-algebra $\mathcal{C}$ and $\omega \in \mathcal{C}^{\mathbf{B}+}$, one defines the scalar form $\langle \cdot, \cdot \rangle_\omega$ on $\mathcal{C}$,

$$\langle x, y \rangle_\omega := \omega(x^*y) \; \; \forall x, y \in \mathcal{C}, \tag{3.5}$$

and the **Gel'fand ideal**

$$\mathcal{I}_\omega := \{x \in \mathcal{C} \mid \omega(x^*x) = 0\} = \{x \in \mathcal{C} \mid \omega(x^*y) = 0 \; \forall y \in \mathcal{C}\}, \tag{3.6}$$

which is a left ideal of $\mathcal{C}$, closed in the norm topology (it is also closed in the weak-$\star$ topology if $\omega \in \mathcal{C}_\star^{\mathbf{B}+}$). The form $\langle \cdot, \cdot \rangle_\omega$ is hermitean on $\mathcal{C}$ and it becomes a scalar product $\langle \cdot, \cdot \rangle_\omega$ on $\mathcal{C}/\mathcal{I}_\omega$.

---

[6]If $\mathcal{N} = \mathfrak{B}(\mathcal{H})$ and $\omega = \mathrm{tr}(\rho_\omega \cdot)$ for $\rho_\omega \in \mathfrak{G}_1(\mathcal{H})^+$, then $\mathrm{supp}(\omega) = \mathrm{ran}(\rho_\omega)$, so for any $\phi = \mathrm{tr}(\rho_\phi \cdot)$ with $\rho_\phi \in \mathfrak{G}_1(\mathcal{H})^+$ one has $\omega \ll \phi$ iff $\mathrm{ran}(\rho_\omega) \subseteq \mathrm{ran}(\rho_\phi)$.

The Hilbert space $\mathcal{H}_\omega$ is obtained by the completion of $\mathcal{C}/\mathcal{I}_\omega$ in the topology of norm generated by $\langle\cdot,\cdot\rangle_\omega$. Consider the morphisms

$$[\cdot]_\omega : \mathcal{C} \ni x \longmapsto [x]_\omega \in \mathcal{C}/\mathcal{I}_\omega, \tag{3.7}$$

$$\pi_\omega(y) : [y]_\omega \longmapsto [xy]_\omega. \tag{3.8}$$

The element $\omega \in \mathcal{C}^{\mathbf{B}+}$ is uniquely represented in terms of $\mathcal{H}_\omega$ by the vector $[\mathbb{I}]_\omega =: \Omega_\omega \in \mathcal{H}_\omega$, which is cyclic for $\pi_\omega(\mathcal{C})$ and satisfies $\|\Omega_\omega\| = \|\omega\|$. Hence

$$\omega(x) = \langle \Omega_\omega, \pi_\omega(x)\Omega_\omega \rangle_\omega \quad \forall x \in \mathcal{C}, \tag{3.9}$$

An analogue of this theorem for weights follows the similar construction, but lacks cyclicity. If $\mathcal{N}$ is a W^*-algebra, and ω is a weight on $\mathcal{N}$, then there exists the Hilbert space $\mathcal{H}_\omega$, defined as the completion of $\mathfrak{n}_\omega/\ker(\omega)$ in the topology of a norm generated by the scalar product $\langle\cdot,\cdot\rangle_\omega : \mathfrak{n}_\omega \times \mathfrak{n}_\omega \ni (x,y) \mapsto \omega(x^*y) \in \mathbb{C}$,

$$\mathcal{H}_\omega := \overline{\mathfrak{n}_\omega/\ker(\omega)} = \overline{\{x \in \mathcal{N} \mid \omega(x^*x) < \infty\}/\{x \in \mathcal{N} \mid \omega(x^*x) = 0\}} = \overline{\mathfrak{n}_\omega/\mathcal{I}_\omega}, \tag{3.10}$$

and there exist the maps

$$[\cdot]_\omega : \mathfrak{n}_\omega \ni x \mapsto [x]_\omega \in \mathcal{H}_\omega, \tag{3.11}$$

$$\pi_\omega : \mathcal{N} \ni x \mapsto ([y]_\omega \mapsto [xy]_\omega) \in \mathfrak{B}(\mathcal{H}_\omega), \tag{3.12}$$

such that $[\cdot]_\omega$ is linear, $\mathrm{ran}([\cdot]_\omega)$ is dense in $\mathcal{H}_\omega$, and $(\mathcal{H}_\omega, \pi_\omega)$ is a representation of $\mathcal{N}$. If $\omega \in \mathcal{W}(\mathcal{N})$ then $(\mathcal{H}_\omega, \pi_\omega)$ is nondegenerate and normal. It is also faithful if $\omega \in \mathcal{W}_0(\mathcal{N})$. The **commutant** of a subalgebra $\mathcal{N}$ of any algebra $\mathcal{C}$ is defined as

$$\mathcal{N}^\bullet := \{y \in \mathcal{C} \mid xy = yx \,\forall x \in \mathcal{N}\}, \tag{3.13}$$

while the **center** of $\mathcal{N}$ is defined as $\mathfrak{Z}_\mathcal{N} := \mathcal{N} \cap \mathcal{N}^\bullet$. A unital $*$-subalgebra $\mathcal{N}$ of an algebra $\mathfrak{B}(\mathcal{H})$ is called the **von Neumann algebra** [119, 99] iff $\mathcal{N} = \mathcal{N}^{\bullet\bullet}$. An image $\pi(\mathcal{N})$ of any representation $(\mathcal{H}, \pi)$ of a W^*-algebra $\mathcal{N}$ is a von Neumann algebra iff π is normal and nondegenerate.

A subspace $\mathcal{D} \subseteq \mathcal{H}$ of a complex Hilbert space $\mathcal{H}$ is called a **cone** iff $\lambda\xi \in \mathcal{D} \,\forall \xi \in \mathcal{D} \,\forall \lambda \geq 0$. A cone $\mathcal{D} \subseteq \mathcal{H}$ is called **self-polar** iff

$$\mathcal{D} = \{\zeta \in \mathcal{H} \mid \langle \xi, \zeta \rangle_\mathcal{H} \geq 0 \,\forall \xi \in \mathcal{D}\}. \tag{3.14}$$

Every self-polar cone $\mathcal{D} \subseteq \mathcal{H}$ is pointed ($\mathcal{D} \cap (-\mathcal{D}) = \{0\}$), spans linearly $\mathcal{H}$ ($\mathrm{span}_\mathbb{C}\mathcal{D} = \mathcal{H}$), and determines a unique conjugation[7] J in $\mathcal{H}$ such that $J\xi = \xi \,\forall \xi \in \mathcal{H}$ [57], as well as a partial order on the set $\mathcal{H}^{\mathrm{sa}} := \{\xi \in \mathcal{H} \mid J\xi = \xi\}$ given by,

$$\xi \leq \zeta \iff \xi - \zeta \in \mathcal{D} \,\forall \xi, \zeta \in \mathcal{H}^{\mathrm{sa}}. \tag{3.15}$$

If $\mathcal{N}$ is a W^*-algebra, $\mathcal{H}$ is a Hilbert space, $\mathcal{H}^\natural \subseteq \mathcal{H}$ is a self-polar cone, π is a nondegenerate faithful normal representation of $\mathcal{N}$ on $\mathcal{H}$, and J is conjugation on $\mathcal{H}$, then the quadruple $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$ is called **standard representation** of $\mathcal{N}$ and $(\mathcal{H}, \pi(\mathcal{N}), J, \mathcal{H}^\natural)$ is called **standard form** of $\mathcal{N}$ iff the conditions [58]

$$J\pi(\mathcal{N})J = \pi(\mathcal{N})^\bullet, \ \xi \in \mathcal{H}^\natural \Rightarrow J\xi = \xi, \ \pi(x)J\pi(x)J\mathcal{H}^\natural \subseteq \mathcal{H}^\natural, \ \pi(x) \in \mathfrak{Z}_{\pi(\mathcal{N})} \Rightarrow J\pi(x)J = \pi(x)^*. \tag{3.16}$$

---

[7] A linear operator $J : \mathrm{dom}(J) \to \mathcal{H}$, where $\mathrm{dom}(J) \subseteq \mathcal{H}$, is called a **conjugation** iff it is antilinear, isometric, and involutive ($J^2 = \mathbb{I}$).

hold. For any standard representation

$$\forall \phi \in \mathcal{N}_\star^+ \ \exists! \xi_\pi(\phi) \in \mathcal{H}^\natural \ \forall x \in \mathcal{N} \ \ \phi(x) = \langle \xi_\pi(\phi), \pi(x)\xi_\pi(\phi)\rangle_\mathcal{H} \tag{3.17}$$

holds. The map $\xi_\pi : \mathcal{N}_\star^+ \to \mathcal{H}^\natural$ is order preserving. If $\mathcal{N} = \mathfrak{B}(\mathcal{K})$ for some Hilbert space $\mathcal{K}$, then the standard representation Hilbert space is given by the space $\mathcal{K} \otimes \mathcal{K}^{\mathbf{B}} \cong \mathfrak{G}_2(\mathcal{K}) := \{x \in \mathfrak{B}(\mathcal{K}) \mid \sqrt{\mathrm{tr}_\mathcal{K}(x^*x)} < \infty\}$ of Hilbert–Schmidt operators equipped with an inner product $\langle x, y\rangle := \mathrm{tr}_\mathcal{K}(x^*y)$, where $(\cdot)^{\mathbf{B}}$ denotes Banach dual space. In this case, the map (3.17) reads $\xi_\pi : \mathfrak{G}_1(\mathcal{H})^+ \ni \rho \mapsto \rho^{1/2} \in \mathfrak{G}_2(\mathcal{H})^+$.

3.2 Relative modular theory

For a given W^*-algebra $\mathcal{N}$, $\phi \in \mathcal{W}(\mathcal{N})$, and $\omega \in \mathcal{W}_0(\mathcal{N})$ the map

$$R_{\phi,\omega} : [x]_\omega \mapsto [x^*]_\phi \ \ \forall x \in \mathfrak{n}_\omega \cap \mathfrak{n}_\phi^* \tag{3.18}$$

is a densely defined, closable antilinear operator. Its closure admits a unique polar decomposition

$$\overline{R}_{\phi,\omega} = J_{\phi,\omega}\Delta_{\phi,\omega}^{1/2}, \tag{3.19}$$

where $J_{\phi,\omega}$ is a conjugation operator, called **relative modular conjugation**, while $\Delta_{\phi,\omega}$ is a positive self-adjoint operator on $\mathrm{dom}(\Delta_{\phi,\omega}) \subseteq \mathcal{H}_\omega$ with $\mathrm{supp}(\Delta_{\phi,\omega}) = \mathrm{supp}(\phi)\mathcal{H}_\omega$, called a **relative modular operator** [4, 27, 35]. The relative modular operators allow to define a one-parameter family of partial isometries in $\mathrm{supp}(\phi)\mathcal{N}$, called **Connes' cocycle** [26],

$$\mathbb{R} \ni t \mapsto [\phi : \omega]_t := \Delta_{\phi,\psi}^{\mathrm{it}}\Delta_{\omega,\psi}^{-\mathrm{it}} = \Delta_{\phi,\omega}^{\mathrm{it}}\Delta_{\omega,\omega}^{-\mathrm{it}} \in \mathrm{supp}(\phi)\mathcal{N}, \tag{3.20}$$

where $\psi \in \mathcal{W}_0(\mathcal{N})$ is arbitrary, so it can be set equal to ω. As shown by Araki and Masuda [10] (see also [92]), the definition of $\Delta_{\phi,\omega}$ and $[\phi : \omega]_t$ can be further extended to the case when $\phi, \omega \in \mathcal{W}(\mathcal{N})$, by means of a densely defined closable antilinear operator

$$R_{\phi,\omega} : [x]_\omega + (\mathbb{I} - \mathrm{supp}(\overline{[\mathfrak{n}_\phi]_\omega}))\zeta \mapsto \mathrm{supp}(\omega)[x^*]_\phi \ \ \forall x \in \mathfrak{n}_\omega \cap \mathfrak{n}_\phi^* \ \forall \zeta \in \mathcal{H}, \tag{3.21}$$

where $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$ is a standard representation of a W^*-algebra $\mathcal{N}$, and $\mathcal{H}_\phi \subseteq \mathcal{H} \supseteq \mathcal{H}_\omega$. For $\phi, \omega \in \mathcal{N}_\star^+$ this becomes a closable antilinear operator [9, 76]

$$R_{\phi,\omega} : x\xi_\pi(\omega) + \zeta \mapsto \mathrm{supp}(\omega)x^*\xi_\pi(\phi) \ \ \forall x \in \pi(\mathcal{N}) \ \forall \zeta \in (\pi(\mathcal{N})\xi_\pi(\omega))^\perp, \tag{3.22}$$

acting on a dense domain $(\pi(\mathcal{N})\xi_\pi(\omega)) \cup (\pi(\mathcal{N})\xi_\pi(\omega))^\perp \subseteq \mathcal{H}$, where $(\pi(\mathcal{N})\xi_\pi(\omega))^\perp$ denotes a complement of the closure in $\mathcal{H}$ of the linear span of the action $\pi(\mathcal{N})$ on $\xi_\pi(\omega)$. In both cases, the relative modular operator is determined by the polar decomposition of the closure $\overline{R}_{\phi,\omega}$ of $R_{\phi,\omega}$,

$$\Delta_{\phi,\omega} := R_{\phi,\omega}^* \overline{R}_{\phi,\omega}. \tag{3.23}$$

If (3.21) or (3.22) is used instead of (3.18), then the formula (3.20) has to be replaced by

$$\mathbb{R} \ni t \mapsto [\phi : \omega]_t \mathrm{supp}(\overline{[\mathfrak{n}_\phi]_\psi}) := \Delta_{\phi,\psi}^{\mathrm{it}}\Delta_{\omega,\psi}^{-\mathrm{it}}, \tag{3.24}$$

and $[\phi : \omega]_t$ is a partial isometry in $\mathrm{supp}(\phi)\mathcal{N}\mathrm{supp}(\omega)$ whenever $[\mathrm{supp}(\phi), \mathrm{supp}(\omega)] = 0$.

3.3 Quantum distances

Given any set X, a *distance* is defined as a map $D : X \times X \to [0, \infty]$ such that $D(x, y) = 0 \iff x = y$. A *relative entropy* is a map $\mathbf{S} : X \times X \to [-\infty, 0]$ such that $(-\mathbf{S})$ is a distance. A distance is called: *bounded* iff $\operatorname{ran}(D) = \mathbb{R}^+$; *symmetric* iff $D(x, y) = D(y, x)$; *metrical* [47] iff is it bounded, symmetric and satisfies *triangle inequality*

$$D(x, y) \leq D(x, y) + D(y, z) \quad \forall x, y, z \in X. \tag{3.25}$$

We will use the symbol d instead of D to denote metrical distances. A distance on a statistical model will be called a *statistical distance*, while a distance on a quantum model will be called a *quantum distance*. A term *information distance* with be used to refer to any of them. Let $\wp(X)$ denote a powerset of X. If $K \subseteq X$ is such that the map

$$\mathfrak{P}_K^D : X \ni \phi \mapsto \arg\inf_{\omega \in K} \{D(\omega, \phi)\} \in \wp(K) \tag{3.26}$$

is a singleton (one-element set, $\{*\}$) for all $\phi \in X$, then we call (3.26) an *entropic projection*[8].

