Quantum collapse rules from the maximum relative entropy principle

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2016 New J. Phys. 18 013022
(http://iopscience.iop.org/1367-2630/18/1/013022)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 193.174.18.5
This content was downloaded on 04/02/2016 at 12:15

Please note that terms and conditions apply.
Quantum collapse rules from the maximum relative entropy principle

Frank Hellmann, Wojciech Kamiński and Ryszard Paweł Kostecki

Abstract

We show that the von Neumann–Lüders collapse rules in quantum mechanics always select the unique state that maximises the quantum relative entropy with respect to the premeasurement state, subject to the constraint that the postmeasurement state has to be compatible with the knowledge gained in the measurement. This way we provide an information theoretic characterisation of quantum collapse rules by means of the maximum relative entropy principle.

1. Introduction

The dynamics of quantum states in the orthodox (von Neumann’s) foundations of quantum mechanics consist of two different prescriptions: the unitary evolution and the so-called ‘collapse’ of a quantum state to a subspace encoding the knowledge gained in the outcome of a measurement. The mappings (rules) describing this collapse were originally formulated by von Neumann [1] and later improved by Lüders [2]. There are two different forms of collapse. When one knows only a measurement corresponding to an observable (a self-adjoint operator with a discrete spectrum) O has taken place, the ‘weak’ rule applies. It is defined as $\rho \mapsto \sum_{i \in I} P_i \rho P_i$, where $\rho$ is the original quantum state (in general, a density operator), while $O = \sum_{i \in I} \lambda_i P_i$ is a spectral decomposition with some countable index set $I$ (hence, $\sum_{i \in I} P_i = 1$, $P_i P_j = P_j P_i$ and $\lambda_i \in \mathbb{R}$ $\forall$ $i, j \in I$). If a measurement corresponding to $O$ has resulted in a specific value $\lambda_i \in \{\lambda_i | i \in I\}$ associated to a projector $P_i \in \{P_i | i \in I\}$, then the ‘strong’ rule, $\rho \mapsto P_i \rho P_i/\text{tr}(P_i)\rho$, is applied.

The negative of Umegaki’s quantum relative entropy [3, 4], $D(\rho, \sigma) = -S(\rho, \sigma) := \text{tr}(\rho \ln \rho - \rho \ln \sigma) \in [0, \infty]$, can be used as a measure of distinguishability, or relative information content, of the quantum state $\sigma$ from the state $\rho$. The use of $D$ instead of $S$ follows Wiener’s idea that the ‘amount of information is the negative of the quantity defined as entropy’ [5]. Note that we call $S = -D$ the relative entropy, following the convention [6] that makes the Gibbs–Shannon and von Neumann entropies the special cases of $S$, after adding a constant: $S_N(\rho) = S(\rho, \mathbb{I}/n) + \log(n)$.

The function $D$ can be considered as a nonsymmetric distance: in general, $D(\rho, \sigma) = D(\sigma, \rho)$. If a given state is $\sigma$ and we believe it to be $\rho$, it can be easier or harder to find our error than if their roles were reversed. Say, $\sigma = P$ with $P$ some projector and $\rho = \mathbb{I}/n$. If we measure the property corresponding to $\mathbb{I} - P$, a single measurement can tell us that the state is not $\sigma$, whereas no single measurement could reveal the same of $\rho$. See e.g. [7, 8] for an overview of reasons for using $D(\rho, \sigma)$ as a measure of distinguishability and relative information content.

A key information theoretic property of the strong collapse rule is that the probability of measuring the value $\lambda_i$ again, after having measured it once, is $1$, which follows from $\text{tr} \left( P_k \frac{P_k \rho P_k}{\text{tr}(P_k)} \right) = 1$. Repeated measurements add no new information. Clearly, the state $P_k \rho P_k/\text{tr}(P_k)$ is not the only state that has this property (note that $P_k$ is not necessary a rank 1 projector). What we demonstrate in this letter is that, among all states that have this property, the strong collapse rule selects the state that is least distinguishable from the initial state $\rho$, that is, it has the minimum relative information $D(\rho, \cdot)$, in a suitably regularised sense. This allows for an information theoretic characterisation of the strong collapse rule: the state after measurement is the state that is least
distinguishable from the previous state, while being compatible with the new information gained by the measurement.

