A DECISION-THEORETIC APPROACH TO DEALING WITH UNCERTAINTY IN QUANTUM MECHANICS

KEANO DE VOS, GERT DE COOMAN, ALEXANDER ERREYGERS, AND JASPER DE BOCK

ABSTRACT. We provide a decision-theoretic framework for dealing with uncertainty in quantum mechanics. This uncertainty is two-fold: on the one hand there may be uncertainty about the state the quantum system is in, and on the other hand, as is essential to quantum mechanical uncertainty, even if the quantum state is known, measurements may still produce an uncertain outcome. In our framework, measurements therefore play the role of acts with an uncertain outcome and our postulates ensure that Born's rule is encapsulated in the utility functions associated with such acts. This approach allows us to uncouple (precise) probability theory from quantum mechanics, in the sense that it leaves room for a more general, so-called imprecise probabilities approach. We discuss the mathematical implications of our findings, which allow us to give a decision-theoretic foundation to recent seminal work by Benavoli, Facchini and Zaffalon, and we compare our approach to earlier and different approaches by Deutsch and Wallace.

1. INTRODUCTION AND OVERVIEW

In dealing with a quantum system, there are various reasons why some subject, whom we'll call You, might be uncertain about the present state it's in: the state may have been prepared by performing a measurement, or You might be uncertain about the dynamics that brought it to its present state, or about the state it started out from, to name a few of them.

This uncertainty is typically represented, and reasoned with, using (quantum) probabilities, in the form of density operators. This type of probabilistic model, pervasive in the literature, has a characteristic that it shares with the probabilistic models that are commonly used in more classical — non-quantum — contexts: it has a very high informational content. That it does, is exemplified by the fact that when such models are used in a decision-theoretic (expected utility) framework, they leave no room for indecision: they allow You to choose between any two options and provide a definite answer to every yes-or-no question. While this isn't a problem or a drawback in itself, it does call into question why such decisiveness is always justified in all contexts where uncertainty is present. And this question is made all the more incisive by the following observation: while probabilities are often seen as constituting a superstructure built on a foundation of classical (propositional) logic, they don't share, in the way they're commonly represented, with this logic one of its most crucial properties: its ability to represent and deal with *partial information* through its conservative *deductive inference* mechanism.

This 'problem' has been recognised for quite some time in classical (non-quantum) uncertainty contexts, and the research field of *imprecise probabilities* (see, for instance, Refs. [2, 46, 50, 51]) has devoted quite some time and effort to allowing the mutually related aspects of indecision, imprecision and partial information to also play their important constitutive role in probability theory. It was, arguably and to the best of our knowledge, Walley who first drew attention to the conservative inference mechanism that is hidden behind classical (non-quantum) probability theory [50, 51] and one of us (De Cooman) who first drew attention to the analogy between that conservative inference mechanism and deductive inference in classical propositional logic [18]. In a nutshell, the so-called

Key words and phrases. uncertainty; decision theory; desirability; lower prevision; prevision; credal set; density operator; quantum mechanics.

imprecise probability models (such as there are: sets of probability measures, coherent lower and upper previsions, coherent sets of desirable gambles and coherent partial preference orderings) represent partial information states that have the same role as deductively closed and logically consistent sets of propositions do in classical propositional logic; and it's the precise models amongst them (such as there are: single probability measures, coherent previsions, maximal coherent sets of desirable gambles and coherent total preference orderings, respectively) that correspond to the maximal (so-called *complete*) deductively closed and logically consistent sets of propositions, to which no further information can be added without producing inconsistencies. On this way of looking at things, insisting that all uncertainty must be modelled by probabilities is very much like claiming that the only reasonable sets of propositions are the complete ones. We'll have ample occasion to explain, and provide pointers to, the relevant technical imprecise probabilities literature dealing with such models as coherent sets of desirable gambles, coherent lower previsions and sets of probability measures further on in the paper, but as far as their connection with conservative inference is concerned, we want to take the opportunity here to also refer readers to Refs. [12, 18, 19, 22], which are relevant explorations of this issue.

We believe that Benavoli, Facchini and Zaffalon were the first to bring the imprecise probabilities framework to bear on dealing with uncertainty in quantum mechanics in a finite-dimensional context. In a number of ground-breaking papers [6, 7], they essentially associated specific types of desirable gambles with measurements, and showed how coherent sets of such desirable gambles can be connected to sets of density operators.

One of the things we want to achieve in this paper, is to provide their uncertainty models with a more solid decision-theoretic foundation and to show that the exact 'Born rule'-like form they give to the desirable gamble that is associated with a measurement on a quantum-mechanical system can be derived from decision-theoretic principles.

We're well aware that we aren't the only ones — and definitely also not the first — to try and provide a decision-theoretic foundation for dealing with uncertainty in quantum mechanics. To name only a few others, Deutsch [23] and Wallace [47–49] have tried to show that in an Everettian (many-worlds) approach to quantum mechanics, a few simple decision-theoretic principles are enough to derive the probabilistic quantum mechanical postulates from the non-probabilistic ones. Their aim in this has been to justify the use of probabilities (Born's rule) in a quantum mechanical context, or to use Deutsch's phrase, to justify

> [...] deriving a probability statement from a factual statement. This could be called deriving a 'tends to' from a 'does'.

In this paper, we don't necessarily want to repeat or improve on their existing arguments and we most decidedly aren't looking for ways to justify using (precise) probabilities in a quantum context. But we are looking for ways to deal with the existing uncertainty in quantum mechanics using a decision-theoretic toolbox. In doing so, we also want to show that more general models than probabilities are reasonable and useful for dealing with epistemic uncertainty in quantum mechanics. As we'll explain in much more detail further on in Section 3, we'll place the decision-theoretic argument in a de Finetti rather than a Savage framework, which will then allow us to deal more easily with partial preference, imprecision and indecision. Indeed, rather than postulate Born's rule (and the concomitant existence of probabilities) directly, we propose (like Deutsch and Wallace) a conglomerate of decision-theoretic postulates that seem less invasive, and which allow us to derive Born's rule as a special case, while allowing (unlike Deutsch and Wallace) for more general (imprecise or partially specified probabilities) models in less restrictive cases. This will allow us to put into perspective the — to our mind unjustifiably strong — uniquely central role that precise probabilities often seem to play in quantum mechanics.

How have we structured the discussion? We begin in Section 2 with a brief review of the basics of quantum mechanics, and in particular of its (non-probabilistic and probabilistic) postulates, which we'll have occasion to refer to and build on later.

In Section 3, we describe the decision-theoretic foundations for our approach to dealing with uncertainty in quantum mechanics, we formulate and motivate our (decision-theoretic) postulates, and we announce the important conclusions that can be drawn from them; we postpone proving these conclusions until Section 10. At the same time, we sketch the decision-theoretic background that will help readers place our subsequent discussion in its proper context and compare it in Sections 7 and 8 to the approaches followed by Deutsch and Wallace. In a nutshell, we show in this section that when measurements $\hat{A}$ are interpreted as acts with an uncertain reward $w_{\hat{A}}(|\Psi\rangle)$, which depends on the (possibly uncertain) quantum state $|\Psi\rangle$ and which is expressed in units of some predetermined linear utility scale, a number of simple postulates fix the form of this reward, as they guarantee that $w_{\hat{A}}(|\Psi\rangle) = \langle\Psi|\hat{A}|\Psi\rangle$ — we'll explain the notations in Section 2.

The decision-theoretic upshot of our results in Section 3 (and their proofs in Section 10) is that uncertainty about the quantum state $|\Psi\rangle$ of a system under consideration can be described by a strict partial vector ordering on the uncertain rewards $w_{\hat{A}}$, or equivalently, on the measurements $\hat{A}$. We describe the mathematical consequences of using such uncertainty models, and alternative representations for them, in Sections 4 to 6. This is where we show that the models introduced by Benavoli, Facchini and Zaffalon [6, 7] can be derived within our decision-theoretic framework. We also provide proofs for our claims there, even though some of them can be found elsewhere in some form or other, mainly in the interest of making this discussion as self-contained as possible.

In Section 4, we show how these strict partial vector orderings can be represented by so-called *coherent sets of desirable measurements*, and that such models provide a means to perform conservative inference based on partial desirability statements about various measurements. In Section 5 we go on to introduce a slightly simpler and less general type of model that has the advantage of being Archimedean, meaning that the preferences can be expressed using the real number scale; these are the coherent lower (and upper) prevision functionals, or equivalently, convex closed sets of density operators. When lower and upper prevision functionals coincide, they turn into coherent previsions; equivalently, when the convex closed sets of density operators collapse to singletons, we recover the more classical case of quantum probability — working with density operators — as a special case, as discussed in Section 6. We'll also discover, incidentally, that the mathematical (almost-)equivalence between working with sets of measurements on the one hand in Section 4 and working with sets of density operators on the other in Section 6, allows us to recover the well-known duality between the Heisenberg and Schrödinger pictures in our decision-theoretic approach as well. Throughout Sections 3 to 6, we illustrate the various models and arguments using a series of simple examples involving qubits.

Section 6.4 is of special importance, because we show there that our general sets of desirable measurements approach, based on the representation result in Section 3, collapses to Born's rule when the state $|\Psi\rangle$ is known. This allows us to compare in some detail our present approach to the earlier ones by Deutsch (in Section 7) and Wallace (in Section 8), and to point out the relevant differences in Section 9.

Finally, we spend some time in Section 11 on rehearsing the main themes in this paper and drawing conclusions from them. We also point to useful and interesting ways that the uncertainty models we're about to discuss in detail, might be used advantageously in practical computational tasks, which also opens up paths for future research.

2. THE QUANTUM-MECHANICAL BASICS

To lay the foundations for the coming discussion and to sketch the necessary context, we start with a concise account of those foundational tenets of quantum mechanics that we'll need, and which we'll present here in the form of seven basic postulates. For a more thorough account of the foundations and more details about what we leave unexplained, we refer to the basic textbooks by Cohen-Tannoudji, Laloe and Diu [9] and Nielsen and Chuang [37], where readers can also find a discussion of those aspects of linear algebra that are necessary for understanding the mathematical underpinnings of these postulates, and which we'll assume them to be familiar with.

We begin with those postulates that don't mention, or rely on, probabilistic notions.

2.1. **The non-probabilistic postulates.** With every quantum system, we associate a complex *Hilbert space* $\mathscr{X}$, with an *inner product* $(\bullet_1,\bullet_2)$ and associated norm $\|\bullet\| := \sqrt{(\bullet,\bullet)}$.

We'll use Dirac's notation and terminology [24] throughout, so we'll call an element $|\psi\rangle$ of $\mathscr{X}$ a *ket*. The corresponding *bra*, or dual ket, is then generically denoted by $\langle\psi|$: it's the continuous[1] linear functional on $\mathscr{X}$ defined by $\langle\psi|(|\phi\rangle) := (|\psi\rangle,|\phi\rangle)$ for all $|\phi\rangle \in \mathscr{X}$. The inner product of two kets $|\psi\rangle$ and $|\phi\rangle$, or in other words, the image of the ket $|\phi\rangle$ under the bra $\langle\psi|$, is then conveniently denoted by $\langle\psi|\phi\rangle$.

For any subset $\mathscr{A}$ of $\mathscr{X}$, its *linear span*

$$\operatorname{span}(\mathscr{A}) := \left\{\sum_{k=1}^{n} \lambda_k|\phi_k\rangle : n \in \mathbb{N}, |\phi_1\rangle,\dots,|\phi_n\rangle \in \mathscr{A}, \lambda_1,\dots,\lambda_n \in \mathbb{C}\right\}$$

is the smallest linear subspace of $\mathscr{X}$ that includes $\mathscr{A}$. Here and in what follows, we denote by $\mathbb{N}$ the set of all natural numbers (without zero), and by $\mathbb{C}$ the set of all complex numbers.

A *state* is a normalised, or normal, ket $|\psi\rangle \in \mathscr{X}$, which means that $\langle\psi|\psi\rangle = 1$. We'll denote the set of possible states — the *state space* — by $\bar{\mathscr{X}} := \{|\psi\rangle \in \mathscr{X} : \langle\psi|\psi\rangle = 1\}$.

Quantum mechanical postulate 1 (QM1)
At any fixed time, the state of a physical system is represented by a normalised ket $|\psi\rangle$*, which is an element of the system's state space* $\bar{\mathscr{X}}$.[2]

To deal with certain aspects of quantum-mechanical systems, such as location, infinite-dimensional Hilbert spaces are essential. But to keep the discussion as simple as possible, *we'll restrict ourselves in this paper to the case of finite-dimensional Hilbert spaces*, which can for instance be used to model such aspects as the spin or (with some extra assumptions, such as ignoring the higher energy levels) the energy of a bounded electron. Such finite-dimensional spaces are particularly useful in quantum computing and quantum cryptography [37].

One way to interact with a quantum system is through measurements, which are represented by Hermitian operators. The *Hermitian conjugate* of a linear operator $\hat{O}$ on $\mathscr{X}$ is the unique linear operator $\hat{O}^\dagger$ on $\mathscr{X}$ such that $(\hat{O}^\dagger|\psi\rangle,|\phi\rangle) = (|\psi\rangle,\hat{O}|\phi\rangle)$ for all $|\psi\rangle,|\phi\rangle \in \mathscr{X}$. A linear operator $\hat{A}$ on $\mathscr{X}$ is called *Hermitian* if it's equal to its Hermitian conjugate $\hat{A}^\dagger$, and we can then use Dirac's notation $\langle\psi|\hat{A}|\phi\rangle := (|\psi\rangle,\hat{A}|\phi\rangle) = (\hat{A}|\psi\rangle,|\phi\rangle)$ without any possible confusion. Observe, by the way, that it follows from the properties of an inner product that[3]

$$\langle\psi|\hat{A}|\phi\rangle = (|\psi\rangle,\hat{A}|\phi\rangle) = (\hat{A}|\phi\rangle,|\psi\rangle)^* = \langle\phi|\hat{A}|\psi\rangle^* \text{ for all } |\psi\rangle,|\phi\rangle \in \mathscr{X}, \tag{1}$$

and therefore also that $\langle\phi|\hat{A}|\phi\rangle$ is real for all $|\phi\rangle \in \mathscr{X}$. We denote the set of all Hermitian operators on $\mathscr{X}$ by $\mathscr{H}$, or by $\mathscr{H}(\mathscr{X})$ should we want to avoid confusion about the Hilbert space $\mathscr{X}$ they're defined on.

[1] In the topology induced by the norm $\|\bullet\|$, the inner product and the norm are continuous.

[2] Strictly speaking, the state of a system is represented by a *ray* of kets, typically characterised by one of its elements: a normalised one. This implies that if the normalised ket $|\psi\rangle$ represents the state of a system, then so does $e^{i\alpha}|\psi\rangle$ for any $\alpha \in \mathbb{R}$; a normalised ket represents a system's state uniquely, up to a phase factor. We'll always implicitly disregard this phase factor when considering the state of a system.

[3] We denote the complex conjugate of $a \in \mathbb{C}$ by a^*.

Quantum mechanical postulate 2 (QM2)
Every measurable physical quantity of a quantum system with state space $\bar{\mathscr{X}}$ can be described by a Hermitian operator $\hat{A} \in \mathscr{H}(\mathscr{X})$.

In fact, in finite-dimensional Hilbert spaces it's also often tacitly assumed that all Hermitian operators correspond to physical measurements, and we'll also make that assumption here.

When we measure a physical quantity associated with a system, this measurement produces a real number, which we'll call the *outcome* of the measurement; we'll denote the set of all real numbers by $\mathbb{R}$. But it turns out that, typically, for a given measurement $\hat{A}$, not all real numbers are possible outcomes for the measurement. This is where the third postulate comes in, for the formulation of which we recall the notion of eigenvalues and eigenkets.

A complex number λ is an *eigenvalue* of a linear operator $\hat{O}$ on $\mathscr{X}$ if there's some non-null ket $|a\rangle \in \mathscr{X}$ such that $\hat{O}|a\rangle = \lambda|a\rangle$. Any ket $|a\rangle$ for which $\hat{O}|a\rangle = \lambda|a\rangle$ is then called an *eigenket* of $\hat{O}$ corresponding to the eigenvalue λ; we'll denote the set of all such eigenkets by $\mathscr{E}_\lambda$. A normal eigenket is called an *eigenstate*. It's clear that any linear combination of eigenkets in $\mathscr{E}_\lambda$ is still an eigenket, so $\mathscr{E}_\lambda$ is a linear subspace of $\mathscr{X}$, called the *eigenspace* of $\hat{O}$ corresponding to the eigenvalue λ. The set of all the eigenvalues of the operator $\hat{O}$ is called its *spectrum* and denoted by $\operatorname{spec}(\hat{O})$.

Quantum mechanical postulate 3 (QM3)
The only possible outcomes of a measurement are the eigenvalues of the corresponding Hermitian operator $\hat{A} \in \mathscr{H}$.

The fact that all eigenvalues of any Hermitian operator are always real and that eigenstates corresponding to different eigenvalues are always orthogonal [37, Exercise 2.17], gives some intuition as to why the measurement operators are (taken to be) Hermitian. Since the outcome of a physical measurement must be a real number, we must therefore consider measurement operators with real eigenvalues, and this is a property that Hermitian operators have. In fact, the knowledge of both the eigenvalues and their corresponding eigenstates is enough to determine the corresponding Hermitian operator. To explain this in the following proposition, we first introduce a convenient new notation. For any ket $|\psi\rangle$, we'll denote by $|\psi\rangle\langle\psi|$ the linear operator on $\mathscr{X}$ that maps any ket $|\phi\rangle$ to the ket $(|\psi\rangle\langle\psi|)(|\phi\rangle) := |\psi\rangle\langle\psi|\phi\rangle$, which is a scalar multiple of the ket $|\psi\rangle$ — in fact, it's the orthogonal projection of $|\phi\rangle$ on the linear subspace spanned by $|\psi\rangle$. It's a trivial exercise to show that $|\psi\rangle\langle\psi|$ is Hermitian. Indeed, any Hermitian operator is a real linear combination of specific operators of this type.