For any W^*-algebra $\mathcal{N}$, and $\phi, \omega \in \mathcal{N}_\star^+$ the following formulas define quantum distances on $\mathcal{N}_\star^+$,

$$D_0(\omega, \phi) := \begin{cases} (\omega - \phi)(\mathbb{I}) - \langle \xi_\pi(\phi), \log(\Delta_{\omega,\phi})\xi_\pi(\phi) \rangle_{\mathcal{H}} & : \omega \ll \phi \\ +\infty & : \text{otherwise}, \end{cases} \tag{3.27}$$

$$D_1(\omega, \phi) := \begin{cases} (\phi - \omega)(\mathbb{I}) - \langle \xi_\pi(\omega), \log(\Delta_{\omega,\phi})\xi_\pi(\omega) \rangle_{\mathcal{H}} & : \omega \ll \phi \\ +\infty & : \text{otherwise}, \end{cases} \tag{3.28}$$

$$D_{1/2}(\psi, \phi) := \begin{cases} 2(\phi + \omega)(\mathbb{I}) - 4 \left\langle \xi_\pi(\phi), \Delta_{\omega,\phi}^{1/2}\xi_\pi(\phi) \right\rangle_{\mathcal{H}} & : \omega \ll \phi \\ +\infty & : \text{otherwise}. \end{cases} \tag{3.29}$$

Hence,

$$\phi \ll \omega \ll \phi \Rightarrow D_0(\omega, \phi) = D_1(\phi, \omega). \tag{3.30}$$

All above examples are special cases of a family of quantum γ-distances D_γ with $\gamma \in [0, 1]$, see [61, 100, 73, 77]. A special case of D_1 is the *Araki distance* [6, 8, 9]

$$D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi) = \begin{cases} - \langle \xi_\pi(\omega), \log(\Delta_{\phi,\omega})\xi_\pi(\omega) \rangle_{\mathcal{H}} & : \omega \ll \phi \\ +\infty & : \text{otherwise}. \end{cases} \tag{3.31}$$

If $D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi) < \infty$, then (3.31) takes the form [102, 103]

$$D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi) = \begin{cases} \mathrm{i} \lim_{t \to +0} \frac{\omega}{t} ([\phi : \omega]_t - \mathbb{I}) & : \omega \ll \phi \\ +\infty & : \text{otherwise}. \end{cases} \tag{3.32}$$

For a semi-finite $\mathcal{N}$, normal faithful semi-finite trace τ on $\mathcal{N}$ and ρ_ϕ and ρ_ω defined by $\phi(\cdot) = \tau(\rho_\phi \cdot)$ and $\omega(\cdot) = \tau(\rho_\omega \cdot)$, the Araki distance (3.31) turns to the *Umegaki distance* [115, 116] (cf. also [6, 7])

$$D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi) = \tau(\rho_\omega(\log \rho_\omega - \log \rho_\phi)) = \tau \left(\rho_\omega^{1/2}(\log \Delta_{\omega,\phi})\rho_\omega^{1/2} \right) \tag{3.33}$$

if $\omega \ll \phi$, and $D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi) = +\infty$ otherwise. A special, but definitely most popular, case of (3.33) is obtained for $\mathcal{N} = \mathfrak{B}(\mathcal{H})$ and $\tau = \operatorname{tr}$.

For a commutative $\mathcal{N} \cong L_\infty(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$, where $(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})$ is any localisable measure space such that $\tilde{\mu}_\phi \ll \tilde{\mu}$ and $\tilde{\mu}_\omega \ll \tilde{\mu}$ where $\psi(x) =: \int_{\mathcal{X}} \tilde{\mu}_\psi(\chi)x(\chi) \, \forall x \in \mathcal{N}^+$ and $\psi \in \{\phi, \omega\}$,

---

[8] So, by definition, every entropic projection $\mathfrak{P}_K^D$ is a unique maximiser of a relative entropy $-D$.

the Araki distance (3.31) takes a form of the **Wald–Good–Kullback–Leibler distance** [121, 53, 83],

$$D_1|_{L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})_1^+}(\omega,\phi) = \begin{cases} \int \tilde{\mu}_\omega \log \frac{\tilde{\mu}_\omega}{\tilde{\mu}_\phi} & : \tilde{\mu}_\omega \ll \tilde{\mu}_\phi \\ +\infty & : \text{otherwise.} \end{cases} \tag{3.34}$$

Under these conditions, and assuming $\tilde{\mu}_\phi \ll \tilde{\mu}_\omega$, one has

$$D_1|_{L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})_1^+}(\omega,\phi) = D_0|_{L_1(\mathcal{X},\mho(\mathcal{X}),\tilde{\mu})_1^+}(\phi,\omega). \tag{3.35}$$

Among above distances, only $D_{1/2}$ is symmetric, and none of them is metrical. Metrical quantum distances will be discussed in Section 5.

4 Lüders' and quantum Jeffrey's rules as entropic projections

Definition 4.1. *If $\rho \in \mathfrak{G}_1(\mathcal{H})^+$, I is a countable set, and $\{P_i \mid i \in I\} \subseteq \mathrm{Proj}(\mathfrak{B}(\mathcal{H}))$ satisfies $\sum_{i \in I} P_i = \mathbb{I} \in \mathfrak{B}(\mathcal{H})$ and $P_i P_j = \delta_{ij} P_i \; \forall i,j \in I$, then the **weak Lüders rule** is defined as a map [84, 110]*

$$\mathfrak{G}_1(\mathcal{H})^+ \ni \rho \mapsto \sum_{i \in I} P_i \rho P_i \in \mathfrak{G}_1(\mathcal{H})^+. \tag{4.1}$$

*If $P \in \mathrm{Proj}(\mathfrak{B}(\mathcal{H}))$, then the **strong Lüders rule** is defined as a map*

$$\mathfrak{G}_1(\mathcal{H})_1^+ \ni \rho \mapsto \frac{P\rho P}{\mathrm{tr}(P\rho P)} \in \mathfrak{G}_1(\mathcal{H})_1^+, \tag{4.2}$$

*with domain restricted by the condition $\mathrm{tr}(P\rho) \neq 0$. The **semi-strong Lüders rule**, called sometimes a 'partial collapse', is defined as a map*

$$\mathfrak{G}_1(\mathcal{H})_1^+ \ni \rho \mapsto \frac{\sum_{j \in J} P_j \rho P_j}{\sum_{j \in J} \mathrm{tr}(P_j \rho P_j)} \in \mathfrak{G}_1(\mathcal{H})_1^+, \tag{4.3}$$

where J is a subset of a countable set I corresponding to an orthogonal decomposition $\sum_{i \in I} P_i = \mathbb{I} \in \mathfrak{B}(\mathcal{H})$, and the domain in (4.3) is restricted by a condition $\sum_{j \in J} \mathrm{tr}(P_j \rho) \neq 0$.

Remark 4.2. If ρ, I, and $\{P_i \mid i \in I\}$ are such as in (4.1), $J \subseteq I$, and the condition

$$\mathrm{tr}(P_i \rho) \neq 0 \iff i \in J \;\; \forall i \in I \tag{4.4}$$

holds, then (4.1) can be written as

$$\mathfrak{G}_1(\mathcal{H})^+ \ni \rho \mapsto \sum_{j \in J} \frac{P_j \rho P_j}{\mathrm{tr}(\rho P_j)} \mathrm{tr}(\rho P_j) \in \mathfrak{G}_1(\mathcal{H})^+. \tag{4.5}$$

In particular, if $J = \{*\}$, then (4.5) turns to the direct extension of strong Lüders' rule (4.2) to $\mathfrak{G}_1(\mathcal{H})^+$, which coincides with (4.2) on $\mathfrak{G}_1(\mathcal{H})_1^+$.

Remark 4.3. In the case when $\rho^2 = \rho$, then $\exists \xi \in \mathcal{H}$ such that $\rho = P_{\mathrm{span}(\xi)}$ (i.e., ρ is a projector onto a closed one-dimensional linear subspace of $\mathcal{H}$ spanned by ξ), and the strong Lüders rule takes the form of the **strong von Neumann rule** [120], called also a 'state vector reduction',

$$\mathcal{H} \ni \xi \mapsto \frac{P\xi}{\langle \xi, P\xi \rangle^{\frac{1}{2}}} \in \mathcal{H}. \tag{4.6}$$

Definition 4.4. *We define* **quantum Jeffrey's rule** *as a map*

$$\mathfrak{G}_1(\mathcal{H})_1^+ \ni \rho \mapsto \rho_{\mathrm{new}} := \sum_{i=1}^n \frac{P_i \rho P_i}{\mathrm{tr}(\rho P_i)} \lambda_i \in \mathfrak{G}_1(\mathcal{H})_1^+, \tag{4.7}$$

where $n \in \mathbb{N}$, $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathfrak{B}(\mathcal{H}))$, $\sum_{i=1}^n P_i = \mathbb{I} \in \mathfrak{B}(\mathcal{H})$, *and* $\mathrm{tr}(\rho P_i) \neq 0$, $P_i P_j = \boldsymbol{\delta}_{ij} P_i$, $\lambda_i = \mathrm{tr}(\rho_{\mathrm{new}} P_i) \ \forall i, j \in \{1, \dots, n\}$.

Remark 4.5. In the case when an orthogonal decomposition of $\mathbb{I} \in \mathfrak{B}(\mathcal{H})$ is given by the set $\{P, \mathbb{I} - P\}$, and $\mathrm{tr}(\rho_{\mathrm{new}}(\mathbb{I} - P)) = 0$, then (4.7) reduces to (4.2). The direct analogy between this property and the conditions under which Jeffrey's rule (2.6) reduces to the Bayes–Laplace rule (2.5), together with the analogy between (4.7) and (2.6), justify the name we gave to (4.7).

Proposition 4.6. *Definition* (4.7) *of* ρ_{new} *is equivalent to a condition*

$$\frac{\mathrm{tr}(\rho_{\mathrm{new}} P_i P)}{\mathrm{tr}(\rho_{\mathrm{new}} P_i)} = \frac{\mathrm{tr}(\rho P_i P)}{\mathrm{tr}(\rho P_i)} \quad \forall P \in \mathrm{Proj}(\mathfrak{B}(\mathcal{H})) \ such \ that \ [P, P_i] = 0 \ \forall i \in \{1, \dots, n\}, \tag{4.8}$$

with $[\rho_{\mathrm{new}}, P_i] = 0 \ \forall i \in \{1, \dots, n\}$.

Proof. An implication from (4.7) to (4.8) is easy. In the opposite direction, let us consider an arbitrary $x \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$ such that $[x, P_i] = 0 \ \forall i \in \{1, \dots, n\}$. For an arbitrary countable set I the condition $[P_i, \rho_{\mathrm{new}}] = 0 \ \forall i \in I$ is equivalent to $\rho_{\mathrm{new}} = \sum_{i \in I} P_i \rho_{\mathrm{new}} P_i$ and to $[\rho_{\mathrm{new}}, x] = 0$, where $x = \sum_{i \in I} \lambda_i P_i \in \mathfrak{B}(\mathcal{H})^{\mathrm{sa}}$ with arbitrary $\{\lambda_i \in \mathbb{R} \mid i \in I\}$. See [64] for a clear discussion of these conditions and their equivalence. Hence,

$$\mathrm{tr}(\rho_{\mathrm{new}} x) = \sum_{i=1}^n \mathrm{tr}(P_i \rho_{\mathrm{new}} P_i x) = \sum_{i=1}^n \mathrm{tr}(\rho_{\mathrm{new}} P_i x P_i) = \sum_{i=1}^n \mathrm{tr}(\rho_{\mathrm{new}} P_i) \frac{\mathrm{tr}(\rho P_i x P_i)}{\mathrm{tr}(\rho P_i)}$$
$$= \sum_{i=1}^n \mathrm{tr}(\rho_{\mathrm{new}} P_i) \frac{\mathrm{tr}(P_i \rho P_i x)}{\mathrm{tr}(\rho P_i)}. \tag{4.9}$$

This gives

$$\rho_{\mathrm{new}} = \sum_{i=1}^n \mathrm{tr}(\rho_{\mathrm{new}} P_i) \frac{P_i \rho P_i}{\mathrm{tr}(\rho P_i)}. \tag{4.10}$$

$\square$

Remark 4.7. In what follows we will prove that the weak Lüders rule is a special case of an entropic projection, determined by constrained minimisation with a $D_0(\phi, \psi)$ distance (Theorems 4.20 and 4.28). Next, we will use different constraints to derive quantum Jeffrey's rule (Theorem 4.29). Finally, we will show that the strong Lüders rule arises as a weakly continuous limit of quantum Jeffrey's rule (Remark 4.30), and can be also obtained by a constrained minimisation of a *regularised* modification of $D_0(\phi, \psi)$ distance (Theorem 4.34). The necessity of recourse to limit or regularisation indicates that, as opposed to weak Lüders' rule, strong Lüders' rule is not directly derivable from minimisation of D_0 distance. However, in Section 5 we will show that in some special cases (which usually turn it to the strong von Neumann rule) it can be derived from minimisation of $D_{1/2}$ distance.

Remark 4.8. According to de Muynck [32] (see also [46, 90, 75, 91]), the strong von Neumann and strong Lüders rules should be viewed as procedures of updating of quantum state, yet not in the *predictive* (inferential) sense, but in the *preparative* (calibrating) sense. If one extends this interpretation to the weak Lüders rule, and assumes that the 'preparative' use of entropic projection should be implemented by constrained minimisation of information distance in its *second* argument, then our result can be interpreted by saying that the weak Lüders rule of *quantum state preparation* can be derived from the constrained minimisation of the Araki distance $D_1|_{\mathcal{N}_{\star 1}^+}(\omega, \phi)$. (We do not consider this interpretation as necessary.)

Remark 4.9. In what follows $\mathcal{M}$ will be an arbitrary set, $\mathcal{Q} \subseteq \mathcal{M}$ its arbitrary subset, and D will be an arbitrary (not necessarily bounded) distance on $\mathcal{M}$. If $\arg\inf_{\phi \in \mathcal{Q}} \{D(\phi, \psi)\}$ consists of a single element, then we will denote it by $\mathfrak{P}_{\mathcal{Q}}^D(\psi)$.