In order to derive the strong collapse rule, we will need two intermediate results. First we will show that the weak collapse rule produces the least distinguishable state among the block diagonal states. We then show that a weighted version of the strong collapse rule, \( \rho \mapsto \sum P_i \rho P_i / \text{tr} (\rho P_i) \), is the least distinguishable amongst the states with blocks of fixed trace. This rule can be interpreted as corresponding to a measurement where we believe that the result \( P_i \) occurred with probability \( p_i \). This intermediate step regularises the problem of a strong collapse, which is then obtained as a limiting case, by taking \( p_i \leftrightarrow \delta_k \) with \( k = 1 \).

Our derivation of the collapse rules from the constrained maximisation of Umegaki’s quantum relative entropy is of special importance in the context of epistemic and information theoretic approaches to the foundations of quantum theory. In this context, collapse rules have been considered as analogues of the Bayes–Laplace rule [9–12]. This analogy rested on mathematical and conceptual similarity, but was not derived from any single unifying principle. In the meantime, the Bayes–Laplace rule has been shown to be a special case of the constrained maximisation of the Kullback–Leibler relative entropy [13–16]. Our result provides the missing piece of the puzzle. Both the Bayes–Laplace and von Neumann–Lüders rules are special cases of a single epistemic principle of inductive inference (or, in other words, information theoretic state updating). This issue will be discussed in more detail in section 5.

After finishing this paper, we were informed about reference [17], where it is shown that a state \( \sigma = \sum P_i \rho P_i \), where \( P_i \) are rank 1 projectors, minimises the functional \( D(\rho, \sigma) \). This is a special case of our result for the weak collapse rule. The generalisation to arbitrary projectors is suggested in [18], but without a proof or an indication of a method of proving this statement. The technique used by us to prove a general theorem is essentially different from one applied in [17] (and it shows that this result for the rank \( \geq 1 \) case is more substantial and nontrivial than for the rank 1 case).

A closely related paper [26] deals with the same type of problem as addressed here, but using a different mathematical approach, allowing for treatment of the infinite dimensional case. Further conceptual and mathematical discussion associated with the results of both papers is carried out there and in [28]. A recent work [19] proves that a partial trace is also a constrained maximiser of quantum relative entropy.

2. The setup

We will consider the finite dimensional case. Hence, quantum states will be identified with non-negative matrices of trace 1, which form the convex set \( \mathcal{D} \) in the space of all hermitian \( n \times n \) complex matrices.

The function \( D(\cdot, \cdot) \) is jointly convex in both arguments [20], which implies that \( D(\rho, \cdot) \) is convex on \( \mathcal{D} \) for all \( \rho \in \mathcal{D} \). Due to the finite dimensionality of the problem, we can use the first order condition for the existence of a minimum of a convex function (see e.g. [21], Theorems 1.2.7 and 2.2.1): if \( V \) is a convex subset of a finite dimensional topological vector space, and \( f : V \rightarrow \mathbb{R} \) is convex then \( x \) is a global minimum of \( f \) on \( V \) if and only if all directional derivatives of \( f \) at \( x \) are non-negative.

For a function differentiable at \( x \) this condition states that if \( x \) is in the interior of \( V \) then the derivatives of \( f \) need to vanish. If \( x \) belongs to some strata of the boundary of \( V \) then all tangential derivatives need to vanish whereas derivatives in inward transversal direction need to be non-negative.

In our minimisation problem we have a subspace \( V \subset \mathcal{D} \) of density matrices that is defined by a linear equation and thus is a subsimplex. The function \( D(\rho, \cdot) \) restricts to a convex and differentiable function on \( V \) and we want to find its minimum. Thus we simply differentiate in the directions preserving \( V \) and set the derivatives to be positive. We will denote this condition by

\[
C^V(\cdot) \Rightarrow \partial_V D(\rho, \cdot) = \partial_V \text{tr} (\rho \ln (\cdot)) \geq 0.
\]

The next two sections will be concerned with evaluating this set of equations.