Proposition 1 ([37, Box 2.2])
Let $\lambda_1, \ldots, \lambda_n$ be any real numbers and let $\{|a_1\rangle, \ldots, |a_n\rangle\}$ be any orthonormal basis for the n-dimensional Hilbert space $\mathscr{X}$. Then $\hat{A} := \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k|$ is the only operator that has $\lambda_1, \ldots, \lambda_n$ as eigenvalues with respective eigenstates $|a_1\rangle, \ldots, |a_n\rangle$, and this operator is Hermitian. Conversely, every Hermitian operator $\hat{A} \in \mathscr{H}$ can be written as $\hat{A} = \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k|$, with $\{|a_1\rangle, \ldots, |a_n\rangle\}$ an orthogonal basis of eigenstates[4] for $\hat{A}$ with corresponding respective real eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

A *projection operator* $\hat{P}$ is a Hermitian operator such that $\hat{P}\hat{P} = \hat{P}$. Due to its hermicity, a projection operator is an orthogonal projection: $(\hat{P}|\phi\rangle, |\phi\rangle - \hat{P}|\phi\rangle) = (\hat{P}|\phi\rangle, |\phi\rangle) - (\hat{P}|\phi\rangle, \hat{P}|\phi\rangle) = (\hat{P}|\phi\rangle, |\phi\rangle) - (\hat{P}^\dagger\hat{P}|\phi\rangle, |\phi\rangle) = 0$ for all $|\phi\rangle \in \mathscr{X}$. Moreover, since for any eigenvalue $\lambda \in \operatorname{spec}(\hat{P})$ with corresponding eigenket $|\phi\rangle$, $\lambda|\phi\rangle = \hat{P}|\phi\rangle = \hat{P}^2|\phi\rangle = \lambda^2|\phi\rangle$, we see that $\operatorname{spec}(\hat{P}) \subseteq \{0, 1\}$.

For any linear subspace $\mathscr{E}$ of $\mathscr{X}$, we'll denote by $\hat{P}_\mathscr{E}$ the unique projection operator $\hat{P}$ whose range $\operatorname{rng}(\hat{P}) := \{\hat{P}|\phi\rangle : |\phi\rangle \in \mathscr{X}\}$ is equal to $\mathscr{E}$. Given an m-dimensional linear

[4] Throughout, we'll use the term 'orthogonal basis of eigenstates' rather than 'orthonormal basis of eigenstates', because the normality is already implied by our using the term 'eigen*states*'.

subspace $\mathscr{E}$ of $\mathscr{X}$ and any orthonormal basis $\{|b_1\rangle, \dots, |b_m\rangle\}$ for $\mathscr{E}$, this projection operator can be written as $\hat{P}_{\mathscr{E}} = \sum_{k=1}^{m} |b_k\rangle\langle b_k|$; see Proposition 1. If $\mathscr{E}_\lambda$ is the eigenspace corresponding to the eigenvalue λ of some Hermitian operator, then we'll also use the notation $\hat{P}_\lambda$ for the projection operator $\hat{P}_{\mathscr{E}_\lambda}$ with range $\mathscr{E}_\lambda$.

Projection operators allow us to represent any Hermitian operator in terms of its eigenvalues and eigenspaces.

Corollary 2 ([37, Box 2.2])
Consider $m \leq n$ distinct real numbers $\lambda_1, \dots, \lambda_m$ and let $\mathscr{E}_1, \dots, \mathscr{E}_m$ be orthogonal subspaces of an n-dimensional Hilbert space $\mathscr{X}$ that span $\mathscr{X}$, then the Hermitian operator $\hat{A} := \sum_{k=1}^{m} \lambda_k \hat{P}_{\mathscr{E}_k}$ is the unique operator with eigenvalues $\lambda_1, \dots, \lambda_m$ and corresponding eigenspaces $\mathscr{E}_{\lambda_1} = \mathscr{E}_1, \dots, \mathscr{E}_{\lambda_m} = \mathscr{E}_m$. Conversely, any Hermitian operator $\hat{A} \in \mathscr{H}$ can be uniquely written as $\hat{A} = \sum_{k=1}^{m} \lambda_k \hat{P}_{\mathscr{E}_{\lambda_k}}$, with $\lambda_1, \dots, \lambda_m$ its distinct real eigenvalues and with orthogonal corresponding eigenspaces $\mathscr{E}_{\lambda_1}, \dots, \mathscr{E}_{\lambda_m}$ that span $\mathscr{X}$.

Postulate QM3 fixes $\mathrm{spec}(\hat{A})$ as the set of all possible outcomes of a measurement $\hat{A}$, but says nothing about which of these outcomes will actually be obtained. The following postulate resolves this uncertainty in an important particular case.

Quantum mechanical postulate 4 (QM4)
Performing a measurement with Hermitian operator $\hat{A} \in \mathscr{H}$ on a quantum system in an eigenstate $|a\rangle$ of $\hat{A}$ results in the corresponding eigenvalue λ being measured with certainty.

In quantum mechanics, measurements seldom leave the state untouched, as is made clear by the following postulate. We won't rely on it for our main argument in Section 3 (and Sections 4 to 6 and 10). But since it allows for ways to make sure that a quantum system is in a known state — by performing a measurement with an outcome that corresponds to a one-dimensional eigenspace — it will enable us in Section 6.4 to consider a special case where we can compare our approach to earlier discussions by Deutsch [23] and Wallace [47–49].

Quantum mechanical postulate 5 (QM5)
If a measurement with Hermitian operator $\hat{A} \in \mathscr{H}$ on a quantum system in a state $|\psi\rangle \in \bar{\mathscr{X}}$ results in the eigenvalue λ, then the state $|\phi\rangle$ of the system immediately after the measurement is given by[5]

$$|\phi\rangle = \frac{\hat{P}_\lambda |\psi\rangle}{\sqrt{\langle\psi|\hat{P}_\lambda|\psi\rangle}}.$$

The last non-probabilistic postulate describes the dynamics of a quantum-mechanical system — how its state changes over time.

Quantum mechanical postulate 6 (QM6)
The temporal evolution of the state $|\psi(t)\rangle$ as a function of the time t, is governed by the Schrödinger equation:[6]

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle,$$

where $\hat{H}(t)$ is the (possibly time-dependent) Hamiltonian of the system.

---

[5]This $|\phi\rangle$ is the normalised version of $\hat{P}_\lambda|\psi\rangle$.

[6]Here, $\hbar = h/2\pi$, where h is Planck's constant.

As this paper is focused on the uncertainty related to measurements in quantum mechanics, the dynamical aspects will have little or no bearing on the present discussion, but we still include this postulate for the sake of completeness.[7]

2.2. **Probabilities in quantum mechanics.** As we've already mentioned, Postulate QM3 tells us what the possible outcomes of a measurement are, but tells us nothing about which of these possible outcomes will actually be observed. Postulate QM4 resolves the remaining uncertainty whenever the quantum state belongs to an eigenspace of the measurement, but fails to do so in more general cases, where the quantum state is a *superposition* — a linear combination — of kets belonging to different eigenspaces. In classical accounts of quantum mechanics, this is where probabilities come into play. The final postulate, also known as *Born's rule*, attaches a specific probability to each possible outcome.

Quantum mechanical postulate 7 (QM7)
When a measurement with corresponding Hermitian operator $\hat{A}$ on $\mathscr{X}$ is executed on a system in state $|\psi\rangle \in \bar{\mathscr{X}}$, then the probability $p(\lambda\,\|\,|\psi\rangle)$ of measuring the eigenvalue λ of $\hat{A}$ is given by

$$p(\lambda\,\|\,|\psi\rangle) := \langle\psi|\hat{P}_\lambda|\psi\rangle. \tag{2}$$

This is in similar spirit to Postulate QM4, where if the system resides in the eigenstate $|a\rangle$, the probability of observing the corresponding eigenvalue λ is 1.

As an immediate consequence of this postulate, the linearity properties of the inner product and Corollary 2, the *expected outcome* $\mathrm{E}(\hat{A}\,\|\,|\psi\rangle)$ of the measurement $\hat{A}$ is then given by

$$\begin{aligned}\mathrm{E}(\hat{A}\,\|\,|\psi\rangle) &= \sum_{\lambda\in\mathrm{spec}(\hat{A})}\lambda\, p(\lambda\,\|\,|\psi\rangle) = \sum_{\lambda\in\mathrm{spec}(\hat{A})}\lambda\langle\psi|\hat{P}_\lambda|\psi\rangle = \left\langle\psi\middle|\sum_{\lambda\in\mathrm{spec}(\hat{A})}\lambda\hat{P}_\lambda\middle|\psi\right\rangle \\ &= \langle\psi|\hat{A}|\psi\rangle.\end{aligned} \tag{3}$$

Born's rule, in the guise of Eqs. (2) and (3),[8] postulates the existence of such probabilities (and expectations), but it leaves open the question of how they should be interpreted. In the many interpretations of quantum mechanics, they tend to acquire a different meaning and/or justification.

Some interpretations of quantum mechanics, such as the Copenhagen interpretation [28], interpret the Born probabilities as physical and frequentist: they tend to insist that these probabilities, and their interpretation as limit frequencies, have been corroborated by numerous experiments, and that they're physical, objective properties attached to the system; on such views, the universe is indeterministic.

Other interpretations, such as QBism [30], insist that these probabilities are epistemic and that they characterise some subject's knowledge, or the lack thereof, about the quantum system. This view is compatible with an agnostic stance about whether the universe is deterministic.

The Everettian world view is deterministic, which might lead us to suspect that the Born probabilities there must necessarily have some epistemic interpretation. Indeed, Deutsch [23] and later Wallace [47–49], have advanced the claim that under certain (rationality) assumptions, and taking into account the non-probabilistic quantum mechanical postulates, any rational decision-maker must bet on the possible outcomes of a measurement using the betting rates supplied by Born's rule. In this sense, on the Everettian view, their claim is that Born's rule follows from the non-probabilistic postulates and basic tenets of rational decision-making. We'll come back to Deutsch and Wallace's arguments in later sections.

[7]Of course, the dynamics will become relevant when we want to account for how the uncertainty changes with time, which we'll leave for future work.

[8]Eq. (2) can be seen as a special case of Eq. (3), for the specific choice $\hat{A} \rightsquigarrow \hat{P}_\lambda$. We'll come across an even more general version, or formulation, further on in Eq. (6).

Leaving aside the different possible interpretations of the Born probabilities for now, we want to stress that the uncertainty about the outcome of a measurement when the system is in a given state $|\psi\rangle$, isn't the only type of uncertainty that can be considered in quantum mechanics: it's also possible that we have imperfect knowledge about the actual state that the system resides in. In the traditional formulation of quantum mechanics, such so-called *epistemic uncertainty* about the system state is described by so-called *mixed states*. It's essential for the later discussion that we recall here in some detail how it can be represented mathematically and how this mathematical representation can be combined with Born's rule in the shape of Eq. (3).

A quantum system is said to be in a *mixed state* if it's believed to be in one of a finite number of possible quantum states $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle \in \bar{\mathscr{X}}$, with respective probabilities $p_1, p_2, \dots, p_m$, where $m \in \mathbb{N}$. These probabilities p_k therefore describe epistemic uncertainty about the actual state the system is in.

The outcome of a measurement on a system in a mixed state is uncertain because there is, on the one hand, this epistemic uncertainty about which of the possible states that system is in, and because in each of these possible states, the outcome of the measurement will be uncertain according to Postulate QM7. To distinguish between these two types of uncertainty, we'll also refer to the uncertainty about the outcome of a measurement on a system in a given pure state as *quantum-mechanical*.[9]

It turns out that, as we'll explain below, the epistemic uncertainty associated with a mixed state can be very conveniently described using its so-called density operator

$$\hat{\rho} := \sum_{k=1}^{m} p_k |\psi_k\rangle\langle\psi_k|. \tag{4}$$

Generally speaking, a *density operator* $\hat{\rho}$ is defined as a Hermitian operator such that $\mathrm{Tr}(\hat{\rho}) = 1$ and $\hat{\rho} \geq \hat{0}$. We'll denote the convex set of all density operators on $\mathscr{X}$ by $\mathscr{R}$, a subset of $\mathscr{H}$.

To explain the notations we've just used, recall that the *trace* $\mathrm{Tr}(\hat{A})$ of a Hermitian operator $\hat{A}$ is the sum of its eigenvalues. Given any orthonormal basis $\{|b_1\rangle, \dots, |b_n\rangle\}$ for the n-dimensional Hilbert space $\mathscr{X}$, the trace can also be written as $\mathrm{Tr}(\hat{A}) = \sum_{k=1}^{n} \langle b_k|\hat{A}|b_k\rangle$. Moreover, a Hermitian operator $\hat{A} \in \mathscr{H}$ is called *positive semidefinite*, which we'll write as $\hat{A} \geq \hat{0}$, if $\langle\psi|\hat{A}|\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathscr{X}$, or equivalently, if all its eigenvalues are nonnegative.

It isn't hard to see that the operator $\hat{\rho}$ in Eq. (4) is indeed a density operator. As a matter of fact, any density operator can be rewritten as some such finite convex mixture of projection operators based on states. But, such finite convex mixtures aren't unique in determining the density operator $\hat{\rho}$.

Proposition 3 ([32, Problem 10.1])
An operator $\hat{\rho}$ is a density operator if and only if there are states $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_m\rangle \in \bar{\mathscr{X}}$ and real numbers $\alpha_1, \alpha_2, \dots, \alpha_m \in [0,1]$ such that $\sum_{k=1}^{m} \alpha_k = 1$ and $\hat{\rho} = \sum_{k=1}^{m} \alpha_k |\psi_k\rangle\langle\psi_k|$.

The so-called *pure states* are then (degenerate) special cases: the density matrix $\hat{\rho}$ corresponding to a pure state $|\psi\rangle \in \bar{\mathscr{X}}$ is the projection operator $\hat{\rho} = |\psi\rangle\langle\psi|$ on that state.

Let's now argue briefly why the concept of a density operator $\hat{\rho}$ associated with a mixed state in Eq. (4), is so useful.

Consider any Hermitian measurement operator $\hat{A} \in \mathscr{H}$ with eigenvalues $\lambda_1, \dots, \lambda_m$ and corresponding respective eigenspaces $\mathscr{E}_{\lambda_1}, \dots, \mathscr{E}_{\lambda_m}$ and corresponding respective projection operators $\hat{P}_{\lambda_1}, \dots, \hat{P}_{\lambda_m}$, then we gather from Corollary 2 that $\hat{A} = \sum_{\ell=1}^{m} \lambda_\ell \hat{P}_{\lambda_\ell}$. We now perform the measurement $\hat{A}$ on the quantum system in its mixed state.

[9] Depending on the interpretation of quantum mechanics, quantum-mechanical uncertainty may be seen as physical, or also as epistemic; see the discussion above.

First, it follows from a straightforward application of the Law of Total Probability that the probability $p(\lambda_\ell)$ of observing eigenvalue λ_ℓ is given by

$$\begin{aligned} p(\lambda_\ell) &= \sum_{k=1}^{m} p\big(\lambda_\ell \big| |\psi_k\rangle\big) p_k = \sum_{k=1}^{m} p_k \langle\psi_k|\hat{P}_{\lambda_\ell}|\psi_k\rangle = \sum_{k=1}^{m} p_k \operatorname{Tr}\big(|\psi_k\rangle\langle\psi_k|\hat{P}_{\lambda_\ell}\big) \\ &= \operatorname{Tr}(\hat{\rho}\hat{P}_{\lambda_\ell}), \end{aligned} \tag{5}$$

where the second equality follows from Born's rule [QM7] in its simplest form (2); the third equality follows from the cyclic nature of the trace;[10] and the final equality follows from the linearity of the trace.

As we now know the probabilities $p(\lambda_\ell)$ of the different possible outcomes λ_ℓ of the measurement $\hat{A}$, it's a simple matter to calculate the corresponding *expected outcome*:

$$\begin{aligned} \mathrm{E}_{\hat{\rho}}(\hat{A}) &:= \sum_{\ell=1}^{m} \lambda_\ell p(\lambda_\ell) = \sum_{\ell=1}^{m} \lambda_\ell \operatorname{Tr}(\hat{\rho}\hat{P}_{\lambda_\ell}) = \operatorname{Tr}\left(\sum_{\ell=1}^{m} \lambda_\ell \hat{\rho}\hat{P}_{\lambda_\ell}\right) = \operatorname{Tr}\left(\hat{\rho}\sum_{\ell=1}^{m} \lambda_\ell \hat{P}_{\lambda_\ell}\right) \\ &= \operatorname{Tr}(\hat{\rho}\hat{A}), \end{aligned} \tag{6}$$

where the second equality follows from Eq. (5) and the third equality from the linearity of the trace. This expression for the expectation operator on measurements associated with a mixed state, or a density operator, is also often referred to as *Born's rule*, and it will provide an anchor point for connecting our developments starting in Section 3 with the more traditional approach in quantum mechanics. This is the version of Born's rule in the presence of *epistemic uncertainty* about the state of the system, as captured by the density operator $\hat{\rho}$. The version of Eq. (3) in the presence of a specific state $|\psi\rangle$, can be recovered from it by letting $\hat{\rho} \rightsquigarrow |\psi\rangle\langle\psi|$, as $\operatorname{Tr}(|\psi\rangle\langle\psi|\hat{A}) = \operatorname{Tr}(\langle\psi|\hat{A}|\psi\rangle) = \langle\psi|\hat{A}|\psi\rangle$.[11] So, Eq. (6) combines the *quantum-mechanical uncertainty* about the outcome of a measurement on a system in a pure state, as expressed in Eq. (3), with the *epistemic uncertainty* about the state of the system, captured by the density operator $\hat{\rho}$.