Definition 4.10. *Let $\psi \in \mathcal{M}$ and $\mathcal{Q} \subseteq \mathcal{M}$. If*

$$\exists \rho \in \mathcal{Q} \; \forall \phi \in \mathcal{Q} \;\; D(\phi, \rho) + D(\rho, \psi) = D(\phi, \psi) \tag{4.11}$$

*holds, then we will say that $\mathcal{Q}$ satisfies **triangle equality** for ψ at $\rho \in \mathcal{Q}$ with respect to D.*

Definition 4.11. *Let $\mathcal{Q}_1, \mathcal{Q}_2 \subseteq \mathcal{M}$. If*

$$\forall \psi \in \mathcal{Q}_1 \quad \arg\inf_{\phi \in \mathcal{Q}_2} \{D(\phi, \psi)\} \cong \{*\} \;\; and \;\; \mathfrak{P}_{\mathcal{Q}_2}^D(\psi) \in \mathcal{Q}_1, \tag{4.12}$$

*then we will say that the pair $(\mathcal{Q}_1, \mathcal{Q}_2)$ satisfies **subprojection property** with respect to D.*

Lemma 4.12. *If $\mathcal{Q} \subseteq \mathcal{M}$ satisfies triangle equality for ψ at $\rho \in \mathcal{Q}$ with respect to D, then*

$$\rho = \arg\inf_{\phi \in \mathcal{Q}} \{D(\phi, \psi)\} \tag{4.13}$$

Proof. From $D(\phi, \rho) \geq 0$ and $D(\phi, \rho) = 0 \iff \phi = \rho$ it follows that $\rho = \arg\inf_{\phi \in \mathcal{Q}} \{D(\phi, \rho)\}$. From

$$\arg\inf_{\phi \in \mathcal{Q}} \{D(\phi, \rho)\} = \arg\inf_{\phi \in \mathcal{Q}} \{D(\rho, \psi) + D(\phi, \rho)\} \tag{4.14}$$

it follows that $\arg\inf_{\phi \in \mathcal{Q}} \{D(\phi, \psi)\}$ exists, is unique, and is equal to ρ. $\qquad \square$

Lemma 4.13. *Let $(\mathcal{Q}_1, \mathcal{Q}_2)$ satisfy the subprojection property with respect to D, and let $\mathcal{Q}_1, \mathcal{Q}_2 \in \mathcal{M}$ satisfy triangle equality for $\psi \in \mathcal{M}$ at $\mathfrak{P}_{\mathcal{Q}_2}^D(\psi)$ with respect to D. Then $\mathcal{Q}_1 \cap \mathcal{Q}_2$ satisfies triangle equality for $\psi \in \mathcal{M}$ at $\mathfrak{P}_{\mathcal{Q}_1 \cap \mathcal{Q}_2}^D(\psi)$ with respect to D and*

$$\arg\inf_{\phi \in \mathcal{Q}_1 \cap \mathcal{Q}_2} \{D(\phi, \psi)\} = \arg\inf_{\phi_2 \in \mathcal{Q}_2} \left\{ D\left(\phi_2, \arg\inf_{\phi_1 \in \mathcal{Q}_1} \{D(\phi_1, \psi)\} \right) \right\}. \tag{4.15}$$

Proof. Triangle equalities in this case read

$$\exists \rho_1 \in \mathcal{Q}_1 \; \forall \phi_1 \in \mathcal{Q}_1 \;\; D(\phi_1, \rho_1) + D(\rho_1, \psi) = D(\phi_1, \psi), \tag{4.16}$$

$$\exists \rho_2 \in \mathcal{Q}_2 \; \forall \phi_2 \in \mathcal{Q}_2 \;\; D(\phi_2, \rho_2) + D(\rho_2, \rho_1) = D(\phi_2, \rho_1). \tag{4.17}$$

Now, let $\rho_2 \in \mathcal{Q}_1 \cap \mathcal{Q}_2$. This gives

$$D(\phi, \rho_1) + D(\rho_2, \rho_1) + D(\rho_1, \psi) = D(\rho_2, \psi) + D_1(\phi, \rho_2) + D(\rho_2, \rho_1). \tag{4.18}$$

For $\phi \in \mathcal{Q}_1 \cap \mathcal{Q}_2$ we have

$$D(\phi, \rho_1) + D(\rho_1, \psi) = D(\phi, \psi). \tag{4.19}$$

This gives

$$D(\phi, \rho_2) + D(\rho_2, \psi) = D(\phi, \psi). \tag{4.20}$$

$\qquad \square$

Proposition 4.14. *If $\mathcal{Q}_i$ satisfies triangle equality for every $i \in \{1, \ldots, n\}$, and $(\mathcal{Q}_i, \mathcal{Q}_j)$ satisfy subprojection property for every $i, j \in \{1, \ldots, n\}$, then*

$$\arg\inf_{\phi \in \mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_n} \{D(\phi, \psi)\} = \rho_n, \tag{4.21}$$

where $\rho_k = \arg\inf_{\phi \in \mathcal{Q}_k} \{D(\phi, \rho_{k-1})\}$, $\rho_0 = \psi$, and $\mathcal{Q}_n$ satisfies triangle equality at ρ_n.

12

Proof. We will prove this lemma by mathematical induction. Let us assume that it holds for some $k \in \mathbb{N}$. Then

$$\rho_k = \underset{\phi \in \mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_k}{\arg\inf} \{D(\phi, \psi)\} \tag{4.22}$$

and $\mathcal{Q}_k$ satisfies triangle equality with ρ_k. Let $\psi_k \in \mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_k$, and consider

$$\mathfrak{P}^D_{\mathcal{Q}_{k+1}}(\psi_k) := \underset{\phi \in \mathcal{Q}_{k+1}}{\arg\inf} \{D(\phi, \psi_k)\}. \tag{4.23}$$

Then $\psi_k \in \mathcal{Q}_i$ for every $i \leq k$, and from the subprojection property for $(\mathcal{Q}_i, \mathcal{Q}_{k+1})$ it follows that $\mathfrak{P}^D_{\mathcal{Q}_{k+1}}(\psi_k) \in \mathcal{Q}_i$, so $\mathfrak{P}^D_{\mathcal{Q}_{k+1}}(\psi_k) \in \mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_k$. Hence, subprojection property holds for $\mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_k$ and $\mathcal{Q}_{k+1}$. Lemma 4.13 applied to $\mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_k$ and $\mathcal{Q}_{k+1}$ gives

$$\mathfrak{P}^D_{\mathcal{Q}_{k+1}}(\rho_k) = \underset{\phi \in (\mathcal{Q}_1 \cap \ldots \cap \mathcal{Q}_n) \cap \mathcal{Q}_{k+1}}{\arg\inf} \{D(\phi, \psi)\}. \tag{4.24}$$

This lemma holds for $k = 2$ by subprojection property of $\mathcal{Q}_1$ and $\mathcal{Q}_2$. $\qquad \square$

Remark 4.15. In what follows, we will assume that $\mathcal{N} = \mathfrak{B}(\mathcal{H})$ and $\phi, \psi, \omega, \rho \in \mathcal{M}(\mathcal{N}) = \mathfrak{B}(\mathcal{H})^+_{\star 1} \cong \mathfrak{G}_1(\mathcal{H})^+_1$ for some Hilbert space $\mathcal{H}$ of arbitrary dimension. The value of $n \in \mathbb{N}$ will be kept arbitrary but fixed.

Lemma 4.16. *If* $\mathrm{tr}(\psi(\log \rho - \log \psi)^2) < \infty$, *then the triangle equality at ρ for ψ with respect to D_0 is equivalent with*

$$\forall \psi \in \mathcal{M}(\mathcal{N}) \; \exists \rho \in \mathcal{Q} \; \forall \phi \in \mathcal{Q} \; \mathrm{tr}(\psi(\log \rho - \log \phi)) = \mathrm{tr}(\rho(\log \rho - \log \phi)). \tag{4.25}$$

Proof. If these distance functionals are finite, then

$$\mathrm{tr}(\psi(\log \psi - \log \phi)) = \mathrm{tr}(\rho(\log \rho - \log \phi)) + \mathrm{tr}(\psi(\log \psi - \log \rho)), \tag{4.26}$$

$$\mathrm{tr}(\psi(\log \rho - \log \phi)) = \mathrm{tr}(\rho(\log \rho - \log \phi)). \tag{4.27}$$

$\qquad \square$

Proposition 4.17. *Let* $\mathcal{Q} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid [\omega, P] = 0\}$, *and let* $\rho = P\psi P + (\mathbb{I} - P)\psi(\mathbb{I} - P)$, *where* $P \in \mathrm{Proj}(\mathcal{N})$. *Then*

$$\forall \phi \in \mathcal{Q} \; D_0(\phi, \psi) = D_0(\phi, \rho) + D_0(\rho, \psi). \tag{4.28}$$

Proof. The operators ϕ, ρ are block-diagonal matrices, so any functions of ρ and ϕ are also block-diagonal. Thus, $[\log \rho - \log \phi, P] = 0$ and $[\log \rho - \log \phi, \mathbb{I} - P] = 0$. It follows

$$(\log \rho - \log \phi) = (\log \rho - \log \phi)(P^2 + (\mathbb{I} - P)^2) = P(\log \rho - \log \phi)P + (\mathbb{I} - P)(\log \rho - \log \phi)(\mathbb{I} - P). \tag{4.29}$$

Substituting $P\rho P = P(P\psi P + (\mathbb{I} - P)\psi(\mathbb{I} - P))P = P\psi P$ we get

$$\mathrm{tr}(\psi P(\log \rho - \log \phi)P) = \mathrm{tr}(\rho P(\log \rho - \log \phi)P). \tag{4.30}$$

Similarly, we obtain

$$\mathrm{tr}(\psi(\mathbb{I} - P)(\log \rho - \log \phi)(\mathbb{I} - P)) = \mathrm{tr}(\rho(\mathbb{I} - P)(\log \rho - \log \phi)(\mathbb{I} - P)), \tag{4.31}$$

and it follows that

$$\mathrm{tr}(\psi(\log \rho - \log \phi)) = \mathrm{tr}(\rho(\log \rho - \log \phi)). \tag{4.32}$$

By Lemma 4.16, this gives (4.28), but it remains to check whether the assumption of this lemma is satisfied. We have

$$\left\|\rho^{1/2}\psi^{-1/2}\psi^{1/2}\right\|_{\mathfrak{G}_2(\mathcal{H})} = \mathrm{tr}(\rho) < \infty. \tag{4.33}$$

Now we want to show that $\left\|\rho^{-1/2}\psi\right\|_{\mathfrak{G}_2(\mathcal{H})}^2 < \infty$. Let us denote in matrix form

$$\psi = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}, \quad \rho = \begin{pmatrix} \psi_{11} & 0 \\ 0 & \psi_{22} \end{pmatrix}. \tag{4.34}$$

From $\psi \geq 0$ it follows that $\psi_{22}\xi = 0 \Rightarrow \psi_{12}\xi = 0$. Moreover,

$$\psi_{11} - \left(\psi_{22}^{-1/2}\psi_{21}\right)^* \left(\psi_{22}^{-1/2}\psi_{21}\right) = \psi_{11} - \psi_{12}\psi_{22}^{-1}\psi_{21}. \tag{4.35}$$

So, for $\xi \in \mathrm{dom}(\psi_{22}^{-1}\psi_{21})$, the corresponding forms satisfy

$$\langle \xi, \psi_{11}\xi\rangle_{\mathcal{H}} - \left\|\psi_{22}^{-1/2}\psi_{21}\xi\right\|_{\mathcal{H}}^2 = \left\langle \begin{pmatrix} \mathbb{I} \\ -\psi_{22}^{-1}\psi_{21} \end{pmatrix}\xi, \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}\begin{pmatrix} \mathbb{I} \\ -\psi_{22}^{-1}\psi_{21} \end{pmatrix}\xi \right\rangle_{\mathcal{H}} \geq 0. \tag{4.36}$$

Hence, as operators,

$$\psi_{11} - \psi_{12}\psi_{22}^{-1}\psi_{21} \geq 0, \tag{4.37}$$

$$\psi_{12}\psi_{22}^{-1}\psi_{21} \leq \psi_{11}, \tag{4.38}$$

$$\mathrm{tr}(\psi_{12}\psi_{22}\psi_{21}^{-1}) \leq \mathrm{tr}(\psi_{11}). \tag{4.39}$$

Hence

$$\left\|(\rho^{-\frac{1}{2}}\psi^{\frac{1}{2}})\psi^{\frac{1}{2}}\right\|_{\mathfrak{G}_2(\mathcal{H})}^2 = \mathrm{tr}(\psi\rho^{-1}\psi) = \mathrm{tr}\begin{pmatrix} \psi_{12}\psi_{22}^{-1}\psi_{21} + \psi_{11} & \psi_{12} + \psi_{12} \\ \psi_{21} + \psi_{21} & \psi_{21}\psi_{11}^{-1}\psi_{12} + \psi_{22} \end{pmatrix}$$

$$\leq 2\mathrm{tr}\begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix} = 2\mathrm{tr}(\psi) < \infty. \tag{4.40}$$

Using

$$\exists \lambda \in \mathbb{R}^+ \ \forall \gamma \in \mathbb{R}^+ \ \ |\log(\gamma)| \leq \lambda\max(\gamma^{1/2}, \gamma^{-1/2}), \tag{4.41}$$

together with (4.33) and (4.40), we obtain

$$\mathrm{tr}(\psi(\log\rho - \log\psi)^2) < \infty, \tag{4.42}$$

which follows from

$$\mathrm{tr}\left(\psi(\log\rho - \log\psi)^2\right) = \left\|(\log\rho - \log\psi)\psi^{1/2}\right\|_{\mathfrak{G}_2(\mathcal{H})}^2$$

$$\leq \left(\left\|\log(\rho)\psi^{1/2}\right\|_{\mathfrak{G}_2(\mathcal{H})} + \left\|\log(\psi)\psi^{1/2}\right\|_{\mathfrak{G}_2(\mathcal{H})}\right)^2 < \infty. \tag{4.43}$$

$\square$

Lemma 4.18. *Given* $P_i, P_j \in \mathrm{Proj}(\mathcal{N})$, *let* $\mathcal{Q}_k := \{\omega \in \mathcal{M}(\mathcal{N}) \mid [P_k, \omega] = 0\}$ *for* $k \in \{i, j\}$, *and* $[P_i, P_j] = 0$. *If* $\psi \in \mathcal{Q}_j$ *then*

$$\arg\inf_{\phi \in \mathcal{Q}_i} \{D_0(\phi, \psi)\} \in \mathcal{Q}_j. \tag{4.44}$$

14

Proof. From Proposition 4.17 it follows that

$$\arg\inf_{\phi \in \mathcal{Q}_i} \{D_0(\phi, \psi)\} = P_i \psi P_i + (\mathbb{I} - P_i)\psi(\mathbb{I} - P_i). \tag{4.45}$$

From $[\psi, P_j] = 0$ we obtain

$$[P_i \psi P_i + (\mathbb{I} - P_i)\psi(\mathbb{I} - P_i), P_j] = P_i[\psi, P_j]P_i + (\mathbb{I} - P_i)[\psi, P_j](\mathbb{I} - P_i). \tag{4.46}$$

$\square$

Lemma 4.19. *Let* $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathcal{N})$, $[P_i, P_j] = 0 \ \forall i, j \in \{1, \ldots, n\}$, *and* $\mathcal{Q} := \{\omega \in \mathcal{M}(\mathcal{N}) \mid [P_i, \omega] = 0 \ \forall i \in \{1, \ldots, n\}\}$. *Then* $\mathcal{Q}$ *satisfies triangle equality and*

$$\arg\inf_{\phi \in \mathcal{Q}} \{D_0(\phi, \psi)\} = \rho_n, \tag{4.47}$$

where $\rho_k = \arg\inf_{\phi \in \mathcal{Q}_k} \{D_0(\phi, \rho_{k-1})\}$ *for* $k \in \{1, \ldots, n\}$ *and* $\rho_0 = \psi$.