3. Weak collapse

In the case of a weak collapse due to the measurement of \( O = \sum \lambda_i P_i \), the constraint set is given by the block diagonal density matrices

\[
\mathcal{V}_w := \{ \sigma \in \mathcal{D} \mid [P_i, \sigma] = 0 \quad \forall P_i \}. \tag{2}
\]

The condition (2) is equivalent with \( \sigma \in \mathcal{V}_w \iff \sigma = \sum P_i \sigma P_i \), as well as with \( \sigma \in \mathcal{V}_w \iff [O, \sigma] = 0 \) (see [22] for a discussion).
We can parametrise $V_n$ in terms of the singular value decomposition of $\sigma$. Every element of $V_n$ is of the form
\[ \sigma = U A U^*, \]
with $A$ a trace 1 diagonal matrix with positive entries, $U$ a unitary that is a product $U = \prod_{i} U_i$, where $U_i$ is an identity on the range of $I - P_i$. We have that $[U, P_i] = 0$, $[A, P_i] = 0$, and thus, writing $\sigma = \sigma_{\text{I}} + \sigma_{\text{II}}$, we have $f(\sigma) = \sigma_{\text{I}} f(\sigma) = \sigma_{\text{II}} f(\sigma)$; that is, functions (in the sense of the functional calculus) act blockwise on the space $V_n$.

Let us consider first the variation $\partial_L \sigma = \text{tr}(\rho \ln(\cdot)) = 0$ in the direction parametrised by the $U_i$. Given a function on a Lie group $f(U)$ we can take the directional derivative by looking at the parameter derivative of a one parameter group of diffeomorphisms on $U$. As multiplication in a Lie group is differentiable we can pick the one parameter group of diffeomorphisms generated by left multiplication with the one-dimensional subgroup $\exp(tL)$,
\[ \phi_t(U) = \exp(tL)(U). \]

We then define the directional derivative in direction $L$ as the derivative of the pushforward of $f$ along $\phi_t$,
\[ \partial_L f(\cdot) = \frac{d}{dt} \phi_t^* f(\cdot) \big|_{t=0}. \]

For a function that is the trace of $U$ in a particular representation this can be easily evaluated:
\[ \frac{d}{dt} \phi_t^* \text{tr}(A U) \big|_{t=0} = \frac{d}{dt} \text{tr}(A e^{tL}) \big|_{t=0} = \sum_{ij} (A)_{ij} \frac{d}{dt} (e^{tL})_{ij} \big|_{t=0} = \sum_{ij} (A)_{ij} (LU)_{ij} = \text{tr}(ALU). \]

A straightforward calculation shows that we further have $\frac{d}{dt} \phi_t^* \text{tr}(AU^*) \big|_{t=0} = \text{tr}(ALU^*) - \text{tr}(AUBU^*)L$.

Note that $[L_i, P_j] = 0$, and in particular $L_i P_j = \delta_{ij} L_i$. The derivative then takes the form
\[ \partial_L \sum_{i} \text{tr}(P_i \rho P_j U_i \ln \Lambda_i T_j) = \text{tr}(\rho_i L_i \ln \sigma_i) - \text{tr}(\rho_i \ln(\sigma_i) L_i) = 0, \]
\[ \text{tr}(L_i \ln \sigma_i, \rho_i) = 0. \]

We thus see that if $\sigma$ and $\sum_i P_i \rho P_i$ are concurrently diagonalisable, the above equation vanishes. In fact, since $[\ln(\sigma_i), \rho_i]$ is traceless and $[L_i, \rho_i]$ spans the space of all traceless matrices in the $i$-th matrix block, this is also a necessary condition.

Let us next consider the variation in the direction of the spectrum, that is the direction of $\Lambda$. We are interested in the case where $\sigma$ and $\sum_i P_i \rho P_i$ are concurrently diagonalisable. Let $\kappa_i^\sigma$ and $\kappa_i^\rho$ be the eigenvalues of $\sigma$ and $\sum_i P_i \rho P_i$ respectively. If $\kappa_i^\sigma \neq 0$ and $\kappa_i^\rho = 0$ then $D(\rho, \sigma) = \infty$, so this can not be the minimum if a state with finite relative entropy exists, and we can disregard this case here.