2.3. **A few elementary results about Hermitian operators.** We conclude this introductory section with a number of basic results about Hilbert spaces and Hermitian operators that will come in useful in later sections, and whose formulations (and proofs) we want to separate off in the interest of didactic clarity.

We begin with a number of elementary lemmas. The first lemma provides a necessary and sufficient condition for kets to be normalised.

Lemma 4 ([37, p. 67])
Let $|a_1\rangle, |a_2\rangle, \dots, |a_m\rangle$ be any collection of mutually orthogonal states in a Hilbert space $\mathscr{X}$. Then any $|\psi\rangle \in \mathscr{X}$ is normal if (and only if when $m = n$) there are $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ such that $\sum_{k=1}^{m} |\alpha_k|^2 = 1$ and $|\psi\rangle = \sum_{k=1}^{m} \alpha_k |a_k\rangle$.

Our second lemma shows that it's possible to write any state as an equal-amplitude superposition of some appropriate collection of orthogonal states.

Lemma 5
If $|\psi\rangle$ is a state in an n-dimensional Hilbert space $\mathscr{X}$, then for all $m \in \{1, 2, \dots, n\}$, there are mutually orthogonal states $|\psi_1\rangle$, $|\psi_2\rangle$, …, $|\psi_m\rangle$ in $\bar{\mathscr{X}}$ such that $|\psi\rangle = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |\psi_k\rangle$.

Proof. Let $\{|b_1\rangle, |b_2\rangle, \dots, |b_n\rangle\}$ be any orthonormal basis of $\mathscr{X}$ with $|b_m\rangle := |\psi\rangle$, and let

$$|\psi_s\rangle := \frac{1}{\sqrt{m}} \sum_{k=1}^{m} e^{2\pi i \frac{sk}{m}} |b_k\rangle, \text{ for } s \in \{1, \dots, m\}.$$

[10] We'll have occasion to use this idea a number of times. It's based on the observation that $\operatorname{Tr}(\hat{A}\hat{B}) = \operatorname{Tr}(\hat{B}\hat{A})$; see [37, Sec. 2.1.8] for more details.

[11] Here and further on, we use the symbol '$\rightsquigarrow$' to denote 'instantiates as' or 'is replaced by'.

Then by Lemma 4, $|\psi_s\rangle$ has norm 1 for all $s \in \{1,2,\dots,m\}$. For all $r,s \in \{1,\dots,m\}$, we use the conjugate symmetry of the inner product to find that

$$\begin{aligned}\langle\psi_r|\psi_s\rangle &= \left(\frac{1}{\sqrt{m}}\sum_{k=1}^{m} e^{-2\pi i\frac{rk}{m}}\langle b_k|\right)\left(\frac{1}{\sqrt{m}}\sum_{\ell=1}^{m} e^{2\pi i\frac{s\ell}{m}}|b_\ell\rangle\right) = \frac{1}{m}\sum_{k=1}^{m}\sum_{\ell=1}^{m} e^{-2\pi i\frac{rk}{m}}e^{2\pi i\frac{s\ell}{m}}\langle b_k|b_\ell\rangle \\ &= \frac{1}{m}\sum_{k=1}^{m}\sum_{\ell=1}^{m} e^{-2\pi i\frac{rk}{m}}e^{2\pi i\frac{s\ell}{m}}\delta_{k\ell} = \frac{1}{m}\sum_{k=1}^{m} e^{-2\pi i\frac{rk}{m}}e^{2\pi i\frac{sk}{m}} = \frac{1}{m}\sum_{k=1}^{m} e^{2\pi i\frac{(s-r)k}{m}} = \delta_{rs},\end{aligned}$$

where the second equality is explained by the linearity properties of the inner product, the third equality follows from the orthonormality of the basis $\{|b_1\rangle,|b_2\rangle,\dots,|b_n\rangle\}$ and the last equality identifies the sum of the terms of a finite geometric sequence. Therefore, $|\psi_r\rangle$ and $|\psi_s\rangle$ are orthogonal for $r \neq s$, and since they all have norm 1, they're normal.

Finally, after identifying once again the sum of the terms of a finite geometric sequence, we find that

$$\begin{aligned}\frac{1}{\sqrt{m}}\sum_{s=1}^{m}|\psi_s\rangle &= \frac{1}{m}\sum_{s=1}^{m}\sum_{k=1}^{m} e^{2\pi i\frac{sk}{m}}|b_k\rangle = \frac{1}{m}\sum_{k=1}^{m}\left(\sum_{s=1}^{m} e^{2\pi i\frac{sk}{m}}\right)|b_k\rangle = \frac{1}{m}\sum_{k=1}^{m} m\delta_{km}|b_k\rangle = |b_m\rangle \\ &= |\psi\rangle.\end{aligned}$$ □

The final lemma provides a useful alternative expression for $\langle\psi|\hat{A}|\psi\rangle$.

Lemma 6
Consider any collection $|a_1\rangle,\dots,|a_m\rangle$ of mutually orthogonal eigenstates of a Hermitian operator $\hat{A}$ on a Hilbert space $\mathscr{X}$ with corresponding eigenvalues $\lambda_1,\dots,\lambda_m$, and any linear combination $|\psi\rangle := \sum_{k=1}^{m}\alpha_k|a_k\rangle$ of these eigenstates with $\alpha_1,\dots,\alpha_m \in \mathbb{C}$ and such that $\sum_{k=1}^{m}|\alpha_k|^2 = 1$. Then $|\psi\rangle$ is normal and $\langle\psi|\hat{A}|\psi\rangle = \sum_{k=1}^{m}|\alpha_k|^2\lambda_k$.

Proof. That $|\psi\rangle$ is normal follows from Lemma 4. Observe that

$$\langle\psi| = \sum_{k=1}^{m}\alpha_k^*\langle a_k| \text{ and } \hat{A}|\psi\rangle = \hat{A}\sum_{\ell=1}^{m}\alpha_\ell|a_\ell\rangle = \sum_{\ell=1}^{m}\alpha_\ell\hat{A}|a_\ell\rangle = \sum_{\ell=1}^{m}\alpha_\ell\lambda_\ell|a_\ell\rangle,$$

where the last equality holds because $|a_\ell\rangle$ is an eigenstate of $\hat{A}$ with eigenvalue λ_ℓ; so if we use the linearity properties of the inner product, we find that, indeed,

$$\langle\psi|\hat{A}|\psi\rangle = \sum_{k=1}^{m}\sum_{\ell=1}^{m}\alpha_k^*\alpha_\ell\lambda_\ell\langle a_k|a_\ell\rangle = \sum_{k=1}^{m}\sum_{\ell=1}^{m}\alpha_k^*\alpha_\ell\lambda_\ell\delta_{k\ell} = \sum_{k=1}^{m}|\alpha_k|^2\lambda_k,$$

where the second equality follows from the orthonormality of the $|a_1\rangle,\dots,|a_m\rangle$. □

It will also be useful to remember that the set $\mathscr{H}$ of all Hermitian operators on a finite-dimensional complex Hilbert space $\mathscr{X}$ constitutes a finite-dimensional real linear space: if $\mathscr{X}$ has dimension n, then $\mathscr{H}$ has dimension n^2. The so-called *Frobenius inner product* [37, Exercise 2.39], defined by

$$(\hat{A},\hat{B}) := \mathrm{Tr}(\hat{A}\hat{B}) \text{ for all } \hat{A},\hat{B} \in \mathscr{H}$$

turns $\mathscr{H}$ into a real Hilbert space.[12] The so-called *Frobenius norm* $\|\bullet\|_{\mathrm{F}}$ that's associated with this inner product, is defined by

$$\|\hat{A}\|_{\mathrm{F}} := \sqrt{(\hat{A},\hat{A})} = \sqrt{\mathrm{Tr}(\hat{A}\hat{A})} \text{ for all } \hat{A} \in \mathscr{H}.$$

It turns $\mathscr{H}$ into a normed linear space, with which we can associate a topology of open sets, a topological interior operator $\mathrm{Int}(\bullet)$ and a topological closure operator $\mathrm{Cl}(\bullet)$.

[12] This is because any finite-dimensional linear space with an inner product is a Hilbert space.

3. Our decision-theoretic approach

Let's now describe our approach to deriving a generalisation of Born's rule, in the more general version of Eq. (6), from a number of assumptions, inspired by QM1–QM4, the relevant non-probabilistic postulates of quantum mechanics.

3.1. **The set-up.** First, we consider a subject, whom we'll call You. You are concerned with the value that the state of a quantum system under consideration takes in the set of all its possible values, which in principle, by QM1, is given by the set $\bar{\mathscr{X}}$ of all states, or normalised kets. *You are uncertain about this state*: You may have some knowledge about the actual value that this state takes in the set $\bar{\mathscr{X}}$, but this knowledge may not be sufficient for You to determine this actual value with certainty. To honour a time-tested tradition, we'll denote this possibly unknown state by a *capital letter* $|\Psi\rangle$. On the other hand, lower case letters will be used to denote, amongst other things, the states $|\psi\rangle \in \bar{\mathscr{X}}$ that are the candidate values for $|\Psi\rangle$.

In a second step, we recall that we can get information about the unknown state $|\Psi\rangle \in \bar{\mathscr{X}}$ through performing measurements, described by Hermitian operators $\hat{A} \in \mathscr{H}$, as postulated by QM2. If You were to perform a measurement $\hat{A} \in \mathscr{H}$ on the system, its outcome would be uncertain too, for two reasons: firstly because You are (or may be) uncertain about the state the system is in, and secondly because even in a perfectly known state, the Postulate QM3 typically leaves unspecified which of the eigenvalues $\lambda \in \operatorname{spec}(\hat{A})$ of the Hermitian operator $\hat{A}$ the measurement would actually yield.

We'll consider that after You've performed the measurement $\hat{A}$, what You'll get as a pay-off will be the actual outcome $\lambda \in \operatorname{spec}(\hat{A})$ of the measurement, expressed in units of some linear utility scale — also called *utiles*. Your entertaining beliefs about which value $|\Psi\rangle$ assumes in $\bar{\mathscr{X}}$ might cause You to prefer performing some measurements over others, because You believe they will lead to better pay-offs.

This is a simple idea, but important for what is to come, so let's look at a simple example to better explain where it comes from. It's the first in a series of instalments of a running example that we'll use to illustrate various ideas throughout the text.

Qubit running example: instalment 1
Consider a qubit [37], which is a quantum system with a two-dimensional Hilbert space, spanned by two orthogonal basis states $|0\rangle$ and $|1\rangle$, so

$$\mathscr{X} = \{a|0\rangle + b|1\rangle : a, b \in \mathbb{C}\},$$

and the set of all possible states is then, due to Lemma 4,

$$\bar{\mathscr{X}} = \{a|0\rangle + b|1\rangle : a, b \in \mathbb{C} \text{ and } |a|^2 + |b|^2 = 1\}.$$

We consider the unknown ket $|\Psi\rangle := A|0\rangle + B|1\rangle$, where A and B are uncertain variables in $\mathbb{C}$ such that $|A|^2 + |B|^2 = 1$. For any given $\alpha, \beta \in \mathbb{R}$, we consider the Hermitian operators $\hat{C}_\alpha, \hat{D}_\beta \in \mathscr{H}$ defined by

$$\hat{C}_\alpha := \alpha|0\rangle\langle 0| - |1\rangle\langle 1| \text{ and } \hat{D}_\beta := -|0\rangle\langle 0| + \beta|1\rangle\langle 1|,$$

so $\hat{C}_\alpha$ has eigenstates $|0\rangle$ and $|1\rangle$ with respective eigenvalues α and -1, and $\hat{D}_\beta$ has eigenstates $|0\rangle$ and $|1\rangle$ with respective eigenvalues -1 and β [see Proposition 1].

Now, suppose that You strongly believe that the qubit is in the $|0\rangle$ state, so in other words, that $|B| = 0$ and $|A| = 1$. Because You accept QM4, this implies a strong belief, on Your part, that the outcome of the measurement $\hat{C}_\alpha$ will be α, and correspondingly, that the outcome of the measurement $\hat{D}_\beta$ will be -1. You'll therefore prefer performing the measurement $\hat{C}_\alpha$ to performing the measurement $\hat{D}_\beta$ as long as α is sufficiently larger than -1. ⅁

It's this idea, namely that *Your having beliefs about what value the (possibly) unknown state $|\Psi\rangle$ assumes in the possibility space $\bar{\mathscr{X}}$ may lead You to having preferences between performing different measurements*, on which we'll build our theory. It brings us squarely into a traditional context of decision-making under uncertainty [1, 3, 4, 17, 19, 36, 50, 52], for which we're about to spell out the more important details, in the form of a number of background assumptions and postulates.

3.2. **Decision-theoretic background.** There are quite a number of ways in which so-called rational decision-making can be approached. Many, if not most, of them start with the basic set-up where, on the one hand, there are *acts a* that You're invited to choose — express a preference — between, and where, on the other hand, the so-called *consequence c* of choosing an act a may depend on what is often referred to as the *state of the world* ω, which is considered to be unknown, or uncertain.

In the simplest case, where there are only a finite number n of acts a_k, $k=1,\dots,n$ and a finite number m of possible states of the world ω_ℓ, $\ell=1,\dots,m$, the decision problem can be summarised in an act-state table, where each entry $c_{k\ell}$ is the consequence of choosing act a_k when the actual state of the world turns out to be ω_ℓ:

	ω_1	ω_2	$\cdots$	ω_m
a_1	c_{11}	c_{12}	$\cdots$	c_{1m}
a_2	c_{21}	c_{22}	$\cdots$	c_{2m}
$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
a_n	c_{n1}	c_{n2}	$\cdots$	c_{nm}

If we denote the set of all consequences by C, then the acts a can be seen as, or identified with, maps from the set of all states Ω to C; in the simple example above, we then have that $a_k(\omega_\ell)=c_{k\ell}$. Savage [42] considers specific preference relations $\trianglerighteq$ on the (set of all) acts A, and his approach consists in providing axioms for the sets of acts A and consequences C and for the preference relation $\trianglerighteq$ on A that guarantee that there's some probability on the state space Ω with corresponding expectation operator E and some so-called *utility function* $U\colon C\to\mathbb{R}$ such that

$$a\trianglerighteq b\Leftrightarrow E(U\circ a)\geq E(U\circ b)\text{ for all }a,b\in A. \tag{7}$$

A number of his axioms have the effect of ensuring that the sets of acts and consequences are *sufficiently rich* and closed under convex mixtures. We see that, in other words, probabilities, utilities and expectations on this way of thinking are tools that can be used to conveniently represent Your preferences between acts; and they can be constructed from the acts, consequences and preferences.

One specific aspect of Savage's approach stands out in the light of what we want to come to next. He only allows for a *total*, or *linear*, *ordering* of acts: it must be that $a\trianglerighteq b$ or $b\trianglerighteq a$, for all $a,b\in A$. This allows him to define a *strict preference* $\triangleright$ by letting $a\triangleright b\Leftrightarrow b\not\trianglerighteq a$, for all $a,b\in A$. In other words, if $a\not\triangleright b$ and $b\not\triangleright a$, then it must be that both $b\trianglerighteq a$ and $a\trianglerighteq b$, which means that You have no option but to consider a and b equivalent: there's no room for incomparability or indecision in Savage's set-up.

We find Savage's blanket totality requirement too strong, generally speaking, for a number of reasons. The first is that, as Savage himself indicates [42, Sec. 2.7], there seems to be no *a priori* reason not to allow for incomparability or indecision when trying to model Your rational preferences: indecision can be perfectly rational, even though allowing for it tends to complicate things. A reluctance to complicate matters seems to have been Savage's prime reason for not pursuing this idea of allowing for indecision:

> There is some temptation to explore the possibilities of analysing preference amongst acts as a **partial ordering**, that is, in effect to replace part 1 of the definition of simple ordering by the very weak proposition f $\leq$ f, admitting that some pairs of acts are incomparable. This would seem to give

> expression to introspective sensations of indecision or vacillation, which we may be reluctant to identify with indifference. My own conjecture is that it would prove a blind alley losing much in power and advancing little, if at all, in realism; but only an enthusiastic exploration could shed real light on the question.

The second reason has a more positive flavour: not allowing for incomparability or indecision tends to hide or ignore important *inferential* aspects of specifying preferences. Indeed, there's a very strong analogy with propositional logic, where it's generally accepted to be of crucial importance to allow for sets of propositions that are deductively closed but *not complete* in the sense that they leave the truth states of some propositions undecided; see for instance the detailed discussion in Ref. [18].

It's interesting to mention already here that Wallace [48, 49], as we'll discuss in Section 8, uses Savage's approach to refine Deutsch's argumentation [23] (discussed in Section 7) in a multiverse context; it's therefore not very surprising that his argument leads to a model for the uncertainty in quantum mechanics that describes Your preferences by a probability–utility pair, as alluded to above.

Here, we want to follow a different route, for the reasons already mentioned, and not impose the totality of preference orderings as a foundational part of our set-up. If we do assume totality of preference relations in certain places and instances, we will be clear about it, and mention it explicitly. In all other cases and instances, we'll assume partially ordered preferences to be the fallback models. To see where this different route could lead, we have a brief look at some of the relevant literature.