Proof. Follows directly from Proposition 4.14, Lemma 4.18 and Proposition 4.17. $\square$

Theorem 4.20. *If* $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathcal{N})$ *satisfies* $P_i P_j = \boldsymbol{\delta}_{ij} P_i \ \forall i, j \in \{1, \ldots, n\}$, $\sum_{i=1}^n P_i = \mathbb{I}$, *and*

$$\mathcal{Q}_{\mathrm{L}} := \{\omega \in \mathcal{M}(\mathcal{N}) \mid [P_i, \omega] = 0 \ \forall i \in \{1, \ldots, n\}\}, \tag{4.48}$$

then

$$\mathfrak{P}_{\mathcal{Q}_{\mathrm{L}}}^{D_0}(\psi) \equiv \arg\inf_{\phi \in \mathcal{Q}_{\mathrm{L}}} \{D_0(\phi, \psi)\} = \sum_{i=1}^n P_i \psi P_i. \tag{4.49}$$

Proof. By mathematical induction. Assume that

$$\rho_k = \sum_{i=1}^k P_i \psi P_i + \left(\mathbb{I} - \sum_{i=1}^k P_i\right) \psi \left(\mathbb{I} - \sum_{i=1}^k P_i\right). \tag{4.50}$$

Then

$$\rho_{k+1} = P_{k+1}\rho_k P_{k+1} + (\mathbb{I} - P_{k+1})\rho(\mathbb{I} - P_{k+1}) = P_{k+1}\psi P_{k+1} + \sum_{i=1}^k P_i \psi P_i + \left(\mathbb{I} - \sum_{i=1}^{k+1} P_i\right) \psi \left(\mathbb{I} - \sum_{i=1}^{k+1} P_i\right), \tag{4.51}$$

what follows from

$$(\mathbb{I} - P_{k+1})\left(\mathbb{I} - \sum_{i=1}^k P_i\right) = \left(\mathbb{I} - \sum_{i=1}^{k+1} P_i\right). \tag{4.52}$$

The first step of this induction is satisfied by Proposition 4.17. $\square$

Lemma 4.21. *If* $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathcal{N})$ *satisfies* $P_i P_j = \boldsymbol{\delta}_{ij} P_i \ \forall i, j \in \{1, \ldots, n\}$ *and* $\sum_{i=1}^n P_i = \mathbb{I}$, *then the conditions*

a) $\mathcal{Q} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid [P_i, \omega] = 0 \ \forall i \in \{1, \ldots, n\}\}$

b) $\mathcal{Q} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid \mathrm{tr}(\omega[f(\{P_i\}), x]) = 0 \ \forall x \in \mathcal{N} \ \forall f : \{1, \ldots, n\} \to \mathbb{C}\}$

are equivalent.

Proof. Using the property $\mathrm{tr}(\rho[P, x]) = \mathrm{tr}([P, \rho]x)$, we have

$$\mathrm{tr}(\rho[P, x]) = 0 \ \forall x \in \mathcal{N} \iff \mathrm{tr}([P, \rho]x) = 0 \ \forall x \in \mathcal{N} \iff [P, \rho] = 0. \tag{4.53}$$

Since there is finitely many projections, every function in a W^*-algebra $\mathcal{N}$ generated by these projections is their finite sum, and is equivalent to a function on n points. $\square$

Remark 4.22. As shown by Lemma 4.21, the weak Lüders rule (4.49) has no equivalent in the commutative case, because then the condition $\phi([x, y]) = 0$ is satisfied trivially for arbitrary $x, y \in \mathcal{N}$. Hence, one *cannot* interpret the weak Lüders rule as a noncommutative *generalisation* of the Bayes–Laplace rule.

Remark 4.23. Now we will generalise Theorem 4.20 to arbitrary W^*-algebras $\mathcal{N}$ and $\rho, \psi, \phi, \omega \in \mathcal{M}(\mathcal{N}) = \mathcal{N}_{\star 1}^+$. In order to prove this theorem, we will need also to generalise Lemma 4.16, Proposition 4.17, and Lemma 4.18. This will be provided, respectively, by Lemma 4.24, Proposition 4.26, and Lemma 4.27. Our main tool will be the expression for $D_1(\psi, \phi) = D_0(\phi, \psi)$ in terms of Connes' cocycle, introduced by Petz in [103].

Lemma 4.24. *Given $\rho, \psi \in \mathcal{M}(\mathcal{N})$, consider a GNS representation $(\mathcal{H}_\psi, \pi_\psi, \Omega_\psi)$. If $\Omega_\psi \in \mathrm{dom}(\log(\Delta_{\rho,\psi}))$, then the triangle equality for ψ at ρ with respect to D_0 is equivalent to*

$$\forall \psi \in \mathcal{M}(\mathcal{N}) \;\; \exists \rho \in \mathcal{Q} \;\; \forall \phi \in \mathcal{Q} \;\; \mathrm{i} \lim_{t \to {}^+0} (\psi(\mathbb{I} - [\phi : \rho]_t)) = \mathrm{i} \lim_{t \to {}^+0} (\rho(\mathbb{I} - [\phi : \rho]_t)). \tag{4.54}$$

Proof. If $\Omega_\psi \in \mathrm{dom}(\log(\Delta_{\rho,\psi}))$, then

$$\lim_{t \to {}^+0} \left\| \frac{1}{t}(\Delta_{\rho,\psi}^{\mathrm{it}} - \mathbb{I})\Omega_\psi \right\|_{\mathcal{H}_\psi} = \lambda \leq \infty. \tag{4.55}$$

Moreover, given

$$x(t) := \mathbb{I} - \pi_\psi([\phi : \rho]_t), \tag{4.56}$$

we have

$$\forall t \in \mathbb{R} \;\; \|x(t)\|_{\mathcal{H}_\psi} \leq 2, \tag{4.57}$$

$$\lim_{t \to {}^+0} x(t)\Omega_\psi = 0, \tag{4.58}$$

$$\lim_{t \to {}^+0} \|x(t)\Omega_\psi\|_{\mathcal{H}_\psi} = 0. \tag{4.59}$$

Hence,

$$\lim_{t \to {}^+0} \left| \left\langle x(t)\Omega_\psi, \frac{1}{t}(\Delta_{\rho,\psi}^{\mathrm{it}} - \mathbb{I})\Omega_\psi \right\rangle_\psi \right| \leq \lim_{t \to {}^+0} \|x(t)\Omega_\psi\|_{\mathcal{H}_\psi} \left\| \frac{1}{t}(\Delta_{\rho,\psi}^{\mathrm{it}} - \mathbb{I})\Omega_\psi \right\|_{\mathcal{H}_\psi} \leq 0 \cdot \lambda = 0. \tag{4.60}$$

So,

$$\begin{aligned}
0 &= \lim_{t \to {}^+0} \left\langle \Omega_\psi, \frac{1}{t} (\mathbb{I} - \pi_\psi([\phi : \rho]_t)) \left(\Delta_{\rho,\psi}^{\mathrm{it}} - \mathbb{I}\right) \Omega_\psi \right\rangle_\psi \\
&= \lim_{t \to {}^+0} \left\langle \Omega_\psi, \frac{1}{t} (\mathbb{I} - \pi_\psi([\phi : \rho]_t)) (\pi_\psi([\rho : \psi]_t) - \mathbb{I}) \Omega_\psi \right\rangle_\psi \\
&= \lim_{t \to {}^+0} \frac{1}{t} \psi \left((\mathbb{I} - [\phi : \rho]_t)([\rho : \psi]_t - \mathbb{I})\right),
\end{aligned} \tag{4.61}$$

where the second equation follows from the property $\Delta_{\rho,\psi}^{\mathrm{it}} \xi_\pi(\psi) = [\rho : \psi]_t \xi_\pi(\psi) \;\; \forall t \in \mathbb{R}$ for a standard representative $\xi_\pi(\psi) \in \mathcal{H}^\natural$ of ψ in a standard representation $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$, which in this case is given by the cyclic vector Ω_ψ of the GNS representation $(\mathcal{H}_\psi, \pi_\psi, \Omega_\psi)$. The triangle equality (4.11) reads

$$\mathrm{i} \lim_{t \to {}^+0} \frac{\psi}{t} ([\rho : \psi]_t - \mathbb{I}) + \mathrm{i} \lim_{t \to {}^+0} \frac{\rho}{t} ([\phi : \rho]_t - \mathbb{I}) = \mathrm{i} \lim_{t \to {}^+0} \frac{\psi}{t} ([\phi : \psi]_t - \mathbb{I}), \tag{4.62}$$

16

and is equivalent to

$$i \lim_{t\to+0} \frac{\psi}{t} \left([\rho:\psi]_t - [\phi:\psi]_t\right) = i \lim_{t\to+0} \frac{\rho}{t} \left(\mathbb{I} - [\phi:\rho]_t\right) \tag{4.63}$$

It remains to calculate

$$\begin{aligned}
&i \lim_{t\to+0} \frac{1}{t} \left(\psi([\rho:\psi]_t - [\phi:\psi]_t)\right) = \\
&i \lim_{t\to+0} \frac{1}{t} \left(\psi([\rho:\psi]_t - [\phi:\rho]_t[\rho:\psi]_t)\right) = \\
&i \lim_{t\to+0} \frac{1}{t} \left(\psi((\mathbb{I} - [\phi:\rho]_t)[\rho:\psi]_t)\right) = \\
&i \lim_{t\to+0} \frac{1}{t} \left(\psi(\mathbb{I} - [\phi:\rho]_t) + \psi((\mathbb{I} - [\phi:\rho]_t)([\rho:\psi]_t - \mathbb{I}))\right) = \\
&i \lim_{t\to+0} \frac{1}{t} \left(\psi(\mathbb{I} - [\phi:\rho]_t)\right).
\end{aligned} \tag{4.64}$$

$\square$

Lemma 4.25. *If $P \in \mathrm{Proj}(\mathcal{N})$ and $\phi \in \mathcal{M}(\mathcal{N})$ then $\phi([P,x]) = 0 \ \forall x \in \mathcal{N}$ iff ϕ is block-diagonal, that is, iff $\phi(Px(\mathbb{I} - P)) = \phi((\mathbb{I} - P)xP) = 0$.*

Proof. If $\phi([P,x]) = 0$, then $\phi(Px) = \phi(xP)$, so, for $x =: y(\mathbb{I} - P)$,

$$\phi(Py(\mathbb{I} - P)) = \phi(y(\mathbb{I} - P)P) = 0, \tag{4.65}$$

and similarly $\phi((\mathbb{I} - P)yP) = 0$. Conversely, every $y \in \mathcal{N}$ has the form

$$y = PyP + (\mathbb{I} - P)yP + Py(\mathbb{I} - P) + (\mathbb{I} - P)y(\mathbb{I} - P). \tag{4.66}$$

Hence,

$$[P,y] = PyP + Py(\mathbb{I} - P) - PyP - (\mathbb{I} - P)yP = Py(\mathbb{I} - P)yP, \tag{4.67}$$
$$\phi([P,y]) = \phi(Py(\mathbb{I} - P) - (\mathbb{I} - P)yP) = 0. \tag{4.68}$$

$\square$

Proposition 4.26. *Let $\mathcal{Q} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid \omega([P,x]) = 0 \ \forall x \in \mathcal{N}\}$, and let $\rho = \psi(P \cdot P) + \psi((\mathbb{I} - P) \cdot (\mathbb{I} - P))$, where $P \in \mathrm{Proj}(\mathcal{N})$. Then*

$$\forall \phi \in \mathcal{Q} \quad D_0(\phi, \psi) = D_0(\phi, \rho) + D_0(\rho, \psi). \tag{4.69}$$

Proof. We will use block decomposition of $\mathcal{N}$ into

$$\mathcal{N} = \begin{pmatrix} P\mathcal{N}P & P\mathcal{N}(\mathbb{I} - P) \\ (\mathbb{I} - P)\mathcal{N}P & (\mathbb{I} - P)\mathcal{N}(\mathbb{I} - P) \end{pmatrix}, \tag{4.70}$$

together with the corresponding notation

$$\forall x \in \mathcal{N} \quad x = PxP + Px(\mathbb{I} - P) + (\mathbb{I} - P)xP + (\mathbb{I} - P)x(\mathbb{I} - P)$$
$$=: x_{11} + x_{21} + x_{12} + x_{22} =: \begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix}, \tag{4.71}$$

$$\forall x \in \mathcal{N} \ \forall \phi \in \mathcal{N}_\star^+ \quad \phi(x) = \phi\begin{pmatrix} x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix} =: \phi_{11}(x_{11}) + \phi_{21}(x_{21}) + \phi_{12}(x_{12}) + \phi_{22}(x_{22})$$
$$=: \begin{pmatrix} \phi_{11} & \phi_{21} \\ \phi_{12} & \phi_{22} \end{pmatrix}(x). \tag{4.72}$$

17

By Lemma 4.25, the states ϕ, ρ are block-diagonal,

$$\phi = \begin{pmatrix} \phi_{11} & 0 \\ 0 & \phi_{22} \end{pmatrix}, \quad \rho = \begin{pmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{pmatrix}. \tag{4.73}$$

Recall from (3.22)-(3.23) that for a given standard representation of $\mathcal{N}$ on a Hilbert space $\mathcal{H}$, and $\xi \in \mathcal{H}$ such that $\xi \perp [\mathcal{N}\Omega_\rho]$, the relative modular operator $\Delta_{\phi,\rho}$ is defined as $\Delta_{\phi,\rho} := R_{\phi,\rho}^* \bar{R}_{\phi,\rho}$, where [9]

$$R_{\phi,\rho}(x\Omega_\rho + \xi) = (\text{supp}(\phi))x^*\Omega_\phi. \tag{4.74}$$