Let us first consider the case that all $\kappa_i^\sigma \neq 0$. We have the condition
\[ \partial_{\Lambda_i} D(\rho, \sigma) = \partial_{\Lambda_i} \left( -\text{tr}(\Lambda_i \ln \Lambda_i) \right) = \partial_{\Lambda_i} \sum_k - \kappa_i^\sigma \ln \kappa_i^\sigma = 0. \]
The derivatives $\partial_{\Lambda_i}$ have to preserve the trace. An overcomplete basis of such derivatives is given by $\partial_{\kappa_i^\sigma} - \partial_{\kappa_i^\rho}$. Thus, for all $k, l, m$
\[ \left( \partial_{\kappa_i^\sigma} - \partial_{\kappa_i^\rho} \right) \sum_{m} \kappa_i^\sigma \ln \kappa_i^m = 0, \]
\[ \frac{\kappa_i^\sigma}{\kappa_i^\rho} \sum_{m} \ln \kappa_i^m = 0. \]

So, the ratios of the eigenvalues of $\sum_i P_i \rho P_i$ and $\sigma$ are fixed. As they both are trace 1, this implies they are the same.

Let assume now that $I$ is the index set of all $i$ such that $\kappa_i^\sigma = 0$. If this set is nonempty then the above conditions cannot be satisfied. However, there is still a possibility that the minimum is on the boundary. The condition for the minimum on the boundary is weaker than the above. Namely, all derivatives in directions pointing in toward the set need to be positive. Such directions can be written as a linear combination
\[ \sum_{i \in I \cup I'} \alpha_{ij} \left( \partial_{\kappa_i^\sigma} - \partial_{\kappa_i^\rho} \right) + \sum_{i \in I} \alpha_{ij} \left( \partial_{\kappa_i^\sigma} - \partial_{\kappa_i^\rho} \right), \]
with $\alpha_{ij} \geq 0$ for $i \in I$, $j \not\in I$ and otherwise $\alpha_{ij}$ arbitrary, since the derivatives with negative coefficients at $\partial_{\kappa_i^\sigma}$ would otherwise point outside the set. For that it is enough to check basis derivatives.
\[
\forall i,j \in I, \left( \partial_{\kappa_i^0} - \partial_{\kappa_j^0} \right) \sum_m - \kappa_i^0 \ln \kappa_i^m \geq 0 \Rightarrow \forall i,j \in I, \frac{\kappa_i^0}{\kappa_j^0} \geq 0, \\
\forall i,j \in I, \left( \partial_{\kappa_i^0} - \partial_{\kappa_j^0} \right) \sum_m - \kappa_i^0 \ln \kappa_i^m = 0 \Rightarrow \forall k,j \in I, \frac{\kappa_i^0}{\kappa_j^0} - \frac{\kappa_k^0}{\kappa_j^0} = 0,
\]

and we see this is a global minimum.

Recall that if \( \kappa_i^0 = 0 \) when \( \kappa_j^0 = 0 \) then \( D(\rho, \sigma) = \infty \). We now also need to consider the case that \( \kappa_i^0 = 0 \) when \( \kappa_j^0 \neq 0 \). In that case we would get the full derivatives in the i direction, thus the equations (11) apply, which can not be satisfied unless all \( \kappa_j^0 = 0 \), which can not occur in \( \mathcal{D} \).

Combining this with the above we have that

\[
C_{\mathcal{V}_i}^i \left( \sum_i P_i \rho P_i \right) = \partial_{\mathcal{V}_i} D \left( \rho, \sum_i P_i \rho P_i \right) \geq 0.
\]

The state \( \sigma = \sum_i P_i \rho P_i \) is the only state \( \sigma \) satisfying \( C_{\mathcal{V}_i}^i (\sigma) \geq 0 \). The set \( \mathcal{V}_i \) is convex, so from (12) and convexity of \( D(\rho, \cdot) \), this is the unique global minimum.