Anscombe and Aumann [1] and Aumann [3, 4], as well as a few authors in their wake, amongst whom Seidenfeld, Schervish and Kadane [44] and Nau [36], have followed a different approach than Savage to modelling decision-making, by (essentially) considering as the set of consequences C the simplex of all probability mass functions — so-called *lotteries* — on some set of rewards R. Taking the set of acts A to be all maps from the set of states of the world Ω to these lotteries makes sure that A is sufficiently rich: it contains all constant maps — the lotteries — and is closed under convex mixtures. The state-dependent lotteries that constitute the acts in A are also called *horse lotteries*. Two differences with Savage's approach stand out: (i) in most of the above-mentioned papers [3, 4, 36, 44], the preference relations $\trianglerighteq$ and $\triangleright$ needn't reflect totality, and therefore still allow for incomparability between acts; and (ii) in some of them [44], the authors also let go of conditions of Archimedeanity that allow the preferences to be represented by (sets of) real-valued utilities.

Bruno de Finetti [15, 16], on the other hand, followed a simpler — and therefore less general — route, where contrary to the above-mentioned approaches, the utility function U isn't derived from Your preferences, but is instead assumed to exist extraneously to the decision problem. This then essentially implies that the consequences $c \in C$ in the decision problem can now themselves be identified with utilities, that is, assumed to be expressed in units of some predetermined linear utility scale. These could, for instance, be tickets in a lottery for a single desirable prize, and in this sense, this simpler take on preference modelling can also be viewed as a special case of the above-mentioned horse lottery approach, where the doubleton set of rewards consists of winning and not winning the prize; see also Refs. [19, 36, 52]. The acts can now be seen as *uncertain rewards*: they're maps $a\colon \Omega \to \mathbb{R}$ that associate with each state of the world ω the reward $a(\omega)$ for taking action a, expressed in utiles. The preference relations $\trianglerighteq$ and $\triangleright$ now express preferences between such uncertain rewards. We'll see further on in Section 7 that Deutsch's take [23] on framing quantum uncertainty in a decision-theoretic context can be seen as inspired by de Finetti's approach.

On de Finetti's approach, Your preference ordering is still ideally total, as well as Archimedean, but later approaches [44, 50, 51] let go of these requirements, to allow for incomparability or indecision, and to incorporate preferences that aren't necessarily representable by the ordering on the real numbers. Doing so has led to a rich literature on so-called *sets of desirable* (or favourable) *gambles*; see for instance Refs. [10, 11, 21, 35, 38, 39, 52]. It's in this generalisation of de Finetti's approach that we intend to develop our argumentation: on the one hand, because it allows for dealing with partial preferences, and on the other, because it's technically the least complicated. We'll leave the more general but technically more involved approaches for future research.

It follows from the discussion near the beginning of this section that each measurement $\hat{A} \in \mathscr{H}$ can be considered as an 'act'

$$\text{act}_{\hat{A}} := \text{ “perform the measurement } \hat{A} \text{ on the quantum system”},$$

and these acts are in a one-to-one correspondence with the elements of the (real) linear space $\mathscr{H}$ of all measurements. If we consider the states of the world ω to be the elements $|\phi\rangle$ of the state space $\bar{\mathscr{X}}$, all that still needs to be specified to complete the basic description of the decision problem, is the consequences, or in other words, the utilities (expressed in utiles) that correspond to any $\text{act}_{\hat{A}}$ in each state of the world $|\phi\rangle$. This brings us to the first background assumption.

Decision-theoretic background assumption 1 (DTB1)
With every measurement $\hat{A} \in \mathscr{H}$ on the Hilbert space $\mathscr{X}$, You associate a reward function*, which is a map $w_{\hat{A}} \colon \bar{\mathscr{X}} \to \mathbb{R}$, such that the real number $w_{\hat{A}}(|\phi\rangle)$ is the reward, expressed in utiles, for* $\text{act}_{\hat{A}}$ *when the quantum system under consideration system is in state $|\phi\rangle$.*

In other words, the uncertain real number $w_{\hat{A}}(|\Psi\rangle)$ is the uncertain reward associated with $\text{act}_{\hat{A}}$; this reward is typically uncertain because the system state $|\Psi\rangle$ is unknown. The reward functions $w_{\hat{A}}$ constitute a subset

$$\mathscr{W} := \{w_{\hat{A}} : \hat{A} \in \mathscr{H}\}$$

of the linear space $\mathscr{L}(\bar{\mathscr{X}})$ of all real-valued maps on $\bar{\mathscr{X}}$. Since we can associate an uncertain reward $w(|\Psi\rangle)$ with every reward function $w \in \mathscr{W}$, we'll agree to call these reward functions *w uncertain rewards* as well. We stress that, in a general decision-making context, the reward functions are Yours to choose or determine, but we'll argue further on that in the specific context of quantum mechanics, a number of simple postulates make sure that this is no longer the case and that You're left with no choice about what these reward functions look like.

This first background assumption fits nicely within de Finetti's approach, but its consequences are stronger than they might seem at first encounter. For a start, it implies that if You know that the quantum system is in a state $|\phi\rangle$ — so if You know that $|\Psi\rangle = |\phi\rangle$ — then Your preferences between acts/measurements must necessarily be *total* in the sense that You'll strictly prefer measurement $\hat{A}$ to measurement $\hat{B}$ if and only if $w_{\hat{A}}(|\phi\rangle) > w_{\hat{B}}(|\phi\rangle)$, and You're indifferent between $\hat{A}$ and $\hat{B}$ if and only if $w_{\hat{A}}(|\phi\rangle) = w_{\hat{B}}(|\phi\rangle)$. In this sense, the *starting point* for our approach is similar to Deutsch's [23] and Wallace's [48, 49], as we'll see in Sections 7 and 8, respectively: *acts/measurements are linearly ordered when You know what state the system is in.*

Moreover, if You know that $|\Psi\rangle = |\phi\rangle$, then $w_{\hat{A}}(|\phi\rangle)$ is a one-number summary of all the possible outcomes $\lambda \in \text{spec}(\hat{A})$ that the measurement $\hat{A}$ may yield for a system in that state $|\phi\rangle$. DTB1 essentially assumes that such a one-number summary is possible, and implies that it's the only thing that matters for Your decisions when You know the state of the system. What it doesn't do, however, is to specify, or impose restrictions on, what these reward functions may look like. For that, we'll use specific postulates about the reward functions in the next section.

In general, when You don't know with certainty what value $|\Psi\rangle$ assumes, we'll still assume that You can order the acts $\mathrm{act}_{\hat{A}}$, but we'll not require, as Savage would have us do, that this ordering should be total.

Decision-theoretic background assumption 2 (DTB2)
Your beliefs about the value of $|\Psi\rangle$ in $\bar{\mathscr{X}}$ lead You to strictly prefer some uncertain rewards in $\mathscr{W}$ to others, which leads to a strict preference relation $\rhd$ on the set of uncertain rewards $\mathscr{W}$ and therefore also on the set of acts, or measurements, $\mathscr{H}$.

By a *strict preference* relation, we mean an irreflexive and transitive binary relation (a so-called *strict partial order*); and when $\mathscr{W}$ is a vector space, as we'll argue that it is further on, we'll assume that it's also a vector ordering. We repeat that there's no totality assumption: it's *not assumed* that You think incomparable measurements — measurements $\hat{A}$ and $\hat{B}$ for which neither $w_{\hat{A}}(|\Psi\rangle) > w_{\hat{B}}(|\Psi\rangle)$ nor $w_{\hat{A}}(|\Psi\rangle) < w_{\hat{B}}(|\Psi\rangle)$ holds — are necessarily equivalent, so that You're indifferent between them. We'll dive into the many and rich details of such belief representations in Sections 4 to 6 further on.

In summary, the background assumptions DTB1 and DTB2 allow You to represent Your beliefs about where the state $|\Psi\rangle$ is by means of a strict partial (vector) ordering on the reward functions, but the ordering is required to be total — or linear — and in particular determined by linear ordering of the real numbers $w_{\bullet}(|\phi\rangle)$ only when You know with certainty that $|\Psi\rangle = |\phi\rangle$.

3.3. **The reward function postulates.** The assumptions DTB1 and DTB2 fix the context for our argument. Against this backdrop, we'll now formulate four decision-theoretic postulates; more precisely, these are postulates that deal with the reward function aspect of our decision-making framework. As we're about to argue, they're inspired by, and in a sense based on, the non-probabilistic postulates of quantum mechanics. We'll prove in Section 10 that they determine what the uncertain rewards $w_{\hat{A}}$ look like: You don't get to play any role in fixing their shape, the postulates RF1–RF4 we're about to introduce, will do that for You.

Every one of these four postulates is in itself intended to capture a simple idea, and we'll devote some attention to trying to point out what the four relevant ideas are. We intend them to capture what is 'essential' about the decision problem in quantum mechanics to allow us to recover Born's rule as a special case. We emphatically don't want to claim that these central ideas can't be clarified further, or stated more succinctly or elegantly. Nor do we necessarily believe that there are no simpler, weaker or more parsimonious sets of postulates that may lead to the same conclusions. *Our aim here is, simply stated, to propose a collection of postulates that we feel are convincing and natural enough, and which allow us to fix the reward functions.*

Let's begin with the simplest postulate, which fixes the utility gauge of the uncertain rewards. It deals directly with the case governed by QM4, where the system resides in an eigenstate $|a\rangle$ of the measurement operator $\hat{A} \in \mathscr{H}$. Since the outcome of the measurement is then necessarily the corresponding eigenvalue λ, we want this to also be the utility You obtain from performing the measurement in that state: the reward received should equal the outcome of the measurement in case of certainty. This can be seen as a convention — as are essentially all ways of fixing a gauge — and it seems the simplest one that allows us to connect the reward — the utility — received from performing a measurement to its outcome.

Reward function postulate 1 (RF1)
Let $\mathscr{E}_{\lambda}$ be the eigenspace corresponding to an eigenvalue λ of a Hermitian operator $\hat{A}$ on a Hilbert space $\mathscr{X}$. Then necessarily $w_{\hat{A}}(|a\rangle) = \lambda$ for all $|a\rangle \in \mathscr{E}_{\lambda}$ with $\langle a|a\rangle = 1$.

The second postulate is a fairly strong one. We'll first give a general formulation and then spend some effort in justifying it, through a number of examples and through an alternative formulation that perhaps captures its essence more clearly. It's important to realise here that this postulate — as does the next one — identifies the values of reward functions across different state spaces, and therefore allows us to express powerful invariance properties.

Reward function postulate 2 (RF2)
Consider any Hermitian operator $\hat{A} := \sum_{k=1}^{r} \lambda_k \hat{P}_{\mathscr{E}_k}$ on a Hilbert space $\mathscr{X}_1$, with (distinct) real eigenvalues $\lambda_1, \dots, \lambda_r$ corresponding to respective mutually orthogonal eigenspaces $\mathscr{E}_1, \dots, \mathscr{E}_r$ that span $\mathscr{X}_1$. Similarly, consider a Hermitian operator $\hat{B} := \sum_{k=1}^{r} \lambda_k \hat{P}_{\mathscr{F}_k}$ on a Hilbert space $\mathscr{X}_2$, with the same eigenvalues $\lambda_1, \dots, \lambda_r$, corresponding to respective mutually orthogonal eigenspaces $\mathscr{F}_1, \dots, \mathscr{F}_r$ that span $\mathscr{X}_2$. Choose any normalised $|a_k\rangle \in \mathscr{E}_k$ and $|b_k\rangle \in \mathscr{F}_k$, and any $\alpha_k \in \mathbb{C}$ such that $\sum_{k=1}^{r} |\alpha_k|^2 = 1$, and consider the states $|\phi_{\hat{A}}\rangle := \sum_{k=1}^{r} \alpha_k |a_k\rangle \in \mathscr{X}_1$ and $|\phi_{\hat{B}}\rangle := \sum_{k=1}^{r} \alpha_k |b_k\rangle \in \mathscr{X}_2$. Then $w_{\hat{A}}(|\phi_{\hat{A}}\rangle) = w_{\hat{B}}(|\phi_{\hat{B}}\rangle)$.

In essence, this requires that if a state $|\phi\rangle$ is a superposition of eigenstates of a measurement $\hat{A}$ corresponding to distinct eigenvalues, then $w_{\hat{A}}(|\phi\rangle)$ should depend *only* on the superposition weights and on these eigenvalues, but not on the eigenstates themselves, nor on the Hilbert space they're embedded in.

The following examples are meant to illustrate two different applications of this second postulate and will provide intuition about what it entails and why it could be considered reasonable as an invariance requirement.

Qubit running example: instalment 2
Let's go back to our example involving a single qubit and assume that it represents the spin of an electron. Assume that the respective states $|0\rangle$ and $|1\rangle$, which constitute an orthonormal basis for the Hilbert space $\mathscr{X}$, represent spin up and spin down in some physical direction, labelled as the z-direction. We consider a measurement in this physical z-direction that is represented by the Hermitian operator $\hat{A} := |1\rangle\langle 1| - |0\rangle\langle 0| \in \mathscr{H}$; it yields the outcome $+1$ for spin up in this direction and -1 for spin down. The corresponding orthogonal eigenspaces are, respectively, $\mathscr{E}_+ = \operatorname{span}(\{|1\rangle\})$ and $\mathscr{E}_- = \operatorname{span}(\{|0\rangle\})$.

However, we can also use the transformed basis with states $|0'\rangle := |1\rangle$ and $|1'\rangle := |0\rangle$ to model the same system. Consider the operator $\hat{B} := |1'\rangle\langle 1'| - |0'\rangle\langle 0'|$, which is now a measurement in the negative physical z-direction. $\hat{B} = -\hat{A}$ has the same eigenvalues $+1$ and -1 as $\hat{A}$, but with exchanged eigenspaces $\mathscr{F}_+ = \operatorname{span}(\{|1'\rangle\}) = \operatorname{span}(\{|0\rangle\}) = \mathscr{E}_-$ and $\mathscr{F}_- = \operatorname{span}(\{|0'\rangle\}) = \operatorname{span}(\{|1\rangle\}) = \mathscr{E}_+$.

Now consider any $\alpha_1, \alpha_2 \in \mathbb{C}$ such that $|\alpha_1|^2 + |\alpha_2|^2 = 1$, and the states $|\phi_{\hat{A}}\rangle := \alpha_1|1\rangle + \alpha_2|0\rangle$ and $|\phi_{\hat{B}}\rangle := \alpha_1|1'\rangle + \alpha_2|0'\rangle = \alpha_1|0\rangle + \alpha_2|1\rangle$.

Even though the measurements $\hat{A}$ and $\hat{B}$, as well as the corresponding states $|\phi_{\hat{A}}\rangle$ and $|\phi_{\hat{B}}\rangle$ are clearly different, the second situation is a mere *relabelling* of the first, and Postulate RF2 therefore requires the corresponding rewards $w_{\hat{B}}(|\phi_{\hat{B}}\rangle)$ and $w_{\hat{A}}(|\phi_{\hat{A}}\rangle)$ in both situations to be the same. ♪

Qubit running example: instalment 3
Consider an experiment where we conduct a measurement on two independent qubits. The first qubit system is the one we're interested in, while the second qubit system isn't really of any interest to us; for example, this could be one of the spins of some free electron in extragalactic space. We can then consider both systems simultaneously and regard them as one larger system. In quantum mechanics, such a composition of two systems with respective Hilbert spaces $\mathscr{X}$ and $\mathscr{Y}$ is described by the tensor product $\mathscr{X} \otimes \mathscr{Y}$; see for instance Ref. [40, Sec. II.4].

We define a measurement on this composite system as follows. On the first system in a state $|\psi\rangle$, we perform a measurement $\hat{C}$, while leaving the second system untouched in some state $|\phi\rangle$. This leads to the tensor product $\hat{C}\otimes\hat{I}$ of the measurement $\hat{C}$ on the first system with the identity measurement $\hat{I}$ on the second one: $(\hat{C}\otimes\hat{I})(|\psi\rangle\otimes|\phi\rangle) = (\hat{C}|\psi\rangle)\otimes|\phi\rangle$. Postulate RF2 then implies that $w_{\hat{C}\otimes\hat{I}}(|\psi\rangle\otimes|\phi\rangle) = w_{\hat{C}}(|\psi\rangle)$.[13] In more words, the reward $w_{\hat{C}\otimes\hat{I}}(|\psi\rangle\otimes|\phi\rangle)$ for this extended measurement on the larger system in the extended (so-called decoherent) state $|\psi\rangle\otimes|\phi\rangle$ shouldn't differ from the reward $w_{\hat{C}}(|\psi\rangle)$ for the measurement on the first system in the state $|\psi\rangle$; this is justifiable because the second system is independent of the first and the extended measurement leaves that second system untouched. ⸙

The first of the two examples above illustrates that RF2 requires reward functions to be *invariant under relabelling*, and the second example hints at *invariance under* a specific type of *compression* of the Hilbert space. We now give an alternative but equivalent formulation of the postulate that elucidates these two types of invariance requirements more clearly.