From

$$\langle x\Omega_\rho, P\xi \rangle_\mathcal{H} = \langle Px\Omega_\rho, \xi \rangle_\mathcal{H} = 0 \ \ \forall x \in \mathcal{N} \tag{4.75}$$

we have $P\xi \perp [\mathcal{N}\Omega_\rho]$, and, analogously, $(\mathbb{I} - P)\xi \perp [\mathcal{N}\Omega_\rho]$, so, because $P \in \mathcal{N}$, it preserves the dense domain. From

$$\begin{aligned} \langle (\mathbb{I} - P)(x_1\Omega_\rho + \xi_1), \Delta_{\phi,\rho}P(x_2\Omega_\rho + \xi_2) \rangle_\mathcal{H} &= \langle \bar{R}_{\phi,\rho}(\mathbb{I} - P)(x_1\Omega_\rho + \xi_1), \bar{R}_{\phi,\rho}P(x_2\Omega_\rho + \xi_2) \rangle_\mathcal{H} \\ &= \langle \bar{R}_{\phi,\rho}((\mathbb{I} - P)x_1\Omega_\rho + (\mathbb{I} - P)\xi_1), \bar{R}_{\phi,\rho}(Px_2\Omega_\rho + P\xi_2) \rangle_\mathcal{H} \\ &= \langle \text{supp}(\phi)x_1^*(\mathbb{I} - P)\Omega_\phi, \text{supp}(\phi)x_2^*P\Omega_\phi \rangle_\mathcal{H} \\ &= \phi((\mathbb{I} - P)x_1\text{supp}(\phi)x_2^*P) \\ &= 0 \end{aligned} \tag{4.76}$$

we have that $[P, \Delta_{\phi,\rho}] = 0$ on the dense domain. Since P is bounded, it follows that $\Delta_{\phi,\rho}$ preserves the decomposition $\text{ran}(P) \oplus \text{ran}(\mathbb{I} - P)$. Hence, $\Delta_{\phi,\rho}$ is block diagonal. The same holds for $\Delta_\rho = \Delta_{\rho,\rho}$. Thus, $[\phi : \rho]_t = \Delta_{\phi,\rho}^{it}\Delta_\rho^{-it}$ is also block diagonal,

$$[\phi : \rho]_t = \begin{pmatrix} ([\phi : \rho]_t)_{11} & 0 \\ 0 & ([\phi : \rho]_t)_{22} \end{pmatrix}, \tag{4.77}$$

$$\mathbb{I} - [\phi : \rho]_t = P(\mathbb{I} - [\phi : \rho]_t)P + (\mathbb{I} - P)(\mathbb{I} - [\phi : \rho]_t)(\mathbb{I} - P). \tag{4.78}$$

From $\psi(PxP + (\mathbb{I} - P)x(\mathbb{I} - P)) = \rho(PxP + (\mathbb{I} - P)x(\mathbb{I} - P)) \ \forall x \in \mathcal{N}$, and we obtain

$$\psi(\mathbb{I} - [\phi : \rho]_t) = \rho(\mathbb{I} - [\phi : \rho]_t). \tag{4.79}$$

Hence, by Lemma 4.24, triangle equality holds if $\Omega_\psi \in \text{dom}(\log(\Delta_{\rho,\psi}))$. It remains to check whether this condition is satisfied. Araki and Masuda [10] and Donald [42] prove this relation, but under stronger conditions (Araki and Masuda assume that ρ is faithful, while Donald assumes also faithfulness of ψ), so we need to provide more general proof. Let $\xi \perp [\mathcal{N}\Omega_\psi]$, and consider operators R_ψ, $R_{\rho,\psi}$ and A given by

$$R_\psi(x\Omega_\psi + \xi) := (\text{supp}(\psi))x^*\Omega_\psi, \tag{4.80}$$

$$R_{\rho,\psi}(x\Omega_\psi + \xi) := (\text{supp}(\psi))x^*\Omega_\psi, \tag{4.81}$$

$$A(x\Omega_\psi) := x\Omega_\rho. \tag{4.82}$$

Then

$$\|A(x\Omega_\psi)\|_{\mathcal{H}_\psi}^2 = \rho(x^*x) \leq \lambda\psi(x^*x) = \lambda\|x\Omega_\psi\|_{\mathcal{H}_\psi}^2, \tag{4.83}$$

where $\lambda \in \mathbb{R}^+$, so A is bounded. From $R_{\rho,\psi} = AR_\psi$, $J_\psi\Delta_\psi^{-1/2} = R_\psi$, and $J_{\rho,\psi}\Delta_{\rho,\psi}^{-1/2} = R_{\rho,\psi}$ we obtain

$$\Delta_{\rho,\psi}^{-1/2} = J_{\rho,\psi}^{-1}AJ_\psi\Delta_\psi^{-1/2}, \tag{4.84}$$

where $J_{\rho,\psi}^{-1}AJ_\psi$ is bounded. So

$$\Omega_\psi \in \text{dom}(\Delta_\psi^{-1/2}) \subseteq \text{dom}(\Delta_{\rho,\psi}^{-1/2}). \tag{4.85}$$

18

Araki [8] proved that $\Omega_\rho = (\Delta_{\rho,\psi})^{1/2}\Delta_\psi$ and $\Delta_\psi \in \mathrm{dom}(\Delta_{\rho,\psi}^{1/2})$. Hence

$$\Omega_\psi \in \mathrm{dom}(\Delta_{\rho,\psi}^{1/2}) \cap \mathrm{dom}(\Delta_{\rho,\psi}^{-1/2}). \tag{4.86}$$

On the other hand,

$$\exists \lambda \in \mathbb{R}^+ \ \forall \gamma \in \mathbb{R}^+ \ \ |\log(\gamma)| \le \lambda \max(\gamma^{1/2}, \gamma^{-1/2}), \tag{4.87}$$

so, for every positive operator x on a Hilbert space $\mathcal{H}$,

$$\mathrm{dom}(\log x) \supseteq \mathrm{dom}(x^{1/2}) \cap \mathrm{dom}(x^{-1/2}). \tag{4.88}$$

It follows that $\Omega_\psi \in \mathrm{dom}(\log(\Delta_{\rho,\psi}))$. $\qquad\square$

Lemma 4.27. *Given $P_i, Pj \in \mathrm{Proj}(\mathcal{N})$, let $\mathcal{Q}_k = \{\omega \in \mathcal{M}(\mathcal{N}) \mid \omega([P_k, x]) = 0 \ \forall x \in \mathcal{N}\}$ for $k \in \{i, j\}$, and $[P_i, P_j] = 0$. If $\psi \in \mathcal{Q}_j$, then*

$$\arg\inf_{\phi \in \mathcal{Q}_i} \{D_0(\phi, \psi)\} \in \mathcal{Q}_j. \tag{4.89}$$

Proof. From Proposition 4.26 it follows that

$$\arg\inf_{\phi \in \mathcal{Q}_i} \{D_0(\phi, \psi)\} = \psi(P(\cdot)P) + \psi((\mathbb{I} - P)(\cdot)(\mathbb{I} - P)). \tag{4.90}$$

Denote the right hand side of (4.90) by φ. We need to check that $\psi([P_j, x]) = 0 \ \forall x \in \mathcal{N} \Rightarrow \varphi([P_j, x]) = 0 \ \forall x \in \mathcal{N}$. But this follows from

$$\begin{aligned}
\varphi([P_j, x]) &= \psi(P_i[P_j, x]P_i + (\mathbb{I} - P_i)[P_j, x](\mathbb{I} - P_i)) \\
&= \psi([P_j, P_i x P_i + (\mathbb{I} - P_i)x(\mathbb{I} - P_i)]) \\
&= 0.
\end{aligned} \tag{4.91}$$

$\qquad\square$

Theorem 4.28. *If $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathcal{N})$ satisfies $P_i P_j = \boldsymbol{\delta}_{ij} P_i \ \forall i, j \in \{1, \dots, n\}$, $\sum_{i=1}^n P_i = \mathbb{I}$, and*

$$\mathcal{Q}_{\mathrm{L}} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid \omega([P_i, x]) = 0 \ \forall x \in \mathcal{N} \ \forall i \in \{1, \dots, n\}\}, \tag{4.92}$$

then

$$\mathfrak{P}_{\mathcal{Q}_{\mathrm{L}}}^{D_0}(\psi) = \sum_{i=1}^n P_i \psi P_i. \tag{4.93}$$

Proof. The same as for Theorem 4.20, with the first step of induction satisfied by Proposition 4.26, and using Lemma 4.27 instead of Lemma 4.18. $\qquad\square$

Theorem 4.29. *Let $\mathcal{N} = \mathfrak{B}(\mathcal{H})$ and $\mathcal{M}(\mathcal{N}) = \mathfrak{B}(\mathcal{H})_{\star 1}^+$. If $\psi \in \mathcal{M}(\mathcal{N})$, $\{P_i\}_{i=1}^n \subseteq \mathrm{Proj}(\mathcal{N})$ satisfy $P_i P_j = \boldsymbol{\delta}_{ij} P_i \ \forall i, j \in \{1, \dots, n\}$, $\sum_{i=1}^n P_i = \mathbb{I}$, and $\mathrm{tr}(\psi P_i) \ne 0 \ \forall i \in \{1, \dots, n\}$, if $\lambda_i \in \mathbb{R}$ satisfy $\sum_{i=1}^n \lambda_i = 1$, and*

$$\mathcal{Q}_{\mathrm{qJ}} := \{\omega \in \mathfrak{G}_1(\mathcal{H})_1^+ \mid \mathrm{tr}(\omega P_i) = \lambda_i\} \tag{4.94}$$

then

$$\mathfrak{P}_{\mathcal{Q}_{\mathrm{qJ}}}^{D_0}(\psi) = \sum_{i=1}^n \lambda_i \frac{P_i \psi P_i}{\mathrm{tr}(\psi P_i)}. \tag{4.95}$$

Proof. Let $\rho = \bigoplus_{i=1}^n \rho_i$, $\phi = \bigoplus_{i=1}^n \phi_i$, and $\rho_i = \lambda_i \frac{P_i \psi P_i}{\mathrm{tr}(P_i \psi_i P_i)}$, $\mathrm{tr}(\phi_i) = \lambda_i \ \forall i \in \{1, \ldots, n\}$. Then, using the fact that each function of block diagonal matrices is block diagonal, we obtain

$$\log \rho - \log \phi = \bigoplus_{i=1}^n \left(\log \rho_i - \log \phi_i\right) =: \bigoplus_{i=1}^n C_i. \tag{4.96}$$

Using $\mathrm{tr}(\frac{\rho_i}{\lambda_i}) = 1 = \mathrm{tr}(\frac{\phi_i}{\lambda_i})$, we obtain

$$
\begin{aligned}
\mathrm{tr}\left(\psi \bigoplus_{i=1}^n C_i\right) &= \sum_{i=1}^n \mathrm{tr}\left(P_i \psi P_i C_i\right) = \sum_{i=1}^n \mathrm{tr}\left(P_i \psi P_i \left(\log \frac{\rho_i}{\lambda_i} - \log \frac{\phi_i}{\lambda_i}\right)\right) \\
&= \sum_{i=1}^n \mathrm{tr}(P_i \psi P_i) \mathrm{tr}\left(\frac{P_i \psi P_i}{\mathrm{tr}(P_i \psi P_i)}\left(\log \frac{\rho_i}{\lambda_i} - \log \frac{\phi_i}{\lambda_i}\right)\right) \\
&= \sum_{i=1}^n \mathrm{tr}(P_i \psi P_i) \mathrm{tr}\left(\widetilde{\rho}_i \left(\log \widetilde{\rho}_i - \log \widetilde{\phi}_i\right)\right) \\
&= \sum_{i=1}^n \mathrm{tr}(P_i \psi P_i) D_0(\widetilde{\phi}_i, \widetilde{\rho}_i) \geq 0, \tag{4.97}
\end{aligned}
$$

where $\widetilde{\rho}_i := \frac{\rho_i}{\lambda_i}$, $\widetilde{\phi}_i := \frac{\phi_i}{\lambda_i}$, and equality is attained iff $\widetilde{\phi}_i = \widetilde{\rho}_i$. Hence,

$$\mathrm{tr}(\psi(\log \rho - \log \phi)) \geq 0, \tag{4.98}$$
$$\mathrm{tr}(\psi(\log \psi - \log \rho)) \leq \mathrm{tr}(\psi(\log \psi - \log \phi)). \tag{4.99}$$

From the condition for equality in (4.97), it follows that ρ is the unique minimiser of $D_0(\cdot, \psi)$. $\square$

Remark 4.30. The strong Lüders rule (4.2) can be obtained from minimisation of D_0 by two different methods. First method amounts to applying quantum Jeffrey's rule (4.95) and taking the limit $\lambda_2, \ldots, \lambda_n \to 0$,

$$\lim_{\lambda_2, \ldots, \lambda_n \to 0} \mathfrak{P}_{\mathcal{Q}_{\mathrm{qJ}}}^{D_0}(\psi) = \frac{P_1 \psi P_1}{\mathrm{tr}(\psi P_1)}. \tag{4.100}$$

Note that (4.95) is a weakly continuous function of λ_i. The limit (4.100) is also weakly continuous. Hence, the strong Lüders rule can be considered as a weakly continuous limit of an entropic projection $\mathfrak{P}^{D_0}$.

Remark 4.31. Despite the observation carried in Remark 4.5, the direct derivation of (4.2) along the lines of Theorem 4.29 with the initial constraints $\lambda_2 = \ldots = \lambda_n = 0$ is impossible, because in such case the necessary assumptions for this theorem do not hold. More precisely, the states ω that satisfy

$$-\mathrm{tr}\left(\omega \log(P \rho P - (\mathbb{I} - P)\rho(\mathbb{I} - P))\right) < \infty \tag{4.101}$$

do not exist if $\mathrm{tr}(\rho P) = 0$. This follows from

$$
\begin{aligned}
-\mathrm{tr}\left(\omega \left(\log|_{P\mathcal{N}P}(P\rho P) + \log|_{(\mathbb{I}-P)\rho(\mathbb{I}-P)}(\mathbb{I} - P)\rho(\mathbb{I} - P)\right)\right) &= -\mathrm{tr}\left(\omega \log|_{P\mathcal{N}P}(P\rho P)\right) \\
&= -\mathrm{tr}(\omega \log 0) = -\infty. \tag{4.102}
\end{aligned}
$$

This situation can be improved by 'regularisation' of the difference of two distance functionals, using Connes' cocycle with respect to some well-behaved 'reference' functional $\omega_0 \in \mathcal{M}(\mathcal{N})$. The natural choice in the case of strong Lüders' rule is $\omega_0 = \rho|_{P\mathcal{N}P}$. (Donald [40, 41] also introduces a distance functional that is dependent on the choice of a subset of C^*-algebra, but his motivation as well as the resulting mathematical construction differ from ours.) Hence, in order to show that the strong Lüders rule is an entropic projection without any limits involved, we will use an analogue of a Theorem 4.28 for a single projection P and a *regularised* distance functional $D_0^P(\phi, \psi)$, defined as a restriction of $D_0(\phi, \psi)$ to a subspace $P\mathcal{N}P$.

Lemma 4.32.
$$D_0^P(\omega, \rho) := \mathrm{i} \lim_{t \to {}^+0} \rho|_{P\mathcal{N}P} \left([\omega|_{P\mathcal{N}P} : \rho|_{P\mathcal{N}P}]_t - \mathbb{I}|_{P\mathcal{N}P}\right). \tag{4.103}$$

is a distance functional on $(P\mathcal{N}P)_{\star 1}^+$, *that is,*

$$D_0^P(\omega, \rho) \geq 0 \quad \forall \rho, \omega \in (P\mathcal{N}P)_{\star 1}^+, \tag{4.104}$$

$$D_0^P(\omega, \rho) = 0 \iff \rho|_{P\mathcal{N}P} = \omega|_{P\mathcal{N}P}. \tag{4.105}$$

Proof. Follows directly from the definition and properties of $D_0|_{\mathcal{N}_{\star 1}^+}$. $\square$

Lemma 4.33. *If* $\mathrm{tr}(\omega P) = \mathrm{tr}(\omega) = 1$ *then* $[\omega, P] = 0$ *and* $\omega = P\omega P$.