4. Strong collapse

The conditions defining ‘strong’ collapse that were specified in Introduction lead us to a troubling situation, because for such states (containing zero eigenvalues) the relative entropy is almost always infinite. We will overcome the problem by deriving a generalised version of the strong collapse rule that is a quantum counterpart of Jeffrey’s rule. The ordinary strong collapse rule will be then obtained by a limiting procedure.

Consider a constraint set given in terms of \( \{ P_i \mid i \in I \} \) such that \( \sum_i P_i = 1 \) by

\[
\mathcal{V}_i := \left\{ \sigma \in \mathcal{D} \mid \left\{ P_i, \sigma \right\} = 0, \quad \text{tr} \left( \sigma P_i \right) = p_i \quad \forall P_i \right\},
\]

where \( \{ P_i \mid i \in I \} \) is again determined by the spectral decomposition of an observable \( O = \sum_i \lambda_i P_i \). The set (13) can be interpreted as encoding the knowledge that the measurement outcome \( \lambda_i \) corresponding to a projection \( P_i \) occurs with a probability \( p_i \).

Here we encounter a problem. If we have a \( p_i \) nonzero but \( \text{tr}(\rho P_i) = 0 \), then every state in \( \mathcal{V}_i \) will have relative entropy \(-\infty\) to \( \rho \). Moreover, even if we subtract the infinite constant, we find that the regularised distance does not depend on the state in the block \( P_i \) and there is no unique minimum. We thus will always assume that \( \text{tr}(\rho P_i) = 0 \) for \( p_i = 0 \).

The variation in the \( U_i \) direction goes through as before. However the variation in the direction of the spectrum changes in that a basis is now given in terms of \( \partial_{\kappa_i^0} - \partial_{\kappa_j^0} \), with \( \kappa_i^0 \) and \( \kappa_j^0 \) belonging to the same block \( P_i \) and thus being eigenvalues of \( \sigma_i \). Thus only the fractions of eigenvalues within each block are fixed. This implies that the eigenvalues of \( \sigma_i \) are uniformly scaled relative to the eigenvalues of \( \rho_i \). The condition

\[
\sum_k \frac{\kappa_i^0}{\kappa_k^0} = p_i \quad \text{fixes} \quad \sigma_i \quad \text{to be} \quad \frac{P_i \rho P_i}{\text{tr}(P_i \rho P_i)}.
\]

This shows that

\[
C_{\mathcal{V}_i}^i \left( \sum_i p_i \frac{P_i \rho P_i}{\text{tr}(P_i \rho P_i)} \right) \geq 0.
\]

The state \( \sigma = \sum_i p_i \frac{P_i \rho P_i}{\text{tr}(P_i \rho P_i)} \) is the only state \( \sigma \) satisfying \( \partial_{\mathcal{V}_i} D(\rho, \sigma) \geq 0 \).

The strong collapse is a limiting case of the above projection, with all \( p_i \) going to zero except of one, \( p_{i_1} \), corresponding to a projection \( P_{i_1} \), that, in turn, corresponds to a measurement result given by an eigenvalue \( \lambda_i \).

We obtain this by taking the weak continuous limit

\[
\lim_{p_1, p_2, \ldots \to 0} \arg \inf_{\sigma \in \mathcal{V}_i} \{ D(\rho, \sigma) \} = \lim_{p_1, p_2, \ldots \to 0} \sum_i p_i \frac{P_i \rho P_i}{\text{tr}(P_i \rho P_i)} = \frac{P_{i_1} \rho P_{i_1}}{\text{tr}(P_{i_1} \rho P_{i_1})}.
\]

Note that in the finite dimensional case that we consider here the weak topology and norm topology coincide.

5. The foundational view

In the orthodox formulation of quantum mechanics the ‘collapse rules’ are postulated. Thus, they are not deduced from any other more fundamental principle. They can be derived from several different conditions, see
for a review, but none of these conditions possesses the status of a fundamental principle of quantum theory. The weak collapse rule can be derived by taking the tensor product with an auxiliary state, followed by unitary evolution and a partial trace. This may serve as a derivation independent of interpretational issues (when this procedure is interpreted as an interaction with some ontic environment, it is usually considered as an instance of decoherence). However, no such construction exists for the strong rule. This fact, as well as the unclear relationship between the strong collapse rule and unitary evolution, renders the orthodox mathematical foundations conceptually insufficient, asking for further insights.