Reward function postulate 2 (RF2*, alternative formulation).
Consider any Hermitian operator $\hat{A} := \sum_{k=1}^{r}\lambda_k\hat{P}_{\mathscr{E}_k}$ on a Hilbert space $\mathscr{X}_1$, with (distinct) real eigenvalues $\lambda_1,\dots,\lambda_r$ corresponding to respective mutually orthogonal eigenspaces $\mathscr{E}_1,\dots,\mathscr{E}_r$ that span $\mathscr{X}_1$. Also consider any r-dimensional Hilbert space $\mathscr{X}_2$ with any orthonormal basis $\{|b_1\rangle\dots,|b_r\rangle\}$ and the Hermitian operator $\hat{B} := \sum_{k=1}^{r}\lambda_k|b_k\rangle\langle b_k|$ on $\mathscr{X}_2$ with the same eigenvalues $\lambda_1,\dots,\lambda_r$ as $\hat{A}$. Choose any normalised $|a_k\rangle\in\mathscr{E}_k$ and any $\alpha_k\in\mathbb{C}$ such that $\sum_{k=1}^{r}|\alpha_k|^2=1$, and consider the states $|\phi_{\hat{A}}\rangle := \sum_{k=1}^{r}\alpha_k|a_k\rangle\in\mathscr{X}_1$ and $|\phi_{\hat{B}}\rangle := \sum_{k=1}^{r}\alpha_k|b_k\rangle\in\mathscr{X}_2$. Then $w_{\hat{A}}(|\phi_{\hat{A}}\rangle) = w_{\hat{B}}(|\phi_{\hat{B}}\rangle)$.

Brief argument that RF2* *and* RF2 *are equivalent.* RF2* clearly follows from RF2 as a special case. That it also implies RF2 can be seen by applying it twice in opposite directions. □

This definitely exhibits the compression part of the invariance requirement: each eigenspace can be *compressed* into a one-dimensional space without affecting the reward function. Since, moreover, the choice of orthonormal basis in $\mathscr{X}_2$ has no effect, this shows that *unitary transformations* (and in particular relabelling) have no impact on the reward function either.

The third postulate exploits the linearity of the utility and considers two measurements $\hat{A}$ and $\hat{B}$ with the same orthogonal eigenspaces $\mathscr{E}_k$ but possibly different corresponding eigenvalues λ_k and μ_k. The sum $\hat{A}+\hat{B}$ then also has the same eigenspaces $\mathscr{E}_k$ with corresponding eigenvalues $\lambda_k+\mu_k$. If the state $|\phi\rangle$ is in one of these eigenspaces $\mathscr{E}_k$, then RF1 guarantees that $w_{\hat{A}+\hat{B}}(|\phi\rangle) = \lambda_k+\mu_k = w_{\hat{A}}(|\phi\rangle)+w_{\hat{B}}(|\phi\rangle)$, so the uncertain reward is additive on each of the eigenspaces $\mathscr{E}_k$. We now require that this additivity should be extended from the eigenspaces to the entire Hilbert space $\mathscr{X}$.

Reward function postulate 3 (RF3)
Consider any two Hermitian operators of the form $\hat{A} := \sum_{k=1}^{r}\lambda_k\hat{P}_{\mathscr{E}_k}$ and $\hat{B} := \sum_{k=1}^{r}\mu_k\hat{P}_{\mathscr{E}_k}$ on a Hilbert space $\mathscr{X}$, where the r eigenspaces $\mathscr{E}_k$ are mutually orthogonal and span $\mathscr{X}$, where $\lambda_k,\mu_k\in\mathbb{R}$ and then all the λ_k are distinct and all the μ_k are distinct. Then $w_{\hat{A}+\hat{B}}(|\phi\rangle) = w_{\hat{A}}(|\phi\rangle)+w_{\hat{B}}(|\phi\rangle)$ for all $|\phi\rangle\in\bar{\mathscr{X}}$.

The fourth and final postulate deals with the continuity of the uncertain rewards, the underlying idea being that if You're no longer able to distinguish between states, You shouldn't be able to distinguish between the corresponding rewards either: the reward functions $w_{\hat{A}}$ should be continuous in their state argument.

[13] Apply the postulate with $\mathscr{X}_1 \rightsquigarrow \mathscr{X}$, $\mathscr{X}_2 \rightsquigarrow \mathscr{X}\otimes\mathscr{Y}$, $\hat{A}\rightsquigarrow\hat{C}$, $\hat{B}\rightsquigarrow\hat{C}\otimes\hat{I}$, $\mathscr{F}_k\rightsquigarrow\mathscr{E}_k\otimes\mathscr{Y}$, $|b_k\rangle\rightsquigarrow|a_k\rangle\otimes|\phi\rangle$, $|\phi_{\hat{A}}\rangle\rightsquigarrow|\psi\rangle$ and $|\phi_{\hat{B}}\rangle\rightsquigarrow|\psi\rangle\otimes|\phi\rangle$.

Reward function postulate 4 (RF4)
Let $\hat{A}$ be a Hermitian operator on a Hilbert space $\mathscr{X}$ and let $|\phi_n\rangle$ be any sequence of states, then $|\phi\rangle = \lim_{n\to+\infty} |\phi_n\rangle$ implies that $w_{\hat{A}}(|\phi\rangle) = \lim_{n\to+\infty} w_{\hat{A}}(|\phi_n\rangle)$.

Observe, by the way, that the continuity of the norm $\|\bullet\|$ in the topology it induces on the Hilbert space $\mathscr{X}$, implies that the limit $|\phi\rangle = \lim_{n\to+\infty} |\phi_n\rangle$ of the sequence of states $|\phi_n\rangle$ is a state as well, so it makes sense to consider the value $w_{\hat{A}}(|\phi\rangle)$ of the reward function $w_{\hat{A}}$ in that limit state $|\phi\rangle$.

We now come to our main result. To formulate it, we introduce the *specific* so-called *reward assignation* $u_\bullet\colon \mathscr{H} \to \mathscr{U} : \hat{A} \mapsto u_{\hat{A}}$, with

$$u_{\hat{A}}(|\phi\rangle) := \langle\phi|\hat{A}|\phi\rangle \text{ for all } |\phi\rangle \in \bar{\mathscr{X}}$$

and with corresponding *set of uncertain rewards*

$$\mathscr{U} := \{u_{\hat{A}} \colon \hat{A} \in \mathscr{H}\} = \{\langle\bullet|\hat{A}|\bullet\rangle \colon \hat{A} \in \mathscr{H}\}.$$

We can and will prove that the postulates RF1–RF4 determine the reward functions $w_{\hat{A}}$ unequivocally, in the sense that they imply that

$$w_{\hat{A}}(|\phi\rangle) = u_{\hat{A}}(|\phi\rangle) = \langle\phi|\hat{A}|\phi\rangle \text{ for all } \hat{A} \in \mathscr{H} \text{ and all } |\phi\rangle \in \bar{\mathscr{X}}. \tag{8}$$

We postpone the detailed and quite formal mathematical argumentation for this interesting result until Section 10, so we can now, in the intervening sections, concentrate on a discussion of its implications.

4. THE BASIC DECISION-THEORETIC MODELS

Let's look at a specific decision problem involving the unknown state $|\Psi\rangle$ of a quantum system with Hilbert space $\mathscr{X}$, with a corresponding set of measurement operators $\mathscr{H}$. In one of its more general forms, decision theory will now use a *strict (partial) vector ordering* to express Your preferences between acts, and therefore, indirectly, Your beliefs about $|\Psi\rangle$.

4.1. Mathematical preliminaries.

Before delving into the details of this ordering, let's take a few moments to contemplate what it is that's being ordered, namely the acts $\text{act}_{\hat{A}}$, which can be identified with the measurements $\hat{A} \in \mathscr{H}$. Our next result is based on the discussion in the previous section, and in particular on Eq. (8), and shows that the acts $\text{act}_{\hat{A}}$, and the measurements $\hat{A}$, are also in a one-to-one correspondence with the uncertain rewards $u_{\hat{A}}$.

Proposition 7
The reward assignation $u_\bullet$ is a linear isomorphism between the real linear spaces $\mathscr{H}$ and $\mathscr{U}$.

The proof is straightforward, but we include it for the sake of completeness.

Proof. Since, clearly, $\mathscr{H}$ is a real linear space, it suffices that prove that (i) the reward assignation $u_\bullet$ is linear; and (ii) that it's a bijection.

For (i), consider any $\hat{A}, \hat{B} \in \mathscr{H}$, any $\alpha, \beta \in \mathbb{R}$ and any $|\psi\rangle \in \bar{\mathscr{X}}$, then $(\alpha\hat{A} + \beta\hat{B})|\psi\rangle = \alpha\hat{A}|\psi\rangle + \beta\hat{B}|\psi\rangle$, so the bi-linearity of the inner product then guarantees that

$$\begin{aligned} u_{\alpha\hat{A}+\beta\hat{B}}(|\psi\rangle) &= \langle\psi|(\alpha\hat{A}+\beta\hat{B})|\psi\rangle = \langle\psi|(\alpha\hat{A}|\psi\rangle + \beta\hat{B}|\psi\rangle) = \alpha\langle\psi|\hat{A}|\psi\rangle + \beta\langle\psi|\hat{B}|\psi\rangle \\ &= \alpha u_{\hat{A}}(|\psi\rangle) + \beta u_{\hat{B}}(|\psi\rangle), \end{aligned}$$

and therefore, indeed, $u_{\alpha\hat{A}+\beta\hat{B}} = \alpha u_{\hat{A}} + \beta u_{\hat{B}}$, so the reward assignation $u_\bullet$ is linear.

For (ii), it suffices to prove that the reward assignation $u_\bullet$ is one-to-one, as it's clearly onto by the definition of $\mathscr{U}$. So, consider any $\hat{A}, \hat{B} \in \mathscr{H}$ and assume that $u_{\hat{A}} = u_{\hat{B}}$, then we must show that $\hat{A} = \hat{B}$. If we let $\hat{C} := \hat{A} - \hat{B}$, then we infer from the linearity of $u_\bullet$ that $u_{\hat{C}} = u_{\hat{A}} - u_{\hat{B}} = 0$, and we must prove that $\hat{C} = \hat{0}$. Since $\hat{C}$ is Hermitian, it has an orthogonal collection of eigenstates $\{|a_1\rangle,\dots,|a_n\rangle\}$ that constitutes a basis for $\mathscr{X}$. We'll denote by λ_k the eigenvalue of $\hat{C}$ that corresponds to the eigenstate $|a_k\rangle$, for $k\in\{1,\dots,n\}$. Consider any complex numbers $\alpha_1,\dots,\alpha_n$ with $\sum_{k=1}^n|\alpha_k|^2=1$, and let $|\psi\rangle:=\sum_{k=1}^n\alpha_k|a_k\rangle$. Then it follows from the assumption and Lemma 6 that $0=u_{\hat{C}}(|\psi\rangle)=\sum_{k=1}^n|\alpha_k|^2\lambda_k$, and since this must hold for all possible choices of the $\alpha_1,\dots,\alpha_n$, we infer that, necessarily, $\lambda_1=\dots=\lambda_n=0$. But then, indeed, by Proposition 1, $\hat{C}=\sum_{k=1}^n\lambda_k|a_k\rangle\langle a_k|=\hat{0}$. □

For this reason, we'll identify the measurements $\hat{A}$ and the corresponding uncertain rewards $u_{\hat{A}}=\langle\bullet|\hat{A}|\bullet\rangle$, and the real linear spaces $\mathscr{H}$ and $\mathscr{U}$. There are two particular aspects of this identification that deserve extra attention.

First, for any real number μ, the Hermitian operator $\mu\hat{I}$ satisfies $(\mu\hat{I})|\phi\rangle=\mu|\phi\rangle$ for all $|\phi\rangle\in\mathscr{X}$ and therefore has a single eigenvalue μ with corresponding eigenspace $\mathscr{X}$. Postulate QM4 — or in this case equivalently QM3 — then guarantees that the corresponding measurement always produces the outcome μ with certainty. The reward assignation $u_\bullet$ takes these constant measurements $\mu\hat{I}$ to the (constant) maps $u_{\mu\hat{I}}=\langle\bullet|\mu\hat{I}|\bullet\rangle=\mu\langle\bullet|\hat{I}|\bullet\rangle=\mu$. We'll identify in our notations the real number μ and the constant map that assumes the value μ.

Second, there are a few (vector) orderings of the real linear space $\mathscr{U}$ that have a natural interpretation and will play an important role in what follows. We begin with the so-called *weak Pareto* or *weak dominance* ordering $\geq$, which is the partial vector ordering on $\mathscr{U}$ defined by

$$u\geq v\Leftrightarrow(\forall|\psi\rangle\in\bar{\mathscr{X}})u(|\psi\rangle)\geq v(|\psi\rangle),\text{ for all }u,v\in\mathscr{U},$$

and the corresponding strict vector ordering $\gneq$, also called *weak strict dominance*, given by

$$\begin{aligned}u\gneq v&\Leftrightarrow u\geq v\text{ and }u\neq v\\&\Leftrightarrow(\forall|\psi\rangle\in\bar{\mathscr{X}})u(|\psi\rangle)\geq v(|\psi\rangle)\text{ and }(\exists|\psi\rangle\in\bar{\mathscr{X}})u(|\psi\rangle)>v(|\psi\rangle),\text{ for all }u,v\in\mathscr{U}.\end{aligned}$$

There's also the *strong dominance* ordering $>$, which is the strict vector ordering on $\mathscr{U}$ defined by

$$u>v\Leftrightarrow(\forall|\psi\rangle\in\bar{\mathscr{X}})u(|\psi\rangle)>v(|\psi\rangle),\text{ for all }u,v\in\mathscr{U}.$$

The (inverse of the) linear isomorphism $u_\bullet$ induces corresponding vector orderings on the Hermitian operators, completely characterised as follows:

$$\left.\begin{aligned}&\hat{A}\geq\hat{0}\Leftrightarrow u_{\hat{A}}\geq u_{\hat{0}}\Leftrightarrow(\forall|\psi\rangle\in\bar{\mathscr{X}})u_{\hat{A}}(|\psi\rangle)\geq0\Leftrightarrow\min\operatorname{spec}(\hat{A})\geq0\\&\hat{A}\gneq\hat{0}\Leftrightarrow\hat{A}\geq\hat{0}\text{ and }\hat{A}\neq\hat{0}\\&\hat{A}>\hat{0}\Leftrightarrow u_{\hat{A}}>u_{\hat{0}}\Leftrightarrow(\forall|\psi\rangle\in\bar{\mathscr{X}})u_{\hat{A}}(|\psi\rangle)>0\Leftrightarrow\min\operatorname{spec}(\hat{A})>0\end{aligned}\right\}\text{ for all }\hat{A}\in\mathscr{H}.$$

These orderings refer to well-known notions for Hermitian operators, some of which we had occasion to mention in our discussion of density operators in Section 2.2. For instance, $\hat{A}\geq\hat{0}$ means that $\hat{A}$ is *positive semidefinite*, $\hat{A}\gneq\hat{0}$ means that $\hat{A}$ is *positive semidefinite and non-zero* and $\hat{A}>\hat{0}$ means that $\hat{A}$ is *positive definite*.

On our way of looking at things, $\hat{A}\geq\hat{B}$, or equivalently $\hat{A}-\hat{B}\geq\hat{0}$, means that $\mathrm{act}_{\hat{A}}$ always produces an uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ that's at least as high as the uncertain reward $u_{\hat{B}}(|\Psi\rangle)$ produced by $\mathrm{act}_{\hat{B}}$, regardless of what value $|\Psi\rangle$ assumes in $\bar{\mathscr{X}}$. That $\hat{A}\gneq\hat{B}$, or equivalently $\hat{A}-\hat{B}\gneq\hat{0}$, means that, in addition, the corresponding rewards aren't equal for all values that $|\Psi\rangle$ may assume in $\bar{\mathscr{X}}$. We'll denote by $\mathscr{H}_{\gneq\hat{0}}$ the set of all non-null measurements with a non-negative uncertain reward, or in other words, all non-null positive semidefinite Hermitian operators. Also, $\mathscr{H}_{\lneq\hat{0}}:=-\mathscr{H}_{\gneq\hat{0}}$ is the set of all non-null negative semidefinite Hermitian operators.

Similarly, $\hat{A}>\hat{B}$, or equivalently $\hat{A}-\hat{B}>\hat{0}$, means that $\mathrm{act}_{\hat{A}}$ always produces an uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ that's strictly higher than the uncertain reward $u_{\hat{B}}(|\Psi\rangle)$ produced by $\mathrm{act}_{\hat{B}}$, regardless of what value $|\Psi\rangle$ assumes in $\bar{\mathscr{X}}$. We'll denote by $\mathscr{H}_{>\hat{0}}$ the set of all measurements with a positive uncertain reward, or in other words, all positive definite Hermitian operators. Also, $\mathscr{H}_{<\hat{0}} \coloneqq -\mathscr{H}_{>\hat{0}}$ is the set of all negative definite Hermitian operators.

Qubit running example: instalment 4
To illustrate these ideas, we go back to the example of a qubit, and we refer to Ref. [37, Sec. 2.1.3] for the mathematical background details.

All Hermitian operators on its Hilbert space $\mathscr{X}$ can be written as linear combinations of the identity operator $\hat{I}$ and the Pauli operators $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $\hat{\sigma}_z$. In other words, for any $\hat{A} \in \mathscr{H}$, there are scalars $w,x,y,z \in \mathbb{R}$ such that $\hat{A} = w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z$.