Proof. We need to check that $\mathrm{tr}((\mathbb{I} - P)\omega P) = 0$, which is equivalent to $\omega((\mathbb{I} - P)xP) = 0$ $\forall x \in \mathcal{N}$. But this follows from

$$|\omega((\mathbb{I} - P)x)|^2 \leq \omega(x^*x)\omega((\mathbb{I} - P)^2) = 0. \tag{4.106}$$

Similarly $\omega(x(\mathbb{I} - P)) = 0 \ \forall x \in \mathcal{N}$. $\square$

Theorem 4.34. *If* $\mathcal{Q}_{\mathrm{sL}} := \{\omega \in \mathcal{M}(\mathcal{N}) \mid \mathrm{tr}(\omega P) = \mathrm{tr}(\omega) = 1\}$, *and* $\psi \in \mathcal{M}(\mathcal{N})$ *satisfies* $\mathrm{tr}(P\psi) \neq 0$, *then*

$$\mathfrak{P}_{\mathcal{Q}_{\mathrm{sL}}}^{D_0^P}(\psi) \equiv \operatorname*{arg\,inf}_{\phi \in \mathcal{Q}_{\mathrm{sL}}} \left\{D_0^P(\phi, \psi)\right\} = \frac{P\psi P}{\mathrm{tr}(P\psi)}. \tag{4.107}$$

Proof. Follows from Theorem 4.29, applied to $(P\mathcal{N}P)_{\star 1}^+$, if we notice that

$$\mathcal{Q}_{\mathrm{sL}} = \{\omega \in \mathcal{M}(\mathcal{N}) \mid \mathrm{tr}(\omega) = 1, \ \omega = P\omega P\}, \tag{4.108}$$

which follows from Lemma 4.33. $\square$

5 Minimisation of quantum metrical distances

In the Hilbert space based quantum mechanics a ***transition probability*** of $\xi_0, \xi_1 \in \mathcal{H}$ such that $\|\xi_0\| = \|\xi_1\| = 1$ is defined as

$$\mathrm{TP}_{\mathcal{H}}(\xi_1, \xi_0) := |\langle \xi_1, \xi_0 \rangle|^2 \in [0, 1]. \tag{5.1}$$

Different generalisations of the notion of transition probability to the case of quantum states over W^*-algebras are possible. The two most important are: the ***Raggio transition probability*** [104]

$$\mathrm{TP}_{\mathrm{R}}(\phi, \psi) := \langle \xi_\pi(\phi), \xi_\pi(\psi) \rangle_{\mathcal{H}} = \frac{1}{2}\left(\phi(\mathbb{I}) + \psi(\mathbb{I}) - \|\xi_\pi(\phi) - \xi_\pi(\psi)\|_{\mathcal{H}}^2\right) \quad \forall \phi, \psi \in \mathcal{N}_\star^+, \tag{5.2}$$

where $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$ is a standard representation of $\mathcal{N}$, and the ***Cantoni–Uhlmann transition probability*** [21, 114]

$$\mathrm{TP}_{\mathrm{CU}}(\phi, \psi) := \sup_{(\mathcal{H}, \pi)} \left\{|\langle \zeta_\pi(\phi), \zeta_\pi(\psi) \rangle_{\mathcal{H}}|^2\right\} \quad \forall \phi, \psi \in \mathcal{N}_\star^+, \tag{5.3}$$

where $\zeta_\pi(\omega) \in \mathcal{H}$ is defined by $\omega(x) = \langle \zeta_\pi(\omega), \pi(x)\zeta_\pi(\omega) \rangle_{\mathcal{H}}$ $\forall x \in \mathcal{N}$ for some representation $(\mathcal{H}, \pi)$ of $\mathcal{N}$, and the supremum varies over all possible representations. For the comparison of (5.2) with (5.3) and with some other possibilities, see [104, 105, 1, 124, 125].

From the geometric perspective it is worth noting that (5.2) is bijectively related to the distance on $\mathcal{N}_\star^+$ defined by the norm of $\mathcal{H} \cong L_2(\mathcal{N})$,

$$d_{L_2(\mathcal{N})}(\phi, \psi) = \|\xi_{\pi_\mathcal{N}}(\phi) - \xi_{\pi_\mathcal{N}}(\psi)\|_{L_2(\mathcal{N})} = \sqrt{\phi(\mathbb{I}) + \psi(\mathbb{I}) - 2\mathrm{TP}_\mathrm{R}(\phi, \psi)}, \qquad (5.4)$$

as well as to the $D_{1/2}$ distance (3.29) on $\mathcal{N}_\star^+$,

$$\begin{aligned} D_{1/2}(\phi, \psi) &= 2(\phi + \psi)(\mathbb{I}) - 4\langle \xi_\pi(\phi), \xi_\pi(\psi)\rangle_\mathcal{H} \\ &= 2(\phi(\mathbb{I}) + \psi(\mathbb{I})) - 4\mathrm{TP}_\mathrm{R}(\phi, \psi) = 2\|\xi_{\pi_\mathcal{N}}(\phi) - \xi_{\pi_\mathcal{N}}(\psi)\|^2_{L_2(\mathcal{N})}, \qquad (5.5) \end{aligned}$$

while (5.3) is bijectively related to the ***Bures distance*** on $\mathcal{N}_\star^+$ [20] (cf. also [3, 5]) defined by

$$d_{\mathrm{Bures}}(\phi, \psi) = \inf_{(\mathcal{H}, \pi)} \{\|\zeta_\pi(\phi) - \zeta_\pi(\psi)\|_\mathcal{H}\} = \sqrt{\phi(\mathbb{I}) + \psi(\mathbb{I}) - 2\sqrt{\mathrm{TP}_{\mathrm{CU}}(\phi, \psi)}}, \qquad (5.6)$$

where ζ_π is defined as above, and inf varies over the same range as sup in (5.3). In the notation above we have used freely the unitary equivalence between any standard representation $(\mathcal{H}, \pi, J, \mathcal{H}^\natural)$ of a W^*-algebra $\mathcal{N}$ and its canonical representation $(L_2(\mathcal{N}), \pi_\mathcal{N}, J_\mathcal{N}, L_2(\mathcal{N})^+)$, see [76, 79]. Both $d_{L_2(\mathcal{N})}(\phi, \psi)$ and $d_{\mathrm{Bures}}(\phi, \psi)$ are metrical distances.

Herbut [64] proved that

$$\mathfrak{P}^{d_{L_2(\mathcal{N})}}_{\mathcal{Q}_\mathrm{L}}(\psi) = \sum_{i=1}^n P_i \rho_\psi P_i \qquad (5.7)$$

for $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, $\rho_\phi, \rho_\psi \in \mathfrak{G}_1(\mathcal{H})^+$, and $\mathcal{Q}_\mathrm{L}$ given by (4.48). This derivation of the weak Lüders rule was the first result of this type in the literature. Minimisation of the same distance function, but under constraints of the type $\mathrm{tr}(\rho_\phi x) = \lambda$ with $x \in \mathfrak{B}(\mathcal{H})^\mathrm{sa}$ and $\lambda \in \mathbb{R}$, was later considered in [34], however with no general results.

Let $\mathcal{N}_0 \subseteq \mathcal{N}$ be the W^*-subalgebras of $\mathfrak{B}(\mathcal{H})$, let $T \in \mathcal{N}^+$ be invertible with $0 < T \leq \mathbb{I}$, and let $\psi \in \mathcal{N}_{\star 1}^+$. Marchand and collaborators [85, 89, 56] considered a quantum inference problem based on

$$\arg\inf_{\omega \in K} \{d_{\mathrm{Bures}}(\omega, \psi)\} = \arg\sup_{\omega \in K} \{\mathrm{TP}_{\mathrm{CU}}(\omega, \psi)\}, \qquad (5.8)$$

with $K = \{\phi \in \mathcal{N}_{\star 1}^+ \mid \phi|_{\mathcal{N}_0} = \psi(T \cdot T)\}$. The algebra $\mathcal{N}_0$ is interpreted as representing operators subjected to a "partial measurement", while T is a noncommutative analogue of the Radon–Nikodým quotient that follows from Sakai's theorem [108] and can be thought of as a generalisation of a projection.[9] They derived in [14, 12, 87] several different "post-measurement" states, dependent on the choice of $\mathcal{N}_0$, T, and initial correlations in ψ.

It was shown by Raggio [105] that the strong Lüders rule (4.2) can be directly derived as a special case of constrained *maximisation* of the Cantoni–Uhlmann transition probability (5.3). Let Y be a convex subset of a real topological vector space X. A subset $F \subseteq Y$ is called a ***face*** iff

$$\forall x \in F \ \exists n \in \mathbb{N} \ \left(\exists \{\lambda_i\}_{i=1}^n \subseteq \mathbb{R}^+ \ x = \sum_{i=1}^n \lambda_i x_i, \ \sum_{i=1}^n \lambda_i = 1\right) \ \Rightarrow \ \{x_i\}_{i=1}^n \subseteq F. \qquad (5.9)$$

Let $\mathcal{N}$ be a W^*-algebra, and let K be a closed, convex subset of $\mathcal{N}_{\star 1}^+$ such that

$$(\lambda\omega + (1-\lambda)\phi \in K \ \forall \lambda \in [0, 1] \ \Rightarrow \ \omega, \phi \in K) \ \forall \omega, \phi \in \mathcal{N}_{\star 1}^+. \qquad (5.10)$$

---

[9]For a generalisation to a setting based on C^*-algebras, see [54]. For a generalisation that does not require a subset $\mathcal{N}_0$ to be a W^*-algebra, see [87].

Such set is a face in $\mathcal{N}_{\star 1}^+$. For each face $K \subseteq \mathcal{N}_{\star 1}^+$ there exists a unique $P \in \text{Proj}(\mathcal{N})$ such that $\omega \in K \iff \omega(P) = 1$. If $\mathcal{Q}_P$ is a face in $\mathcal{N}_{\star 1}^+$ with a corresponding $P \in \text{Proj}(\mathcal{N})$, $\psi \in \mathcal{N}_{\star 1}^+$, $\mathcal{N} = \mathfrak{B}(\mathcal{H})$, and $\text{tr}(\rho_\psi \cdot) \equiv \psi$, then [105]

$$\frac{P\rho_\psi P}{\text{tr}(P\rho_\psi)} = \arg\sup_{\omega \in \mathcal{Q}_P} \{\text{TP}_{\text{CU}}(\omega, \psi)\}. \tag{5.11}$$

This corresponds to Domotor's observation [38] that the faces in $L_1(\mathcal{X}, \mho(\mathcal{X}), \tilde{\mu})_1^+$ form the correct constraints for the Bayes–Laplace rule. On the other hand, (5.5) gives us

$$\arg\sup_{\omega \in \mathcal{Q}_P} \{\text{TP}_{\text{R}}(\omega, \psi)\} = \arg\inf_{\omega \in \mathcal{Q}_P} \{D_{1/2}(\omega, \psi)\}. \tag{5.12}$$

From the fact [105] that

$$\arg\sup_{\omega \in \mathcal{Q}_P} \{\text{TP}_{\text{R}}(\omega, \psi)\} = \arg\sup_{\omega \in \mathcal{Q}_P} \{\text{TP}_{\text{CU}}(\omega, \psi)\} \tag{5.13}$$

whenever $\mathcal{N}$ is commutative, or ψ is pure, or $\mathcal{Q}_P = \{*\}$, or $P \in \{x \in \mathcal{N} \mid \sigma_t^\psi(x) = x \ \forall t \in \mathbb{R}\}$ for $\sigma_t^\psi := \pi_\omega^{-1}(\Delta_{\omega,\omega}^{it} \pi_\omega(x) \Delta_{\omega,\omega}^{-it})$, we can conclude that in any of these cases

$$\mathfrak{P}_{\mathcal{Q}_P}^{D_{1/2}}(\psi) = \frac{P\rho_\psi P}{\text{tr}(P\rho_\psi)}. \tag{5.14}$$

Thus, under the above conditions, the strong Lüders rule can be derived as a result of constrained minimisation of $D_{1/2}(\omega, \psi)$. However, these conditions are so restrictive (eliminating e.g. nonpure density operators) that it is more proper to say that the equation (5.14) expresses a derivation of the strong von Neumann rule. Raggio [105] showed also that, for a general ψ and K the same as in (5.11), (5.12) leads to a different result than (5.11). This is also the case for a general ψ and $K = \mathcal{Q}_{\text{sL}}$. Nevertheless, we can prove the following:

Proposition 5.1. *For $\psi \in \mathfrak{G}_1(\mathcal{H})^+$ and $\mathcal{Q}_{\text{L}}$ given by (4.48),*

$$\mathfrak{P}_{\mathcal{Q}_{\text{L}}}^{D_{1/2}}(\psi) = \sum_{i=1}^n P_i \rho_\psi P_i. \tag{5.15}$$

Proof. Follows directly from (5.7) and

$$(d_{L_2(\mathcal{N})}(\phi, \psi))^2 = 2D_{1/2}(\phi, \psi). \tag{5.16}$$

$\square$

Another result was obtained by Hadjisavvas [59, 60], who showed that the strong von Neumann rule for pure $\psi \in \mathfrak{G}_1(\mathcal{H})_1^+$ can be derived as

$$\mathfrak{P}_{\mathcal{Q}_{\text{sL}}}^{d_{L_1(\mathcal{N})}} = \frac{P\rho_\psi P}{\text{tr}(P\rho_\psi)}, \tag{5.17}$$

where $d_{L_1(\mathcal{N})}$ is the metrical ***Jauch–Misra–Gibson–Kronfli distance*** [69, 81, 60] on $\mathcal{N}_\star \cong L_1(\mathcal{N})$,

$$d_{L_1(\mathcal{N})}(\phi, \psi) := \frac{1}{2}\|\phi - \psi\|_{\mathcal{N}_\star}. \tag{5.18}$$

It is worth noticing that the original definition [69] of $d_{L_1(\mathcal{N})}$ was provided over the measures on orthonormal orthomodular lattice $\text{Proj}(\mathfrak{B}(\mathcal{H}))$, where it takes a form

$$\sup_{x \in \text{Proj}(\mathfrak{B}(\mathcal{H}))} |p_1(x) - p_2(x)|, \tag{5.19}$$

while the distance minimised in [93, 96, 94, 95] can be represented in a form

$$\sup_{x \in \mathcal{A}} |p_1(x) - p_2(x)|, \tag{5.20}$$

where $\mathcal{A}$ is a boolean algebra[10]. In this sense, Hadjisavvas' derivation of the strong von Neumann rule from minimisation of $d_{L_1(\mathcal{N})}$ is similar to Burris' derivation of Jeffrey's rule as a (nonunique) minimiser of (5.20) (reported in [96, 95][11]). In [36] Diu showed that, when applied to nonpure states ψ, $\mathfrak{P}^{d_{L_1(\mathcal{N})}}_{\mathcal{Q}_{\mathrm{sL}}}(\psi)$ does not lead to the strong Lüders rule (in [37] he extended this result to a more general family of metrical distances on quantum states). In our opinion, Diu's result cannot be used as a general argument against using constrained minimisation of $d_{L_1(\mathcal{N})}$, or any other information distance, because strong Lüders' rule is not a uniquely "correct" quantum state change rule. However, on the positive side, Diu's result and the similar result by Raggio on $\mathfrak{P}^{D_{1/2}}_{\mathcal{Q}_{\mathrm{sL}}}(\psi)$ for nonpure ψ exemplify that the choice of an information distance subjected to minimisation preselects the type of possible constraints and the class of admissible results. In this sense, the choice of an information distance (metrical or not) amounts to the choice of a specific convention of inference, which in turn determines some range of possible forms of information dynamics and their output states.