In general, an ontic interpretation of the quantum state leads to considering quantum collapse as a change of the ‘state of being’ of some ‘material object/thing’. On the other hand an epistemic interpretation leads to considering quantum collapse as a change of the ‘state of information’ of some ‘experiencing user/agent’. (There also is a corresponding difference in the meaning of the term ‘measurement’.) In particular, the dynamical reduction approach of [25], belongs to the former class, providing an ontic explanation by means of a general dynamical principle from which the quantum collapse rule is derived. On the other hand, an epistemic interpretation of collapse rules as quantum mechanical analogues of the Bayes–Laplace rule $p(x) \mapsto p(x)p(b|x)/p(b)$ was proposed in [9–12]. However, no epistemic explanation, understood as a derivation from some fundamental principle of information theory (or statistical inference theory) has been offered. Our paper (as well as the closely related paper [26]) provides such a derivation.

Following the postulates of [27, 28] (which aim at reapproaching the foundations of quantum theory in the spirit of [29–32]), we demonstrated that the mapping to the unique solution of constrained minimisation of the relative information $D,$

$$\rho \mapsto \arg\inf_{\sigma \in \mathcal{Q}} \{D(\rho, \sigma)\}, \quad (16)$$

can serve as the general principle of quantum state change due to the acquisition of new information (represented by the constraints $\mathcal{Q}$). This amounts to selecting the quantum state that is the least distinguishable from the original state among all states that are in a strict agreement with the new knowledge (represented by the constraints).

In order to derive the quantum collapse rules from the principle (16), we needed to identify the information theoretic constraints that define the situations of weak and strong collapse. The ‘weak’ collapse amounts to encoding the information that a specific observable $O$ has been subjected to measurement. A quantum state $\sigma$ that carries such information has to be compatible with the possibility of measuring all eigenvalues of $O$ precisely. Such a situation can be characterised by the condition $[\sigma, O] = 0$ (or, equivalently, $[P_i, \sigma] = 0 \forall P_i$). The ‘strong’ collapse should additionally result in a state that would reproduce the result of measurement of a particular eigenvalue with certainty (that is, with probability equal 1). That is, given a projector $P$ encoding the outcome $\lambda$ of the measurement, the post-collapse density operator $\sigma$ should satisfy the condition of a ‘weak’ collapse, as well as $\text{tr}(Pr) = 1.$ This provides an interesting general insight into a structure of quantum theory: why it is possible to use (16) in order to derive various quantum state change rules without assuming the probabilistic interpretation carried under the label of the ‘Bayes rule’, the latter seems to be required for justification of the choice of constraints leading to a specific class of rules, including ‘strong’ collapse.

Our results can be considered as a quantum counterpart of derivations [13–16] of the Bayes–Laplace rule from the constrained maximisation of the Kullback–Leibler relative entropy [33], $S(p, q) := -\int_{\mathcal{X}} \mu(x)p(x)\log(p(x)/q(x))$, where $x \in \mathcal{X}$, while $p$ and $q$ are densities of probability measures with respect to a measure $\mu$ on $\mathcal{X}$. The functional $S(p, q)$ is a special case of Umegaki’s quantum relative entropy $S(\sigma, \rho)$ for discrete $\mathcal{X}$ and $[\sigma, \rho] = 0.$ This strengthens the analogy between the Bayes–Laplace and the von Neumann–Lüders rules: they are just two special cases of a single general principle of inductive inference, given by (16). From the Bayesian perspective, the state $\rho$ is a prior, while $\sigma$, satisfying the constraints and maximising $S(\rho, \sigma)$, is a posterior.

6. Remarks

It has been known for quite a long time (see e.g. [34]) that a ‘weak’ collapse leads to an increase of the absolute entropy $-\text{tr}(\rho \log \rho).$ Our result uncovers an unexpectedly strong inverse of this fact: a ‘weak’ collapse is a result of maximisation of the absolute entropy $-\text{tr}(\rho \log \rho) - \rho \log \rho$ under specific constraints.