Moreover, for any $\alpha, \beta \in \mathbb{C}$, it holds that

$$\hat{\sigma}_x\big(\alpha|0\rangle + \beta|1\rangle\big) = \beta|0\rangle + \alpha|1\rangle \text{ and } \hat{\sigma}_y\big(\alpha|0\rangle + \beta|1\rangle\big) = -i\beta|0\rangle + i\alpha|1\rangle \text{ and } \hat{\sigma}_z\big(\alpha|0\rangle + \beta|1\rangle\big) = \alpha|0\rangle - \beta|1\rangle,$$

and therefore

$$u_{\hat{\sigma}_x}\big(\alpha|0\rangle + \beta|1\rangle\big) = \alpha\beta^* + \alpha^*\beta \text{ and } u_{\hat{\sigma}_y}\big(\alpha|0\rangle + \beta|1\rangle\big) = i(\alpha\beta^* - \alpha^*\beta) \text{ and } u_{\hat{\sigma}_z}\big(\alpha|0\rangle + \beta|1\rangle\big) = \alpha\alpha^* - \beta\beta^*,$$

so

$$u_{\hat{A}}\big((\alpha|0\rangle + \beta|1\rangle\big) = \alpha\alpha^*(w+z) + \beta\beta^*(w-z) + \alpha\beta^*(x+iy) + \alpha^*\beta(x-iy).$$

We also find after some algebraic manipulations that the eigenvalues λ of $\hat{A}$ are given by

$$\lambda = w \pm \sqrt{x^2+y^2+z^2}, \tag{9}$$

and therefore

$$\mathscr{H}_{\geq\hat{0}} = \{w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z \colon w,x,y,z \in \mathbb{R} \text{ and } \sqrt{x^2+y^2+z^2} \leq w\} \tag{10}$$

$$\mathscr{H}_{\gneq\hat{0}} = \{w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z \colon w,x,y,z \in \mathbb{R} \text{ and } \sqrt{x^2+y^2+z^2} \leq w \neq 0\} \tag{11}$$

$$\mathscr{H}_{>\hat{0}} = \{w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z \colon w,x,y,z \in \mathbb{R} \text{ and } \sqrt{x^2+y^2+z^2} < w\}.$$

We close this mathematical digression with a very brief foray into topology. The so-called *supremum* norm of a map $g \colon \bar{\mathscr{X}} \to \mathbb{R}$ is defined by $\|g\|_\infty \coloneqq \sup_{|\phi\rangle\in\bar{\mathscr{X}}} |g(|\phi\rangle)|$. It turns $\mathscr{U}$, and therefore indirectly also $\mathscr{H}$ via the (inverse of the) linear isomorphism $u_\bullet$, into a normed linear space; simply observe that all the elements $u_{\hat{A}}$ of $\mathscr{U}$ are *bounded* — have bounded supremum norm $\|u_{\hat{A}}\|_\infty$. Indeed, for any $\hat{A} \in \mathscr{H}$, we know from Proposition 1 that there's a basis of eigenvectors $|a_1\rangle, \dots, |a_n\rangle$ of $\hat{A}$ for the Hilbert space $\mathscr{X}$, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$, so taking into account Lemmas 4 and 6,

$$\begin{aligned}\|\hat{A}\|_\infty \coloneqq \|u_{\hat{A}}\|_\infty = \sup_{|\phi\rangle\in\bar{\mathscr{X}}} |u_{\hat{A}}(|\phi\rangle)| = \sup\left\{\left|\sum_{k=1}^n |\alpha_k|^2\lambda_k\right| \colon \sum_{k=1}^n |\alpha_k|^2 = 1\right\} \\ = \max\{|\lambda| \colon \lambda \in \operatorname{spec}(\hat{A})\}.\end{aligned} \tag{12}$$

Observe, by the way, that for the standard definition of the so-called *operator norm* [43, Ch. 23] on the normed linear space $\mathscr{X}$, there's a related result: taking into account that $\hat{A}^\dagger\hat{A} = \hat{A}^2$ is also Hermitian with the same eigenstates as $\hat{A}$, and with eigenvalues that are the squares of the corresponding eigenvalues of $\hat{A}$, we find that

$$\begin{aligned}\|\hat{A}\|_{\mathrm{op}} \coloneqq \sup_{|\phi\rangle\in\mathscr{X}\setminus\{0\}} \frac{\|\hat{A}|\phi\rangle\|}{\||\phi\rangle\|} = \sup_{|\phi\rangle\in\mathscr{X}\setminus\{0\}} \frac{\sqrt{\langle\phi|\hat{A}^\dagger\hat{A}|\phi\rangle}}{\sqrt{\langle\phi|\phi\rangle}} = \sup_{|\phi\rangle\in\bar{\mathscr{X}}} \sqrt{\langle\phi|\hat{A}^\dagger\hat{A}|\phi\rangle} = \|\hat{A}^\dagger\hat{A}\|_\infty \\ = \max\{|\lambda| \colon \lambda \in \operatorname{spec}(\hat{A}^\dagger\hat{A})\} = \max\{\lambda^2 \colon \lambda \in \operatorname{spec}(\hat{A})\}.\end{aligned} \tag{13}$$

Since the real linear space $\mathscr{H}$ is finite-dimensional, all norms are equivalent — lead to the same topology. In Section 2.3 we came across the Frobenius norm $\|\bullet\|_{\mathrm{F}}$ on $\mathscr{H}$, which therefore gives rise to the same topology of open sets, the same topological interior operator $\mathrm{Int}(\bullet)$ and the same topological closure operator $\mathrm{Cl}(\bullet)$, as the supremum norm $\|\bullet\|_\infty$ and the operator norm $\|\bullet\|_{\mathrm{op}}$ do.

Observe, by the way, that a similar argumentation, again based on Lemmas 4 and 6, allows us to infer that

$$\begin{aligned}\inf u_{\hat{A}} &:= \inf_{|\phi\rangle\in\bar{\mathscr{X}}} u_{\hat{A}}(|\phi\rangle) = \inf\left\{\sum_{k=1}^{n}|\alpha_k|^2\lambda_k\colon \sum_{k=1}^{n}|\alpha_k|^2=1\right\}\\ &= \min\{\lambda\colon \lambda\in\operatorname{spec}(\hat{A})\} = \min\operatorname{spec}(\hat{A}).\end{aligned}\tag{14}$$

and similarly that $\sup u_{\hat{A}} = \max\operatorname{spec}(\hat{A})$.

4.2. **Sets of desirable measurements.** We'll take $u_{\hat{A}} \rhd u_{\hat{B}}$ to mean that, based on Your beliefs about the value of $|\Psi\rangle$ in $\bar{\mathscr{X}}$, You strictly prefer[14] the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ corresponding to measurement $\hat{A}$ to the uncertain reward $u_{\hat{B}}(|\Psi\rangle)$ corresponding to measurement $\hat{B}$. Thus, Your beliefs lead You to a (strict) preference relation $\rhd$ on the set $\mathscr{U}$: it collects those couples $(u_{\hat{A}}, u_{\hat{B}})$ for which You strictly prefer $u_{\hat{A}}(|\Psi\rangle)$ to $u_{\hat{B}}(|\Psi\rangle)$. *We'll say that Your beliefs are represented by the preference ordering* $\rhd$.

As we'll see further on in Section 4.3, we'll want to allow for the possibility that the ordering $\rhd$ is only a partial representation (also called an *assessment*) of Your beliefs: You may not have all the time and resources needed to give an account of all Your preferences between the (infinitely many) uncertain rewards in $\mathscr{U}$. This is one practical reason why we don't require the ordering $\rhd$ to be total. Another, more fundamental, reason is that You may not have at Your disposal all the information that would lead You to impose a total ordering on the uncertain rewards in $\mathscr{U}$.

But we do require the ordering $\rhd$ to take into account the rational implications of those preferences that You do express, and to also satisfy certain non-inferential rationality requirements, such as respecting those *background preferences* that should always be present, regardless of any information or beliefs You might have about the value of $|\Psi\rangle$ in $\bar{\mathscr{X}}$. These requirements, both inferential and non-inferential, are captured by the notion of *coherence*, as expressed by the axioms PO1 to PO5, or equivalently D1 to D4, further on.

Because we've argued above that the linear isomorphism $u_\bullet$ allows us to move freely from measurements $\hat{A}$ to uncertain rewards $u_{\hat{A}}$ and backwards, we see that we can readily interpret a preference $u_{\hat{A}} \rhd u_{\hat{B}}$ as a preference between measurements (or between the corresponding acts): $\hat{A} \rhd \hat{B} \Leftrightarrow u_{\hat{A}} \rhd u_{\hat{B}}$.

Once the step of focusing on such preference relations is taken, there's not much else we can do but apply the existing theory for choosing between uncertain rewards — see, for instance Refs. [2, 12, 17, 39, 44, 46, 50] and the discussion in Section 3.2 — and translating everything back to preferences between measurements. In doing so, we'll mainly follow the lead taken by Benavoli, Facchini and Zaffalon in their earlier work [6, 7], but we'll also be adding a few interesting details as we go along. This is our programme for the remainder of this section and Sections 5 and 6.

The notion of coherence captures the *minimal* rationality requirements that we'll want Your preferences to satisfy. We call a binary preference ordering $\rhd$ on the space of measurements $\mathscr{H}$ *coherent* if it satisfies the following conditions:

PO1. $\hat{A} \not\rhd \hat{A}$ for all $\hat{A}\in\mathscr{H}$; [irreflexivity]
PO2. $\hat{A} \rhd \hat{B}$ and $\hat{B} \rhd \hat{C} \Rightarrow \hat{A} \rhd \hat{C}$ for all $\hat{A},\hat{B},\hat{C}\in\mathscr{H}$; [transitivity]
PO3. $\hat{A} \rhd \hat{B} \Rightarrow c\hat{A} \rhd c\hat{B}$ for all $\hat{A},\hat{B}\in\mathscr{H}$ and all $c\in\mathbb{R}_{>0}$; [positive scaling]
PO4. $\hat{A} \gneq \hat{B} \Rightarrow \hat{A} \rhd \hat{B}$ for all $\hat{A},\hat{B}\in\mathscr{H}$; [monotonicity]

[14] Such *strict preference* can be given the following operationalisable meaning: You strictly prefer $u_{\hat{A}}(|\Psi\rangle)$ to $u_{\hat{B}}(|\Psi\rangle)$ if You accept the uncertain reward $u_{\hat{A}}(|\Psi\rangle) - u_{\hat{B}}(|\Psi\rangle)$ but don't want to give it away; see for instance Ref. [39] for a thorough discussion of such preferences, also leading to a justification for the axioms PO1 to PO5 and D1 to D4.

PO5. $\hat{A} \rhd \hat{B} \Rightarrow (\hat{A}+\hat{C}) \rhd (\hat{B}+\hat{C})$ for all $\hat{A},\hat{B},\hat{C} \in \mathscr{H}$. [additivity]

What lies behind these coherence requirements?

Axioms PO1 and PO2 reflect the strict partial order aspect of the preference. We don't require the ordering to be *total*, by the way: $\hat{A} \ntriangleright \hat{B}$ *and* $\hat{B} \ntriangleright \hat{A}$ needn't imply that $\hat{A}=\hat{B}$; it may be that You're *indifferent* between the different uncertain rewards $u_{\hat{A}}(|\Psi\rangle)$ and $u_{\hat{B}}(|\Psi\rangle)$ in the sense that You're willing to exchange any one of the two for the other, but alternatively *also* that to You, $\hat{A}$ and $\hat{B}$ are *incomparable* in the sense that You don't feel able or compelled to express a (weak or strict) preference between the uncertain rewards $u_{\hat{A}}(|\Psi\rangle)$ and $u_{\hat{B}}(|\Psi\rangle)$.

Axioms PO3 and PO5 turn the preference ordering into a *vector* ordering that's compatible with, or preserved under, the addition and scalar multiplication of measurements. This is a reflection of the linearity of the utility scale that rewards are expressed in, and which we assumed from the outset in Section 3.

And, finally, Axiom PO4 expresses that Your preferences should take into account the natural *background ordering* $\gneq$ between measurements. If $\hat{A} \gneq \hat{B}$, or equivalently, $u_{\hat{A}} \gneq u_{\hat{B}}$ then, regardless of Your beliefs about the unknown state $|\Psi\rangle$, You ought to always strictly prefer $\hat{A}$ over $\hat{B}$, as the uncertain reward $u_{\hat{A}}$ can never be lower than the uncertain reward $u_{\hat{B}}$ in any given state, while for some states it will be strictly higher.[15]

Because a coherent $\rhd$ is a strict vector ordering, we can represent it in a mathematically equivalent manner by the following convex cone of measurements:

$$\mathscr{D} := \{\hat{A} \in \mathscr{H} : \hat{A} \rhd \hat{0}\}.$$

Its elements are the measurements that You (strictly) prefer to the zero measurement $\hat{0}$, which gives You zero utility and is therefore often called the *status quo*. We'll therefore call such measurements *desirable* (to You), and the corresponding convex cone $\mathscr{D}$ a *set of desirable measurements* (for You). Observe that, by PO5,

$$\hat{A} \rhd \hat{B} \Leftrightarrow \hat{A}-\hat{B} \rhd \hat{0} \Leftrightarrow \hat{A}-\hat{B} \in \mathscr{D}, \text{ for any two measurements } \hat{A},\hat{B} \in \mathscr{H}, \quad (15)$$

which confirms that the ordering $\rhd$ is indeed completely determined by the convex cone $\mathscr{D}$, and vice versa.

Such sets of desirable measurements $\mathscr{D}$ therefore constitute an equivalent representation to preference orderings $\rhd$. The basic rationality requirements that are typically imposed on them are:

D1. $\hat{0} \notin \mathscr{D}$;
D2. if $\hat{A},\hat{B} \in \mathscr{D}$ then also $\hat{A}+\hat{B} \in \mathscr{D}$;
D3. if $\hat{A} \in \mathscr{D}$ and $\lambda > 0$ then also $\lambda\hat{A} \in \mathscr{D}$;
D4. if $\hat{A} \gneq \hat{0}$ then $\hat{A} \in \mathscr{D}$.

They're the one-to-one counterparts, in that order and via Eq. (15) and therefore PO5, for the axioms PO1 to PO4. Any subset $\mathscr{D}$ of $\mathscr{H}$ with these four properties is called a *coherent* set of desirable measurements (in $\mathscr{H}$). Axiom D1 essentially states that Your ordering $\rhd$ must be strict, and together with D2 guarantees, via Eq. (15), that You have a strict preference ordering between acts. Axioms D2 and D3 state that $\mathscr{D}$ must be a convex cone, so they — in combination with Eq. (15) and therefore PO5 — also reflect the linearity of the utility scale. The final axiom D4 simply requires that non-null measurements that never yield a negative outcome — that are positive semidefinite — must be desirable to You.[16]

[15] Most, if not all, of what we'll discuss in Sections 4.2, 4.3, 5 and 6 remains, mutatis mutandis, valid if we consider $>$ rather than $\gneq$ as the background ordering and add the monotonicity requirement that $\hat{A} \rhd \hat{B}$ and $\hat{B} \geq \hat{C}$ should imply that also $\hat{A} \rhd \hat{C}$, or equivalently, that $\hat{A} \rhd \hat{0}$ and $\hat{B} \geq \hat{A}$ should imply that also $\hat{B} \rhd \hat{0}$; this connects directly to the remark in footnote 16.

[16] For some applications, it may be preferable to replace D4 with the weaker condition that $\mathscr{D}$ should include all positive definite measurements, and then also to explicitly add the requirement that $\hat{A} \in \mathscr{D}$ and $\hat{B} \geq \hat{A}$ imply that $\hat{B} \in \mathscr{D}$; see D6 further on. Doing this wouldn't affect the main conclusions in this paper. Again, most of what

It follows readily from the coherence axioms D1, D2 and D4 that a coherent set of desirable measurements $\mathscr{D}$ also has the following properties, where we denote by $\mathscr{H}_{\leq\hat{0}} := -\mathscr{H}_{\geq\hat{0}}$ the set of all negative semidefinite measurements:

D5. $\mathscr{D} \cap \mathscr{H}_{\leq\hat{0}} = \emptyset$;
D6. if $\hat{A} \in \mathscr{D}$ and $\hat{B} \geq \hat{A}$ then also $\hat{B} \in \mathscr{D}$.

We've been talking about desirable *measurements*, but we want to remind readers at this point that a measurement $\hat{A}$ is considered to be desirable because You strictly prefer the associated uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ to the status quo 0. The desirability of measurements always goes back to the desirability of the associated uncertain rewards and thus to the desirability of *gambles* — bounded real-valued maps. The coherence axioms D1 to D4 are direct translations of coherence axioms for the desirability of gambles that appear in the relevant imprecise probabilities literature [7, 19–21, 38, 39, 51], and in that sense the desirable measurements framework we're about to explore below, can be seen as a direct application of imprecise probabilities ideas to quantum theory.

It's important to stress here that there's *nothing essentially probabilistic* about this representation of Your beliefs. To give a simple example, the so-called *vacuous* set of desirable measurements $\mathscr{D}_{\text{vac}} := \{\hat{A} \in \mathscr{H} : \hat{A} \gneq \hat{0}\} = \mathscr{H}_{\gneq\hat{0}}$ is coherent and therefore allowable as a candidate model for Your preferences. It represents Your complete ignorance about $|\Psi\rangle$ in the sense that it reflects Your (quite conservative) lack of any inclination to engage in any act$_{\hat{A}}$ whose reward function $u_{\hat{A}}$ has a negative value in some states. Recall that, for the corresponding *vacuous strict preference ordering* $\gneq$,

$$\hat{A} \gneq \hat{B} \Leftrightarrow \hat{A} - \hat{B} \gneq \hat{0} \Leftrightarrow \hat{A} \neq \hat{B} \text{ and } \min\operatorname{spec}(\hat{A} - \hat{B}) \geq 0, \text{ for all } \hat{A}, \hat{B} \in \mathscr{H}.$$

How this condition is expressed in terms of the eigenvalues of *different* $\hat{A}$ and $\hat{B}$ depends on how 'commensurate' $\hat{A}$ and $\hat{B}$ are: the simplest condition arises when $\hat{A}$ and $\hat{B}$ commute, as then all eigenvalues of $\hat{A}$ must be at least as large as the *corresponding* eigenvalues of $\hat{B}$ (with at least one eigenvalue being strictly larger).