6 Discussion

The information state changes conditioned on certainties (yes/no truth values corresponding to elements of a boolean algebra or orthonormal lattice of projections) are definitely not the only way, and also not the most useful way, of defining information dynamics of information states, both statistical and quantum. The possibility of a derivation of the Bayes–Laplace, Jeffrey's, Lüders', and quantum Jeffrey's rules from minimisation of different distance functionals subjected to various constraints shows the explanatory (semantic) strength of the approach based on entropic projections.

While *in principle* any quantum state change rule (such as Luders' rule, quantum Jeffrey's rule, or some quantum channel[12]) is an *ad hoc* postulate, not derived from any other, more fundamental, property of quantum theoretic formalism, *in practice* the choice of a particular form of this rule reflects certain assumptions about the relationship between the knowledge about outcomes of experimental procedure and the contents of a quantum theoretical model. For example, weak Lüders' rule assumes the specific type of knowledge about the "measurement result", requiring specification of the choice of the projection operators $\{P_i \mid i \in I\} \subseteq \mathfrak{B}(\mathcal{H})$. This restricts the allowed form of 'experimental evidence' to nonempty convex closed linear subspaces of a commutative L_2 space.

A virtue of the approach based on quantum entropic projections is that it allows for a vast generalisation beyond the above restriction, while keeping clear underlying conceptual principles as well as strong mathematical useability. Each information distance functional D expresses a choice of a specific convention on the preferred/relevant and unpreferred/irrelevant aspects of information states: the relevant aspects are those that more strongly participate in the values of D (one can think of a conventional character of a least squares distance, which expresses certain arbitrary preferences regarding the information content of the data). To every choice of an information distance there corresponds a preferred type of constraint (preferred geometric form of 'experimental evidence') for which this information distance is capable of obtaining a unique

---

[10]More precisely, it is a finite boolean algebra that is defined as a Lindenbaum–Tarski algebra of a predicate calculus language.

[11]In [118] this result is incorrectly attributed to [68], despite the clear statement in [96, 95]. Note that this result was proven only for Jeffrey's rule (2.6) with $n = 2$.

[12]That is, a completely positive trace preserving map between quantum states.

24

minimum. The general way to introduce constraints is then to provide a mapping $\Xi \to \mathcal{M}(\mathcal{N})$ from the space Ξ of ('epistemic'/'experimental'/'registration') parameters describing the sets of "possible outcomes" into nonempty convex closed subsets of $\mathcal{M}(\mathcal{N})$.

This way, as opposed to the Bayes–Laplace rule and Lüders' rules, in our framework the conditioning is provided not upon the abstract 'event' that belongs to a boolean algebra or to an orthomodular lattice, respectively, but upon the value taken in the space Ξ of parameters describing the "possible outcomes". This is similar to the semi-spectral (povm/cp-map) approach, which uses linear povm-instruments conditioned upon the spaces of "possible outcomes" that belong to $\mho(\mathcal{X})$. This difference (conditioning upon 'quantitative information' as opposed to conditioning upon 'event') is the key insight. Note that in the semi-spectral approach there is also no bijection between the space of 'abstract events' and the space of "possible outcomes": a single effect can correspond to various elements $\mathcal{Y} \in \mho(\mathcal{X})$ [15]: the 'eigenstate-eigenvalue link' [45] breaks down here. However, as opposed to the semi-spectral approach, our approach completely detaches from the reliance on the use of spectral theory in foundations of quantum theory, allowing for more flexible operational specification of the 'experimental evidence', and for deriving "quantum measurement" rules from a single underlying principle (entropic projection), which is alternative to quantum channels. In particular, the use of quantum distance as an underlying mathematical structure allows for geometric analysis and justification of the choice of a specific convention of quantum inference/dynamics. This is in contrast with the general lack of clear geometric justification for a choice of a specific quantum channel in the semi-spectral approach.

An extended development of the approach based on quantum entropic projections as an alternative to povm-instruments, and as a replacement for Lüders' rules, will be carried out in [80]. For an alternative derivation of Lüders' rules from quantum entropic projections based on D_0, see [62].

What does it all mean for quantum bayesianism? Quite often 'bayesianism' is understood as a subjective interpretation of probability equipped with the requirement of using the Bayes–Laplace rule for the purpose of changing probabilities due to learning new information. However, the appearance of other updating rules, such as Jeffrey's rule, Field's rule [44], and constrained maximisation of the WGKL distance [82, 65, 74, 123], has undermined the universality of a second component. The usual perspective on the meaning of 'quantum bayesianism' starts from the semi-spectral approach to mathematical foundations of quantum theory, and aims at recasting (some suitable class of) povm-instruments as (a modified form of) the Bayes–Laplace rule, while keeping the subjective interpretation of probabilities. Such perspective assumes that probability theory and spectral theory should be fundamental constitutents of quantum theory. Moreover, it also does not provide justification for using povm-instruments: the mathematical foundations are just taken for granted. The results contained in this paper are intended to serve as a guideline (via quantum Jeffrey's rule) and as a backwards compatibility proof (via recovery of Lüders' rules) for an alternative approach to the mathematical and conceptual meaning of 'quantum bayesianism'. According to our point of view, quantum states should be used as carriers of intersubjective knowledge on their own mathematical right (as elements of a noncommutative L_1 space), without reference to probabilities (elements of a commutative L_1 space), while the processes of inductive inference (information dynamics) can be fruitfully modelled using quantum entropic projections as a nonlinear alternative to the Bayes–Laplace rule and linear povm-instruments. See [80] for a detailed account.

Acknowledgments

References

[1] Alberti P.M., Uhlmann A., 1984, *Transition probabilities on C^*- and W^*-algebras*, in: Baumgärtel H., Laßner G., Pietsch A., Uhlmann A. (eds.), *Proceedings of the second international conference on operator algebras, ideals, and their applications in theoretical physics (Leipzig 1983)*, Teubner, Stuttgart, p.5. www.physik.uni-leipzig.de/~uhlmann/PDF/Uh84a.pdf. ↑ 21.

[2] 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. ↑ 5.

[3] Araki H., 1972, *Bures distance function and a generalization of Sakai's noncommutative Radon–Nikodym theorem*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **8**, 335. dx.doi.org/10.2977/prims/1195193113. ↑ 22.

[4] Araki H., 1973, *Relative hamiltonian for faithful normal states of von Neumann algebra*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **9**, 165. dx.doi.org/10.2977/prims/1195192744. ↑ 8.

[5] Araki H., 1974, *Some properties of the modular conjugation operator of von Neumann algebras and a noncommutative Radon–Nikodym theorem with a chain rule*, Pacific J. Math. **50**, 309. euclid:pjm/1102913224. ↑ 22.

[6] Araki H., 1976, *Introduction to relative hamiltonian and relative entropy*, in: Guerra F., Robinson D.W., Stora R. (eds.), *Les methodes mathematiques de la theorie quantique des champs, Marseille, 23-27 juin 1975*, Colloques Internationaux C.N.R.S. **248**, Éditions du C.N.R.S., Paris, p.782. ↑ 9.

[7] Araki H., 1976, *Positive cone, Radon–Nikodym theorems, relative hamiltonian and the Gibbs condition in statistical mechanics*, in: Kastler D. (ed.), *C^*-algebras and their applications to statistical mechanics and quantum field theory*, North-Holland, Amsterdam. ↑ 9.

[8] Araki H., 1976, *Relative entropy for states of von Neumann algebras I*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **11**, 809. dx.doi.org/10.2977/prims/1195191148. ↑ 9, 19.

[9] Araki H., 1977, *Relative entropy for states of von Neumann algebras II*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **13**, 173. dx.doi.org/10.2977/prims/1195190105. ↑ 8, 9, 18.

[10] Araki H., Masuda T., 1982, *Positive cones and L_p-spaces for von Neumann algebras*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **18**, 339. dx.doi.org/10.2977/prims/1195183577. ↑ 8, 18.

[11] 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). ↑ 1, 3.

[12] Benoist R.W., Marchand J.-P., 1979, *Statistical inference in coupled quantum systems*, Lett. Math. Phys. **3**, 93. ↑ 2, 22.

[13] Benoist R.W., Marchand J.-P., Wyss W., 1979, *A note on relative entropy*, Lett. Math. Phys. **3**, 169. ↑ 2.

[14] Benoist R.W., Marchand J.-P., Yourgrau W., 1977, *Statistical inference and quantum mechanical measurement*, Found. Phys. **7**, 827 (addendum: 1978, Found. Phys. **8**, 117). ↑ 2, 22.

[15] Braunstein S.L., Caves C.M., 1988, *Quantum rules: an effect can have more than one operation*, Found. Phys. Lett. **1**, 3. ↑ 25.

[16] Bub J., 1977, *Von Neumann's projection postulate as a probability conditionalization rule in quantum mechanics*, J. Phil. Logic **6**, 381. ↑ 1, 2.

[17] Bub J., 1979, *Conditional probabilities in non-boolean possibility structures*, in: Hooker C.A. (ed.), *The logico-algebraic approach to quantum mechanics*, Vol.2, Reidel, Dordrecht, p.209. ↑ 1, 2.

[18] Bub J., 1979, *The measurement problem in quantum mechanics*, in: Toraldo di Francia G. (ed.), *Problems in the foundations of physics*, North-Holland, Amsterdam, p.71. ↑ 1, 2.

[19] Bub J., 2007, *Quantum probabilities as degrees of belief*, Stud. Hist. Phil. Mod. Phys. **38**, 232. ↑ 1.

[20] Bures D.J.C., 1969, *An extension of Kakutani's theorem on infinite product measures to the tensor product of semiinfinite W^* algebras*, Trans. Am. Math. Soc. **135**, 194. ↑ 2, 22.

[21] Cantoni V., 1975, *Generalized "transition probability"*, Comm. Math. Phys. **44**, 125. euclid:cmp/1103899296. ↑ 21.

[22] Carazza B., Casartelli M., D'Elia A., 1977, *Segal entropy and the principle of least interference*, Phys. Lett. A **62**, 205. ↑ 2.

[23] 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. ↑ 1, 4.

[24] Caves C.M., Fuchs C.A., Schack R., 2001, *Quantum probabilities as bayesian probabilities*, Phys. Rev. A **65**, 022305. arXiv:quant-ph/0106133. ↑ 1.

[25] Coles P.J., 2012, *Unification of different views of decoherence and discord*, Phys. Rev. A **85**, 042103. arXiv:1110.1664. ↑ 3, 26.

[26] 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. ↑ 8.

[27] Connes A., 1974, *Caractérisation des espaces vectoriels ordonnés sous-jacents aux algèbres de von Neumann*, Ann. Inst. Fourier Grenoble **24**, 121. numdam:AIF_1974__24_4_121_0. ↑ 8.

[28] Csiszár 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. ↑ 5.

[29] de Laplace P.-S., 1774, *Mémoire sur la probabilité des causes par les événements*, Mém. Acad. Roy. Sci. Paris (Savants Étrangers) **6**, 621. (reprinted in: 1986, Stat. Sci. **1**, 359.). ↑ 1, 3.

[30] de Laplace P.-S., 1812, *Théorie analytique des probabilités*, Courcier, Paris. ↑ 3.

[31] de Laplace P.-S., 1814, *Essai philosophique sur les probabilités*, Courcier, Paris. (Engl. transl. 1951, *A philosophical essay on probabilities*, Dover, New York). ↑ 1, 3.

[32] de Muynck W.M., 2002, *Foundations of quantum mechanics: an empiricist approach*, Kluwer, Dordrecht. ↑ 11.

[33] Diaconis P., Zabell S., 1982, *Updating subjective probability*, J. Amer. Statist. Assoc. **77**, 822. ↑ 1, 5.

[34] Dieks D., Veltkamp P., 1983, *Distance between quantum states, statistical inference and the projection postulate*, Phys. Lett. A **97**, 24. ↑ 2, 22.

[35] Digernes T., 1975, *Duality for weights on covariant systems and its applications*, Ph.D. thesis, University of California, Los Angeles. ↑ 8.

[36] Diu B., 1982, *Note on a recent proposal concerning statistical inference in quantum theory*, Ann. Inst. Henri Poincaré A **37**, 59. ↑ 24.

[37] Diu B., 1983, *Statistical inference and distance between states in quantum mechanics*, Ann. Inst. Henri Poincaré A **38**, 167. ↑ 24.

[38] Domotor Z., 1985, *Probability kinematics, conditionals, and entropy principles*, Synthese **63**, 75. ↑ 1, 23.

[39] Domotor Z., Zanotti H., Graves H., 1980, *Probability kinematics*, Synthese **44**, 421. ↑ 1.

[40] Donald M.J., 1986, *On the relative entropy*, Commun. Math. Phys. **105**, 13. euclid:cmp/1104115254. ↑ 2, 20.

[41] Donald M.J., 1987, *Further results on the relative entropy*, Math. Proc. Cambridge Phil. Soc. **101**, 363. ↑ 2, 20.

[42] Donald M.J., 1990, *Relative hamiltonians which are not bounded from above*, J. Funct. Anal. **91**, 143. ↑ 18.

[43] Douven I., Romeijn J.-W., 2012, *A new resolution of the Judy Benjamin problem*, Mind **479**, 637. eprints.lse.ac.uk/27004/1/A_new_resoltuion_(LSERO).pdf. ↑ 1, 5.

[44] Field H., 1978, *A note on Jeffrey conditionalization*, Phil. Sci. **45**, 361. ↑ 25.

[45] Fine T.L., 1973, *Theories of probability*, Academic Press, New York. ↑ 25.

[46] Fok V.A., 1932, *Nachala kvantovoi mekhaniki*, Kubuch, Leningrad (Engl. transl. 1978, Fundamentals of quantum mechanics, Mir, Moskva). ↑ 11.