All earlier results on derivation of weak and strong collapse rules from minimisation of two point functionals on the space of quantum states [22, 35–42] were obtained for (various) symmetric quantum information distances. The importance of our result stems from the importance of (the negative of) Umegaki’s relative entropy in quantum information theory as opposed to symmetric quantum information distances, which do not carry a similar semantic significance. This statement can be approached either axiomatically or pragmatically. On the axiomatic side, $D(\rho, \sigma)$ is characterised [43] by the direct sum property, invariance under
autormorphisms (so, in particular, unitaries), additive decomposition under conditional expectations (onto subalgebra), and measurability over the state space. These properties eliminate all above symmetric information distances. An extensive discussion of the reasons for (and, in particular, applications of) these properties can be found in [48]. Furthermore, \( D(\rho, \sigma) \) is a direct quantum generalisation of \( -S(\rho, q) \), which can be also characterised ‘dynamically’ as a unique functional \( \Phi(p, q) \) such that the mapping \( q \mapsto \arg\min_{p \in \mathbb{C}} \{ \Phi(p, q) \} \) satisfies a few very reasonable desiderata for information processing \([15, 32, 44]\). On the pragmatic side, \( D(\rho, \sigma) \) is widely used in quantum information theory as the most fundamental measure of distinguishability of quantum states (see e.g. [7, 8, 45–47]). Hence, from the perspective of quantum information theoretic approaches to foundations of quantum theory, our results provide an essential, new perspective on the mathematical form of collapse rules due to quantum measurement.

As noted by one of the referees, this leads to a question whether the results of this paper can be reproduced (or extended) in the setting of generalised probabilistic theories \([49, 50] \). This setting lacks a general analogue of the collapse rules, but it allows us to introduce a well defined notion of information distance \([51, 52] \) (which reduces to the Umegaki and Kullback–Leibler distances in quantum mechanical and probabilistic case, respectively). Hence, the possible extension of our result to generalised probabilistic theories can bring in new foundational insights (in particular—as suggested by a referee—one can ask whether defining a post-measurement state as a minimiser of a specific information distance given some type of constraints preselects some type of theories). The main open technical problem is how to replace the use of block diagonal decomposition and variational analysis of the spectrum of operators by some other method. It may be possible that a restriction to a subclass of theories satisfying some sort of spectral condition (see e.g. \([56, 57] \)) will be necessary for this. We hope to return to this problem in another paper.

Acknowledgments

We would like to thank Carlos S Guedes for many important and insightful discussions throughout the development of this result. We thank also Patrick Coles for informing us about \([17, 18] \), and Daniel Ranford for some suggestions and comments. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Research and Innovation. This research was also partially financed by the National Science Center of the Republic of Poland (Narodowe Centrum Nauki) through the grant number DEC2011/01/N/HS3/03273.

References

[8] Vedral V 2002 The role of relative entropy in quantum information theory Rev. Mod. Phys. 74 197
[16] Douven I and Romeijn JW 2012 A new resolution of the Judy Benjamin problem Mind 479 637

This problem seems to reflect quite similar issue in quantum logic, where the collapse rules are also not a part of the basic framework, so they require an additional justification or derivation (see e.g. \([53] \)), pointing towards a possibility of some more fundamental principle. Such derivations were provided (under some assumptions) by means of minimisation of suitable symmetric distances (see e.g. \([38, 39, 42] \)), corresponding (in some cases) to symmetric transition probability functionals. It was shown in \([34] \) that nonsymmetric transition probability \([33] \) plays more fundamental role in quantum logic, but its relationship to nonsymmetric distances and to the problem of derivation of collapse rules was not investigated.
[22] Herbst F 1969 Derivation of the change of state in measurement from the concept of minimal measurement Ann. Phys. 35 271
[34] Wehrl A 1978 General properties of entropy Rev. Mod. Phys. 50 221
[53] Stairs A 1982 Discussion: quantum logic and the Lüders rule Phil. Sci. 49 422