Of course, many other, and arguably more interesting, types of preference orders are also allowed by this set of desirable measurements approach, and we'll have occasion to study quite a few of them in the following sections. But before doing so, we want to briefly draw attention to an interesting aspect of our allowing Your preference ordering to be partial and not insisting on totality, as is often done [17, 23, 42, 47–49]. This is the type of approach advocated in the so-called imprecise probabilities literature [2, 12, 46, 50, 52], which allows the *inferential aspects* of uncertain reasoning to come to the fore, as already hinted at in the Introduction.

4.3. **Conservative inference.** When we want to model Your beliefs about the system under consideration, it may not be feasible to ask You to come up with all the measurements that are desirable to You — or all Your strict preferences for that matter. It may be much more reasonable to expect only a *partial assessment* $\mathscr{A} \subseteq \mathscr{H}$ of the measurements — or uncertain rewards — You deem desirable. And because the coherence axioms D1 to D4 are closed under taking arbitrary intersections, we can use the idea of closure, or conservative inference, to infer from this partial assessment the smallest — most conservative — coherent sets of desirable measurements that it includes, similarly to what is done in propositional logic by taking the deductive closure of a collection of propositions. In this analogy, the measurements in $\mathscr{H}_{\gneq\hat{0}}$ play the role of the logical tautologies, and the measurements in $\mathscr{H}_{\leq\hat{0}}$ that of the logical contradictions.

---

we'll discuss in Sections 4.2, 4.3, 5 and 6 remains, mutatis mutandis, valid if we consider $>$ rather than $\gneq$ as the background ordering.

To see how this works, we start by defining the *positive hull* operator, which associates with any set of measurements $\mathscr{A} \subseteq \mathscr{H}$ the smallest convex cone that includes $\mathscr{A}$:

$$\operatorname{posi}(\mathscr{A}) := \left\{ \sum_{k=1}^{n} \lambda_k \hat{A}_k \colon n \in \mathbb{N}, \lambda_k > 0, \hat{A}_k \in \mathscr{A} \right\}.$$

It's now clear that $\mathscr{C}(\mathscr{A}) := \operatorname{posi}(\mathscr{A} \cup \mathscr{H}_{\gneq \hat{0}})$ is the smallest set of measurements that includes $\mathscr{A}$ and satisfies axioms D2 to D4. Observe that

$$\begin{aligned}\mathscr{C}(\mathscr{A}) &= \operatorname{posi}\big(\mathscr{A} \cup \mathscr{H}_{\gneq \hat{0}}\big) = \operatorname{posi}\big(\mathscr{H}_{\gneq \hat{0}}\big) \cup \operatorname{posi}(\mathscr{A}) \cup \big(\operatorname{posi}(\mathscr{A}) + \operatorname{posi}(\mathscr{H}_{\gneq \hat{0}})\big) \\ &= \mathscr{H}_{\gneq \hat{0}} \cup \operatorname{posi}(\mathscr{A}) \cup \big(\operatorname{posi}(\mathscr{A}) + \mathscr{H}_{\gneq \hat{0}}\big) = \mathscr{H}_{\gneq \hat{0}} \cup \big(\operatorname{posi}(\mathscr{A}) + \mathscr{H}_{\geq \hat{0}}\big). \end{aligned} \tag{16}$$

If $\hat{0} \notin \mathscr{C}(\mathscr{A})$, then $\mathscr{C}(\mathscr{A})$ is the smallest coherent set of desirable measurements that includes $\mathscr{A}$, and therefore the most conservative coherent model that's compatible with Your assessment. There are then typically infinitely many coherent extensions of $\mathscr{A}$, each one of which will include $\mathscr{C}(\mathscr{A})$ and will therefore be more committal, or less conservative, in the sense that more measurements will be included. We'll then say that Your assessment is *consistent*, because it can be extended to a coherent set of measurements. If, on the other hand, $\hat{0} \in \mathscr{C}(\mathscr{A})$ then there's no compatible coherent model, meaning that Your assessment $\mathscr{A}$ is *inconsistent*. Using Eq. (16), we can simplify the consistency condition as follows:

$$\hat{0} \notin \mathscr{C}(\mathscr{A}) \Leftrightarrow \hat{0} \notin \operatorname{posi}(\mathscr{A}) + \mathscr{H}_{\geq \hat{0}} \Leftrightarrow \operatorname{posi}(\mathscr{A}) \cap \mathscr{H}_{\leq \hat{0}} = \emptyset. \tag{17}$$

It should be clear from this discussion that the operator $\mathscr{C}$ acts like a deductive closure, or conservative inference, operator that associates with any assessment $\mathscr{A}$ the set $\mathscr{C}(\mathscr{A})$ of all measurements whose desirability can be deduced from $\mathscr{A}$ taking into account the 'production axioms' D2 to D4.

Qubit running example: instalment 5

To illustrate these ideas, we go back to the example of a qubit. If You have no knowledge at all about the system's state, then the set $\mathscr{H}_{\gneq \hat{0}}$ of all non-null positive semidefinite Hermitian operators is a reasonable model for Your desirable measurements.

Suppose You make the assessment that $\hat{A}_o = w_o \hat{I} + x_o \hat{\sigma}_x + y_o \hat{\sigma}_y + z_o \hat{\sigma}_z$ is desirable as well, for some $w_o, x_o, y_o, z_o \in \mathbb{R}$. If we recall Eq. (16), then we see that the most conservative set of desirable measurements that corresponds to the assessment $\mathscr{A} := \{\hat{A}_o\}$ is

$$\begin{aligned}\mathscr{C}(\mathscr{A}) &= \mathscr{H}_{\gneq \hat{0}} \cup \big(\operatorname{posi}(\mathscr{A}) + \mathscr{H}_{\geq \hat{0}}\big) = \mathscr{H}_{\gneq \hat{0}} \cup \big(\{a\hat{A}_o \colon a > 0\} + \mathscr{H}_{\geq \hat{0}}\big) \qquad (18) \\ &= \Big\{ (aw_o + w)\hat{I} + (ax_o + x)\hat{\sigma}_x + (ay_o + y)\hat{\sigma}_y + (az_o + z)\hat{\sigma}_z \colon \\ &\qquad\qquad (a, w) > 0 \text{ and } w \geq \sqrt{x^2 + y^2 + z^2} \Big\} \\ &= \Big\{ w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z \colon \\ &\qquad (a, w - aw_o) > 0 \text{ and } w - aw_o \geq \sqrt{(x - ax_o)^2 + (y - ay_o)^2 + (z - az_o)^2} \Big\}, \qquad (19)\end{aligned}$$

where we let '$(a,b) > 0$' be a shorthand for '$a, b \geq 0$ and $a + b > 0$'; the third equality follows from Eqs. (10) and (11) and the final equality from the replacement $ax_o + x \rightsquigarrow x$, $ay_o + y \rightsquigarrow y$, $az_o + z \rightsquigarrow z$ and $aw_o + w \rightsquigarrow w$. Eq. (17) guarantees that the consistency condition $\hat{0} \notin \mathscr{C}(\mathscr{A})$ will be satisfied if and only if $\hat{A}_o \notin \mathscr{H}_{\leq \hat{0}}$, which is in its turn equivalent to $\sqrt{x_o^2 + y_o^2 + z_o^2} > -w_o$, by Eq. (10). In that case, Your assessment that $\hat{A}_o$ is desirable, is consistent, and $\mathscr{C}(\{\hat{A}_o\})$ is the set of all measurements that You ought to then also deem desirable, as a result of Your making the desirability assessment $\mathscr{A} = \{\hat{A}_o\}$. ∮

5. Archimedean models: coherent lower and upper previsions

Coherent sets of desirable measurements $\mathscr{D}$ are our most general models for describing Your uncertainty about the system state $|\Psi\rangle$. They are, indeed, very general and can be quite complex. Their complexity can be reduced somewhat by limiting ourselves to the special case of so-called *Archimedean* models, where the preference ordering is essentially reduced to comparing real numbers. This simplification is achieved by looking at *coherent lower* and *upper previsions*, which are studied in the imprecise probabilities literature; see for instance, Refs. [2, 19, 46, 50, 51]. Doing so will take us closer to the language of probabilities, but we'll warn in Section 6.3 against too naive an enthusiasm in making such interpretational jumps.

Here and in Section 6, we'll discuss how coherent (lower and upper) previsions can be derived from coherent sets of desirable measurements, what their properties are, and how they can be seen to provide a connection with the existing probabilistic models for quantum mechanics, such as Born's rule and the use of density matrices.

In the imprecise probabilities literature, coherent lower and upper previsions are typically defined for *gambles*, which are bounded real-valued maps interpreted as uncertain rewards. To apply the theory to the present context, where we'll define them for *measurements*, rather than gambles, two roads lie open to us, and we could venture upon both. The more circuitous one consists in recalling from the discussion in Section 4.1 that measurements $\hat{A}$ are in a one-to-one relationship with their reward functions $u_{\hat{A}}$, which are gambles. Everything that can be done for gambles [2, 46, 50, 51] can therefore also indirectly be done for measurements, through the identification between the two. The more direct road, which we'll follow below for the most part, exploits the fact that coherent lower and upper previsions have also been defined and studied directly for abstract vectors that live in a normed linear space [19], rather than for gambles. The discussion below for measurements can therefore also be seen as a direct instantiation of that study. Even though this means that, in much of the discussion below, we could content ourselves by merely stating results and referring to the above-mentioned literature for their proofs, we'll nevertheless provide short proofs for all of them, to make the discussion sufficiently self-contained.

Let's start with a coherent set of desirable measurements $\mathscr{D}$ as a model for Your beliefs about $|\Psi\rangle$. Its elements $\hat{A} \in \mathscr{D}$ are the measurements (acts) You strictly prefer to the status quo $\hat{0}$. We now define Your *supremum buying price* $\underline{\Lambda}_{\mathscr{D}}(\hat{A})$ for measurement $\hat{A} \in \mathscr{H}$ as

$$\underline{\Lambda}_{\mathscr{D}}(\hat{A}) := \sup\{\alpha \in \mathbb{R} \colon \hat{A} - \alpha\hat{I} \in \mathscr{D}\} = \sup\{\alpha \in \mathbb{R} \colon \hat{A} \rhd \alpha\hat{I}\}, \tag{20}$$

the supremum (certain) price — amount of utility — You're willing to pay to acquire the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ associated with performing measurement $\hat{A}$. Let's explain better. The statement '$\hat{A} - \alpha\hat{I} \in \mathscr{D}$' means that the uncertain reward $u_{\hat{A}-\alpha\hat{I}}(|\Psi\rangle) = u_{\hat{A}}(|\Psi\rangle) - \alpha u_{\hat{I}}(|\Psi\rangle) = u_{\hat{A}}(|\Psi\rangle) - \alpha$ is desirable to You, or in other words, that You find it desirable to get the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ and give away the fixed amount of utility α — to buy the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ for a price α. Hence, our use of the name 'supremum buying price' for $\underline{\Lambda}_{\mathscr{D}}(\hat{A})$.

Similarly, we define Your *infimum selling price* $\overline{\Lambda}_{\mathscr{D}}(\hat{A})$ for measurement $\hat{A} \in \mathscr{H}$ as

$$\overline{\Lambda}_{\mathscr{D}}(\hat{A}) := \inf\{\beta \in \mathbb{R} \colon \beta\hat{I} - \hat{A} \in \mathscr{D}\} = \inf\{\beta \in \mathbb{R} \colon \beta\hat{I} \rhd \hat{A}\} = -\underline{\Lambda}_{\mathscr{D}}(-\hat{A}), \tag{21}$$

the infimum (certain) price — amount of utility — You're willing to receive to give away the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ associated with performing measurement $\hat{A}$. We'll also call $\underline{\Lambda}_{\mathscr{D}}$ and $\overline{\Lambda}_{\mathscr{D}}$ *price functionals* for the set of desirable measurements $\mathscr{D}$.

Qubit running example: instalment 6

Let's return to our running example about the qubit system, where we recall the assessment $\mathscr{A} := \{\hat{A}_o\} = \{w_o\hat{I} + x_o\hat{\sigma}_x + y_o\hat{\sigma}_y + z_o\hat{\sigma}_z\}$ considered earlier in Instalment 5.

Let $\hat{A}_o = \hat{\sigma}_x$, so $(w_o, x_o, y_o, z_o) = (0,1,0,0)$. With this assessment, You express that You consider it more likely that measuring the spin in the x-direction will yield the outcome $+1$ than that it will yield the other possible outcome -1.

It's then easy to see that $0 = -w_o < \sqrt{x_o^2 + y_o^2 + z_o^2} = 1$, so this assessment is consistent, and therefore $\mathscr{C}(\mathscr{A})$ as characterised by Eq. (19) is the smallest coherent set of desirable measurements that contains $\hat{A}_o = \hat{\sigma}_x$ and therefore represents Your assessment. Observe that, with $\hat{C} := w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z$ and $\alpha \in \mathbb{R}$,

$$\hat{C} - \alpha\hat{I} \in \operatorname{posi}(\mathscr{A}) + \mathscr{H}_{\geq \hat{0}} \Leftrightarrow (\exists a > 0)\alpha \leq w - \sqrt{(x-a)^2 + y^2 + z^2}$$

and

$$\hat{C} - \alpha\hat{I} \in \mathscr{H}_{\gneq \hat{0}} \Leftrightarrow \big(w \neq \alpha \text{ and } \alpha \leq w - \sqrt{x^2 + y^2 + z^2}\big),$$

where we started from Eq. (18) and used Eqs. (10) and (11). These conditions, together with Eq. (20), allow us to find the values for $\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C})$ and $\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C}) = -\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(-\hat{C})$ in any measurement $\hat{C} \in \mathscr{H}$. After some calculations, we find that

$$\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z) = \begin{cases} w - \sqrt{y^2 + z^2} & \text{if } x \geq 0 \\ w - \sqrt{x^2 + y^2 + z^2} & \text{if } x \leq 0 \end{cases} \text{ and}$$

$$\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z) = \begin{cases} w + \sqrt{x^2 + y^2 + z^2} & \text{if } x \geq 0 \\ w + \sqrt{y^2 + z^2} & \text{if } x \leq 0, \end{cases} \tag{22}$$

so, in particular,

$$\begin{cases} \underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_x) = 0 \\ \overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_x) = 1 \end{cases} \text{and } \begin{cases} \underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_y) = -1 \\ \overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_y) = 1 \end{cases} \text{and } \begin{cases} \underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_x - \hat{\sigma}_y) = -1 \\ \overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{\sigma}_x - \hat{\sigma}_y) = \sqrt{2}. \end{cases} \tag{23}$$

The discussion in Section 6, and in particular in Section 6.3, will allow us to conclude that, on a more traditional view of probability in quantum mechanics, $\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{A})$ and $\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{A})$ *can* be interpreted as tight lower and upper bounds on the expected values of the outcome of the measurement $\hat{A}$, based on Your assessment $\mathscr{A}$. We now see in the expressions above that Your assessment that 1 is a likelier outcome than -1 for the measurement $\hat{\sigma}_x$, allows You to tightly bound the expected value of its outcome to lie between 0 and 1. For the expected outcome of the measurement $\hat{\sigma}_y$, Your assessment $\mathscr{A}$ leads to the *vacuous* bounds -1 and 1, so it's really totally non-informative about it.[17]

By the way, the difference $\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C}) - \underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C})$ is called the *imprecision*, or *bid-ask spread*, for $\hat{C} = w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z$, and it's given by $\sqrt{x^2 + y^2 + z^2} + \sqrt{y^2 + z^2}$.