[47] Fréchet M., 1906, *Sur quelques points du calcul fonctionnel*, Rend. Circ. Matem. Palermo **22**, 1. webpages.ursinus.edu/nscoville/Frechet Thesis French I.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French II.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French III.pdf, webpages.ursinus.edu/nscoville/Frechet Thesis French IV.pdf. ↑ 9.

[48] Fuchs C.A., 2002, *Quantum mechanics as quantum information (and only a little more)*, arXiv:quant-ph/0205039. ↑ 1.

[49] Fuchs C.A., 2003, *Quantum mechanics as quantum information, mostly*, J. Mod. Opt. **50**, 987. arXiv:quant-ph/0205039. ↑ 1.

[50] Gel'fand I.M., Naĭmark M.A., 1943, *On the imbedding of normed rings into the ring of operators in Hilbert space*, Matem. Sbornik (N.S.) **12**, 197. mathnet.ru:msb6155. ↑ 6.

[51] Giffin A., 2008, *Maximum entropy: the universal method for inference*, Ph.D. thesis, State University of New York, Albany. arXiv:0901.2987. ↑ 1, 4.

[52] Giffin A., Caticha A., 2007, *Updating probabilities with data and moments*, in: Knuth K. et al. (eds.), *Bayesian inference and maximum entropy methods in science and engineering*, AIP Conf. Proc. **954**, 74. arXiv:0708.1593. ↑ 5.

[53] Good I.J., 1950, *Probability and the weighing of evidence*, Griffin, London. ↑ 10.

[54] Gudder S.P., 1980, *Statistical inference in quantum mechanics*, Rep. Math. Phys. **17**, 265. ↑ 2, 22.

[55] Gudder S.P., Marchand J.-P., 1977, *Conditional expectations on von Neumann algebras: a new approach*, Rep. Math. Phys. **12**, 317. ↑ 2.

[56] Gudder S.P., Marchand J.-P., Wyss W., 1979, *Bures distance and relative entropy*, J. Math. Phys. **20**, 1963. ↑ 2, 22.

[57] Haagerup U., 1973, *The standard form of von Neumann algebras*, Preprint Ser. 1973 No. **15**, Københavns Universitet Matematisk Institut, København. ↑ 7.

[58] Haagerup U., 1975, *The standard form of von Neumann algebras*, Math. Scand. **37**, 271. www.mscand.dk/article.php?id=2275. ↑ 7.

[59] Hadjisavvas N., 1978, *Étude de certaines consequences d'une interprétation subjective de la notion d'état*, Ann. Fond. Louis de Broglie **3**, 155. ↑ 2, 23.

[60] Hadjisavvas N., 1981, *Distance between states and statistical inference in quantum theory*, Ann. Inst. Henri Poincaré A **38**, 167. numdam:AIHPA_1981__35_4_287_0.pdf. ↑ 2, 23.

[61] Hasegawa H., 1993, *α-divergence of the non-commutative information geometry*, Rep. Math. Phys. **33**, 87. ↑ 9.

[62] Hellmann F., Kamiński W., Kostecki R.P., 2014, *Quantum collapse rules from the maximum relative entropy principle*, arXiv:1407.7766. ↑ 2, 25.

[63] Henderson L., 2010, *Bayesian updating and information gain in quantum measurement*, in: Bokulich A., Jaeger G. (eds.), *Philosophy of quantum information and entanglement*, Cambridge University Press, Cambridge, p.151. ↑ 1.

[64] Herbut F., 1969, *Derivation of the change of state in measurement from the concept of minimal measurement*, Ann. Phys. **55**, 271. ↑ 2, 11, 22.

[65] Hobson A., Cheng B.-K., 1973, *A comparison of the Shannon and Kullback information measures*, J. Stat. Phys. **7**, 301. ↑ 25.

[66] Hughes R.I.G., van Fraassen B.C., 1984, *Symmetry arguments in probability kinematics*, in: Kitcher P., Asquith P. (ed.), *Proceedings of the biennal meeting of the Philosophy of Science Association*, Vol.2, Philosophy of Science Association, East Lausing, p.851. ↑ 1.

[67] Jacobs K., 2002, *How do two observers pool their knowledge about a quantum system?*, Quant. Inf. Proc. **1**, 73. arXiv:quant-ph/0201096. ↑ 1.

[68] Jamison B., 1974, *A Martin boundary interpretation of a minimum change principle to probability kinematics*, Z. Warschein. Geb. **30**, 265. ↑ 24.

[69] Jauch J.M., Misra B., Gibson A.G., 1968, *On the asymptotic condition of scattering theory*, Helv. Phys. Acta **41**, 513. ↑ 23.

[70] Jeffrey R.C., 1957, *Contributions to the theory of inductive probability*, Ph.D. thesis, Princeton University, Princeton. ↑ 3.

[71] Jeffrey R.C., 1965, *The logic of decision*, Chicago University Press, Chicago. ↑ 3.

[72] Jeffrey R.C., 1968, *Probable knowledge*, in: Lakatos I. (ed.), *The problem of inductive logic*, North-Holland, Amsterdam, p.166. ↑ 3.

[73] Jenčová A., 2005, *Quantum information geometry and non-commutative L_p spaces*, Inf. Dim. Anal. Quant. Prob. Relat. Top. **8**, 215. www.mat.savba.sk/~jencova/lpspaces.pdf. ↑ 9.

[74] Johnson R.W., 1979, *Axiomatic characterization of the directed divergences and their linear combinations*, IEEE Trans. Inf. Theory **25**, 709. ↑ 25.

[75] Kemble E.C., 1937, *The fundamental principles of quantum mechanics*, McGraw–Hill, New York. ↑ 11.

[76] Kosaki H., 1980, *Canonical L^p-spaces associated with an arbitrary abstract von Neumann algebra*, Ph.D. thesis, University of California, Los Angeles. dmitripavlov.org/scans/kosaki-thesis.pdf. ↑ 8, 22.

[77] Kostecki R.P., 2011, *The general form of γ-family of quantum relative entropies*, Open Sys. Inf. Dyn. **18**, 191. arXiv:1106.2225. ↑ 9.

[78] Kostecki R.P., 2011, *Quantum information geometric foundations of quantum theory and space-time*, talk given at Quantum foundations seminar (July 12), Perimeter Institute, Waterloo. ↑ 1.

[79] Kostecki R.P., 2013, *W^*-algebras and noncommutative integration*, arXiv:1307.4818. ↑ 1, 5, 22.

[80] Kostecki R.P., 2014, *Towards quantum information geometric foundations*, in preparation. ↑ 25.

[81] Kronfli N.S., 1970, *States on generalised logics*, Int. J. Theor. Phys. **3**, 191. ↑ 23.

[82] Kullback S., 1959, *Information theory and statistics*, Wiley, New York (2nd ed. 1968). libgen.org:66875BFC74B54B9B73E51375EF11FEDB. ↑ 25.

[83] Kullback S., Leibler R.A., 1951, *On information and sufficiency*, Ann. Math. Statist. **22**, 79. euclid:aoms/1177729694. ↑ 10.

[84] Lüders G., 1951, *Über die Zustandsänderung durch den Messprozess*, Ann. Phys. Leipzig **8**, 322. (Engl. transl.: 2004, *Concerning the state-change due to the measurement process*, arXiv:quant-ph/0403007). ↑ 1, 10.

[85] Marchand J.-P., 1977, *Relative coarse-graining*, Found. Phys. **7**, 35. ↑ 2, 22.

[86] Marchand J.-P., 1981, *Statistical inference in quantum mechanics*, in: Gustafson K.E. et al (eds.), *Quantum mechanics in mathematics, chemistry and physics*, Plenum, New York, p.73. ↑ 2.

[87] Marchand J.-P., 1983, *Statistical inference by minimal Bures distance*, in: van der Merwe (ed.), *Old and new questions in physics, cosmology, philosophy, and theoretical biology. Essays in honor of Wolfgang Yourgreau*, Plenum, New York, p.275. libgen.org:e97f930eac263d74a4c065bbc15287ae. ↑ 2, 22.

[88] Marchand J.-P., 1983, *Statistical inference in non-commutative probability*, Rend. Sem. Math. Fis. Milano **52**, 551. ↑ 2.

[89] Marchand J.-P., Wyss W., 1977, *Statistical inference and entropy*, J. Stat. Phys. **16**, 349. ↑ 2, 22.

[90] Margenau H., 1936, *Quantum mechanical descriptions*, Phys. Rev. **49**, 240. ↑ 11.

[91] Margenau H., 1963, *Measurements and quantum states: part I, II*, Phil. Sci. **30**, 1, 138. ↑ 11.

[92] Masuda T., 1984, *A note on a theorem of A. Connes on Radon-Nikodym cocycles*, Publ. Res. Inst. Math. Sci. Kyōto Univ. **20**, 131. doi:10.2977/prims/1195181833. ↑ 8.

[93] May S.J., 1973, *On the application of a minimum change principle to probability kinematics*, Ph.D. thesis, University of Waterloo, Waterloo. ↑ 24.

[94] May S.J., 1976, *Probability kinematics: a constrained optimization problem*, J. Phil. Log. **5**, 395. ↑ 1, 24.

[95] May S.J., 1979, *An application of Neustadt's abstract maximum principle to probability kinematics*, J. Optim. Th. Appl. **27**, 249. ↑ 1, 24.

[96] May S.J., Harper W.L., 1976, *Towards an optimisation procedure for applying minimum change principles in probability kinematics*, in: Harper W.L., Hooker C.A. (eds.), *Foundations of probability theory, statistical inference, and statistical theories of science*, Vol.1, Reidel, Dordrecht, p.137. ↑ 1, 24.

[97] Modi K., Paterek T., Son W., Vedral V., Williamson M., 2010, *Unified view of quantum and classical correlations*, Phys. Rev. Lett. **104**, 080501. arXiv:0911.5417. ↑ 2, 26.

[98] Morimoto T., 1963, *Markov processes and the H-theorem*, J. Phys. Soc. Jap. **12**, 328. ↑ 5.

[99] Murray F.J., von Neumann J., 1936, *On rings of operators*, Ann. Math. **37**, 116. ↑ 7.

[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. ↑ 9.

[101] Palge V., Konrad T., 2008, *A remark on Fuchs' bayesian interpretation of quantum mechanics*, Stud. Hist. Phil. Mod. Phys. **39**, 273. ↑ 1.

[102] Petz D., 1985, *Properties of quantum entropy*, in: Accardi L., von Waldenfels W. (eds.), *Quantum probability and applications II*, LNM **1136**, Springer, Berlin, p.428. ↑ 9.

[103] Petz D., 1986, *Properties of the relative entropy of states of von Neumann algebra*, Acta Math. Hungar. **47**, 65. ↑ 9, 16.

[104] Raggio G.A., 1982, *Comparison of Uhlmann's transition probability with the one induced by the natural positive cone of von Neumann algebras in standard form*, Lett. Math. Phys. **6**, 233. ↑ 21.

[105] Raggio G.A., 1984, *Generalized transition probabilities and applications*, in: Accardi L. et al (eds.), *Quantum probability and appplications to quantum theory of irreversible processes*, Springer, Berlin, p.327. libgen.org:c4a12cdc49b77c7d0d999f6bfb49c1c1. ↑ 2, 21, 22, 23.

[106] Rédei M., 1992, *When can non-commutative statistical inference be bayesian?*, Int. Stud. Phil. Sci. **6**, 129. ↑ 1.

[107] Rédei M., Summers S.J., 2007, *Quantum probability theory*, Stud. Hist. Phil. Mod. Phys. **38**, 390. arXiv:quant-ph/0601158. ↑ 1.

[108] Sakai S., 1965, *A Radon–Nikodym theorem in W^*-algebras*, Bull. Amer. Math. Soc. **71**, 149. euclid:bams/1183526404. ↑ 22.

[109] Schack R., Brun T.A., Caves C.M., 2001, *Quantum Bayes rule*, Phys. Rev. A **64**, 014305. arXiv:quant-ph/0008113. ↑ 1.

[110] Schwinger J., 1959, *The algebra of microscopic measurement*, Proc. Nat. Acad. Sci. U.S.A. **45**, 1542. ↑ 10.

[111] Segal I.E., 1947, *Irreducible representations of operator algebras*, Bull. Amer. Math. Soc. **61**, 69.

www.ams.org/journals/bull/1947-53-02/S0002-9904-1947-08742-5/S0002-9904-1947-08742-5.pdf. ↑ 6.

[112] Streater R.F., 2007, *Lost causes in and beyond physics*, Springer, Berlin. ↑ 1, 2.

[113] Tribus M., Rossi R., 1973, *On the Kullback information measure as a basis for information theory: comments on a proposal by Hobson and Chang*, J. Stat. Phys. **9**, 331. ↑ 5.

[114] Uhlmann A., 1976, *The "transition probability" in the state space of a *-algebra*, Rep. Math. Phys. **9**, 273. www.physik.uni-leipzig.de/~uhlmann/PDF/Uh76a.pdf. ↑ 21.

[115] Umegaki H., 1961, *On information in operator algebras*, Proc. Jap. Acad. **37**, 459. euclid:pja/1195523632. ↑ 9.

[116] Umegaki H., 1962, *Conditional expectation in an operator algebra, IV (entropy and information)*, Kōdai Math. Sem. Rep. **14**, 59. euclid:kmj/1138844604. ↑ 9.

[117] Valente G., 2007, *Is there a stability problem for bayesian noncommutative probabilities?*, Stud. Hist. Phil. Mod. Phys. **38**, 832. ↑ 1.

[118] van Fraassen B.C., 1981, *A problem for relative information minimizers in probability kinematics*, Brit. J. Phil. Sci. **32**, 375. ↑ 4, 24.

[119] von Neumann J., 1930, *Zur algebra der Funktionaloperatoren und Theorie der normalen Operatoren*, Math. Ann. **102**, 370. gdz.sub.uni-goettingen.de/dms/load/img/?PPN=GDZPPN002273675&IDDOC=38466. ↑ 7.

[120] von Neumann J., 1932, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin. libgen.org:DF8F17426E6D36B4AD2D350970158BD7 (Engl. transl. 1955, *Mathematical foundations of quantum mechanics*, Princeton University Press, Princeton. libgen.org:C23114E46AB9284F70E789F93BC0512D). ↑ 1, 10.

[121] Wald A., 1947, *Sequential analysis*, Wiley, New York. ↑ 10.

[122] 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** *(NIPS 05)*, MIT Press, p.1457. users.soe.ucsc.edu/~manfred/pubs/C72.pdf. ↑ 1, 2.

[123] Williams P.M., 1980, *Bayesian conditionalisation and the principle of minimum information*, Brit. J. Phil. Sci. **31**, 131. ↑ 1, 4, 25.

[124] Yamagami S., 2008, *Geometric mean of states and transition amplitudes*, Lett. Math. Phys. **84**, 123. arXiv:0801.0858. ↑ 21.

[125] Yamagami S., 2010, *Geometry of quasi-free states of CCR algebras*, Int. J. Math. **21**, 875. arXiv:0801.1739. ↑ 21.

[126] Zellner A., 1988, *Optimal information processing and Bayes' theorem*, Amer. Stat. **42**, 278. ↑ 1.