Although typically somewhat more involved, essentially the same type of arguments can be used to find $\mathscr{C}(\mathscr{A})$ and $\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}$ and $\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}$ for more involved assessments $\mathscr{A}$ and more general Hilbert spaces $\mathscr{X}$, and they will typically involve solving some kind of semi-definite programming problem [8, Sec. 4.6.2]. ♪

The price functionals $\underline{\Lambda}_{\mathscr{D}}$ that are obtained by letting $\mathscr{D}$ range over all the coherent sets of desirable measurements are exactly all the real functionals $\underline{\Lambda} : \mathscr{H} \to \mathbb{R}$ that satisfy the following properties:

LP1. $\underline{\Lambda}(\hat{A}) \geq \min \operatorname{spec}(\hat{A})$ for all $\hat{A} \in \mathscr{H}$; [bounds]
LP2. $\underline{\Lambda}(\hat{A} + \hat{B}) \geq \underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B})$ for all $\hat{A}, \hat{B} \in \mathscr{H}$; [super-additivity]
LP3. $\underline{\Lambda}(\lambda\hat{A}) = \lambda\underline{\Lambda}(\hat{A})$ for all $\hat{A} \in \mathscr{H}$ and all real $\lambda \geq 0$. [non-negative homogeneity]

Proof. First, let's assume that $\mathscr{D}$ is a coherent set of desirable measurements, and show that the corresponding price functional $\underline{\Lambda}_{\mathscr{D}}$ is real-valued and satisfies LP1 to LP3. First, consider any $\hat{A} \in \mathscr{H}$, then it follows from D6 that the set $L_{\hat{A}} := \{\alpha \in \mathbb{R} \colon \hat{A} - \alpha\hat{I} \in \mathscr{D}\}$ is

[17] This can be related to the *incommensurability* — non-commuting character — of $\hat{\sigma}_x$ and $\hat{\sigma}_y$.

decreasing and therefore bounded above by any element of its set-theoretic complement $\mathbb{R} \setminus L_{\hat{A}} = \{\alpha \in \mathbb{R} \colon \hat{A} - \alpha \hat{I} \notin \mathscr{D}\}$. Since it readily follows from D5 that $\max \operatorname{spec}(\hat{A}) \notin L_{\hat{A}}$ and from D4 that $(-\infty, \min \operatorname{spec}(\hat{A})) \subseteq L_{\hat{A}}$, we find that $\min \operatorname{spec}(\hat{A}) \leq \sup L_{\hat{A}} \leq \max \operatorname{spec}(\hat{A})$ and therefore that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}) = \sup L_{\hat{A}}$ is real. This also shows that $\underline{\Lambda}_{\mathscr{D}}$ satisfies LP1. Also, consider any $\alpha, \beta \in \mathbb{R}$ such that $\hat{A} - \alpha \hat{I} \in \mathscr{D}$ and $\hat{B} - \beta \hat{I} \in \mathscr{D}$, then $\hat{A} + \hat{B} - (\alpha + \beta)\hat{I} \in \mathscr{D}$ by D2. Invoking Eq. (20) now leads to the conclusion that $\underline{\Lambda}_{\mathscr{D}}$ satisfies LP2. For LP3, first infer from LP1 and LP2 for $\hat{A} = \hat{B} = \hat{0}$ that $\underline{\Lambda}_{\mathscr{D}}(\hat{0}) = 0$, so we need only check that $\underline{\Lambda}_{\mathscr{D}}$ satisfies LP3 for $\lambda > 0$. Now observe that, by D3, $\lambda \hat{A} - \alpha \hat{I} \in \mathscr{D} \Leftrightarrow \hat{A} - {}^{\alpha}\!/\!{}_{\lambda} \hat{I} \in \mathscr{D}$, and invoke Eq. (20).

Conversely, assume that the real functional $\underline{\Lambda} \colon \mathscr{H} \to \mathbb{R}$ satisfies LP1 to LP3, then we must show that there's some coherent set of desirable measurements $\mathscr{D}$ such that $\underline{\Lambda} = \underline{\Lambda}_{\mathscr{D}}$. If we let

$$\mathscr{D}_{\underline{\Lambda}} := \{\hat{A} \in \mathscr{H} \colon \underline{\Lambda}(\hat{A}) > 0 \text{ or } \hat{A} \gneq \hat{0}\}, \tag{24}$$

then

$$\mathscr{D}'_{\underline{\Lambda}} := \{\hat{A} \in \mathscr{H} \colon \underline{\Lambda}(\hat{A}) > 0\} \subseteq \mathscr{D}_{\underline{\Lambda}} \subseteq \{\hat{A} \in \mathscr{H} \colon \underline{\Lambda}(\hat{A}) \geq 0\} =: \mathscr{D}''_{\underline{\Lambda}}, \tag{25}$$

where the second inclusion follows from LP1. Also observe that by the constant additivity property LP6, which as we'll show further on follows from LP1 to LP3, it holds that

$$\hat{A} - \alpha \hat{I} \in \mathscr{D}'_{\underline{\Lambda}} \Leftrightarrow \underline{\Lambda}(\hat{A}) > \alpha \text{ and } \hat{A} - \alpha \hat{I} \in \mathscr{D}''_{\underline{\Lambda}} \Leftrightarrow \underline{\Lambda}(\hat{A}) \geq \alpha \text{ for all } \alpha \in \mathbb{R}. \tag{26}$$

If we now combine Eqs. (25) and (26) with Eq. (20), we find that $\underline{\Lambda} = \underline{\Lambda}_{\mathscr{D}_{\underline{\Lambda}}}$, so it only remains to show that $\mathscr{D}_{\underline{\Lambda}}$ is (a) coherent (set of desirable measurements), or in other words satisfies D1 to D4. For D1, simply observe that we've already proved above that $\underline{\Lambda}(\hat{0}) = 0$, so indeed $\hat{0} \notin \mathscr{D}_{\underline{\Lambda}}$. For D2, consider any $\hat{A}, \hat{B} \in \mathscr{D}_{\underline{\Lambda}}$, then there are a number of possibilities. If $\hat{A} \gneq \hat{0}$ and $\hat{B} \gneq \hat{0}$, then clearly also $\hat{A} + \hat{B} \gneq \hat{0}$, and therefore $\hat{A} + \hat{B} \in \mathscr{D}_{\underline{\Lambda}}$. If $\underline{\Lambda}(\hat{A}) > 0$ and $\underline{\Lambda}(\hat{B}) > 0$, then by LP2 also $\underline{\Lambda}(\hat{A} + \hat{B}) \geq \underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) > 0$, and therefore $\hat{A} + \hat{B} \in \mathscr{D}_{\underline{\Lambda}}$. For the remaining possibilities, we may assume without loss of generality that $\underline{\Lambda}(\hat{A}) > 0$ and $\hat{B} \gneq \hat{0}$. But then $\underline{\Lambda}(\hat{B}) \geq 0$ by LP1, and therefore by LP2 also $\underline{\Lambda}(\hat{A} + \hat{B}) \geq \underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) > 0$, whence, here too, $\hat{A} + \hat{B} \in \mathscr{D}_{\underline{\Lambda}}$. That $\mathscr{D}_{\underline{\Lambda}}$ satisfies D3 follows readily from LP3, and, finally, $\mathscr{D}_{\underline{\Lambda}}$ satisfies D4 by construction. □

Adopting Walley's [50] nomenclature,[18] we'll call any map $\underline{\Lambda} \colon \mathscr{H} \to \mathbb{R}$ that satisfies the so-called *coherence conditions* LP1 to LP3 a *coherent lower prevision* on $\mathscr{H}$.

The corresponding *coherent upper previsions* $\overline{\Lambda}$ are related to the coherent lower previsions $\underline{\Lambda}$ through the *conjugacy relationship*: $\overline{\Lambda}(\hat{A}) = -\underline{\Lambda}(-\hat{A})$ for all $\hat{A} \in \mathscr{H}$.

Coherent lower and upper previsions also always satisfy the following useful properties, which are immediate consequences of the coherence conditions LP1 to LP3. The convergence in LP8 is convergence in supremum norm $\|\bullet\|_{\infty}$, or equivalently, since all norms are equivalent — lead to the same topology — on the finite dimensional space $\mathscr{H}$, convergence in Frobenius norm $\|\bullet\|_{\mathrm{F}}$ or in operator norm $\|\bullet\|_{\mathrm{op}}$.

LP4. $\min \operatorname{spec}(\hat{A}) \leq \underline{\Lambda}(\hat{A}) \leq \overline{\Lambda}(\hat{A}) \leq \max \operatorname{spec}(\hat{A})$; [more bounds]
LP5. $\underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) \leq \underline{\Lambda}(\hat{A} + \hat{B}) \leq \underline{\Lambda}(\hat{A}) + \overline{\Lambda}(\hat{B}) \leq \overline{\Lambda}(\hat{A} + \hat{B}) \leq \overline{\Lambda}(\hat{A}) + \overline{\Lambda}(\hat{B})$;
LP6. $\overline{\Lambda}(\hat{A} + \mu \hat{I}) = \overline{\Lambda}(\hat{A}) + \mu \hat{I}$ and $\underline{\Lambda}(\hat{A} + \mu) = \underline{\Lambda}(\hat{A}) + \mu$; [constant additivity]
LP7. if $\hat{A} \geq \hat{B}$ then $\overline{\Lambda}(\hat{A}) \geq \overline{\Lambda}(\hat{B})$ and $\underline{\Lambda}(\hat{A}) \geq \underline{\Lambda}(\hat{B})$; [monotonicity]
LP8. if the sequence $\hat{A}_n \in \mathscr{H}$ converges to the measurement $\hat{A}$, then $\overline{\Lambda}(\hat{A}_n) \to \overline{\Lambda}(\hat{A})$ and $\underline{\Lambda}(\hat{A}_n) \to \underline{\Lambda}(\hat{A})$. [continuity]

[18] Since, as we've shown in Eq. (14), $\min \operatorname{spec}(\hat{A}) = \inf u_{\hat{A}}$ for all $\hat{A} \in \mathscr{H}$, the correspondence between our coherence requirements for lower previsions on the linear space of measurements and Walley's coherence result for lower previsions on a subspace of the space of all gambles, is immediate; see [50, Thm. 2.8.4].

Proof. LP4. The first and third inequalities follow directly from LP1 and conjugacy, so we concentrate on the second inequality. Simply observe that

$$0=\underline{\Lambda}(\hat{0})=\underline{\Lambda}(\hat{A}-\hat{A})\geq\underline{\Lambda}(\hat{A})+\underline{\Lambda}(-\hat{A})=\underline{\Lambda}(\hat{A})-\overline{\Lambda}(\hat{A}),$$

where the first equality follows from LP3 and the first inequality from LP2.

LP5. It suffices to prove the second equality, as all the other inequalities will follow then by also invoking LP2 and conjugacy. Now simply observe that

$$\underline{\Lambda}(\hat{A})=\underline{\Lambda}(\hat{A}+\hat{B}-\hat{B})\geq\underline{\Lambda}(\hat{A}+\hat{B})+\underline{\Lambda}(-\hat{B})=\underline{\Lambda}(\hat{A}+\hat{B})-\overline{\Lambda}(\hat{B}),$$

where the first inequality follows from LP2.

LP6. It suffices to prove the second equality, as the first then follows by conjugacy. First, infer from LP4 that $\underline{\Lambda}(\mu\hat{I})=\overline{\Lambda}(\mu\hat{I})=\mu$, and then invoke LP5 to find that

$$\underline{\Lambda}(\hat{A})+\mu=\underline{\Lambda}(\hat{A})+\underline{\Lambda}(\mu\hat{I})\leq\underline{\Lambda}(\hat{A}+\mu\hat{I})\leq\underline{\Lambda}(\hat{A})+\overline{\Lambda}(\mu\hat{I})=\underline{\Lambda}(\hat{A})+\mu.$$

LP7. We concentrate on proving the second implication, as the first will then follow from conjugacy. Assume that $\hat{A}\geq\hat{B}$, or equivalently, $\hat{A}-\hat{B}\geq\hat{0}$, then

$$0\leq\min\operatorname{spec}(\hat{A}-\hat{B})\leq\underline{\Lambda}(\hat{A}-\hat{B})\leq\underline{\Lambda}(\hat{A})+\overline{\Lambda}(-\hat{B})=\underline{\Lambda}(\hat{A})-\underline{\Lambda}(\hat{B}),$$

where the first inequality follows from the definition of the vector ordering $\geq$, the second inequality follows from LP1 and the third inequality from LP5.

LP8. We concentrate on proving the second implication, as the first will then follow from conjugacy. Simply observe that

$$\underline{\Lambda}(\hat{A})-\underline{\Lambda}(\hat{B})=\underline{\Lambda}(\hat{A})+\overline{\Lambda}(-\hat{B})\leq\overline{\Lambda}(\hat{A}-\hat{B})\leq\max\operatorname{spec}(\hat{A}-\hat{B})\leq\|\hat{A}-\hat{B}\|_\infty,$$

where the first inequality follows from LP5, the second one from LP4 and the third one from Eq. (12). Since, similarly, $\underline{\Lambda}(\hat{B})-\underline{\Lambda}(\hat{A})\leq\|\hat{A}-\hat{B}\|_\infty$, we find that

$$|\underline{\Lambda}(\hat{A})-\underline{\Lambda}(\hat{B})|\leq\|\hat{A}-\hat{B}\|_\infty\text{ for all }\hat{A},\hat{B}\in\mathscr{H},$$

which even proves the Lipschitz continuity of $\underline{\Lambda}$. □

To better understand the relationship between coherent sets of desirable measurements and coherent lower previsions, we observe that

$$\hat{A}\in\operatorname{Int}(\mathscr{D})\Leftrightarrow\underline{\Lambda}_{\mathscr{D}}(\hat{A})>0\text{ and }\hat{A}\in\operatorname{Cl}(\mathscr{D})\Leftrightarrow\underline{\Lambda}_{\mathscr{D}}(\hat{A})\geq0,\text{ for all }\hat{A}\in\mathscr{H}.\tag{27}$$

Proof. To prove Eq. (27), consider any $\hat{A}\in\mathscr{H}$.

Assume that $\hat{A}\in\operatorname{Int}(\mathscr{D})$, implying that there's some real $\varepsilon>0$ such that the open ball $B_\varepsilon(\hat{A}):=\{\hat{B}\in\mathscr{H}\colon\|\hat{B}-\hat{A}\|_\infty<\varepsilon\}\subseteq\mathscr{D}$. Since $\|(\hat{A}-\varepsilon/2\hat{I})-\hat{A}\|_\infty=\varepsilon/2$, we find that $\hat{A}-\varepsilon/2\hat{I}\in\mathscr{D}$, and therefore, indeed, $\underline{\Lambda}_{\mathscr{D}}(\hat{A})\geq\varepsilon/2>0$.

Assume that $\hat{A}\notin\operatorname{Int}(\mathscr{D})$, and therefore $\hat{A}\in\operatorname{Cl}(\mathscr{H}\setminus\mathscr{D})$, implying that there's some sequence $\hat{A}_n\in\mathscr{H}\setminus\mathscr{D}$ such that $\|\hat{A}-\hat{A}_n\|_\infty\to0$. But $\hat{A}_n\notin\mathscr{D}$ implies that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}_n)\leq0$ [use Lemma 8], and therefore the continuity of $\underline{\Lambda}_{\mathscr{D}}$ [use LP8] implies that also $\underline{\Lambda}_{\mathscr{D}}(\hat{A})\leq0$.

Assume that $\hat{A}\in\operatorname{Cl}(\mathscr{D})$, which implies that there's some sequence $\hat{A}_n\in\mathscr{D}$ such that $\|\hat{A}-\hat{A}_n\|_\infty\to0$. But $\hat{A}_n\in\mathscr{D}$ implies that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}_n)\geq0$, and therefore the continuity of $\underline{\Lambda}_{\mathscr{D}}$ [use LP8] implies that also $\underline{\Lambda}_{\mathscr{D}}(\hat{A})\geq0$.

Assume that $\hat{A}\notin\operatorname{Cl}(\mathscr{D})$ and therefore $\hat{A}\in\operatorname{Int}(\mathscr{H}\setminus\mathscr{D})$, implying that there's some real $\varepsilon>0$ such that the open ball $B_\varepsilon(\hat{A}):=\{\hat{B}\in\mathscr{H}\colon\|\hat{B}-\hat{A}\|_\infty<\varepsilon\}\subseteq\mathscr{H}\setminus\mathscr{D}$. Since $\|(\hat{A}+\varepsilon/2\hat{I})-\hat{A}\|_\infty=\varepsilon/2$, we find that $\hat{A}+\varepsilon/2\hat{I}\notin\mathscr{D}$, which implies that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}+\varepsilon/2\hat{I})\leq0$ [use Lemma 8] and therefore $\underline{\Lambda}_{\mathscr{D}}(\hat{A})\leq-\varepsilon/2<0$ [use LP6]. □

Lemma 8

Consider any coherent set of desirable measurements $\mathscr{D}$ and let $\underline{\Lambda}_{\mathscr{D}}$ be the corresponding price functional. Then, for any $\hat{A}\in\mathscr{H}$, we have that $\hat{A}\notin\mathscr{D}\Rightarrow\underline{\Lambda}_{\mathscr{D}}(\hat{A})\leq0$.

Proof. That $\hat{A} \notin \mathscr{D}$, or in other words, that $\hat{A} - 0\hat{I} \notin \mathscr{D}$, implies that 0 dominates all elements of the decreasing set $L_{\hat{A}} = \{\alpha \in \mathbb{R} \colon \hat{A} - \alpha\hat{I} \in \mathscr{D}\}$ and therefore dominates its supremum $\underline{\Lambda}_{\mathscr{D}}(\hat{A})$. □

Eq. (27) tells us that all measurements $\hat{A}$ with a positive supremum buying price $\underline{\Lambda}_{\mathscr{D}}(\hat{A})$ lie definitely inside $\mathscr{D}$, and those with a negative supremum buying price definitely outside $\mathscr{D}$. It follows at once from Eq. (27) that

$$\hat{A} \in \mathrm{Cl}(\mathscr{D}) \setminus \mathrm{Int}(\mathscr{D}) \Leftrightarrow \underline{\Lambda}_{\mathscr{D}}(\hat{A}) = 0, \text{ for all } \hat{A} \in \mathscr{H},$$

so for the so-called *marginally desirable* measurements in the topological boundary of the convex cone $\mathscr{D}$, which are the measurements with supremum buying price zero, their supremum buying price typically won't determine whether they belong to $\mathscr{D}$, unless $\mathscr{D}$ is known to satisfy additional properties, besides coherence.[19] The upshot of this is that the price functional $\underline{\Lambda}_{\mathscr{D}}$ characterises the set of desirable measurements $\mathscr{D}$ it's derived from *only up to the measurements in its topological boundary*. Coherent lower previsions therefore have at most as much, and typically (slightly) less, representational power than coherent sets of desirable measurements: the correspondence between coherent sets of desirable measurements and coherent lower previsions is typically many-to-one. As is apparent from Eq. (24) in the proof above that investigates the correspondence between the two models, one way to make a coherent lower prevision $\underline{\Lambda}$ correspond to exactly one strict preference $\rhd$, is to pick the one given by

$$\hat{A} \rhd \hat{B} \Leftrightarrow \hat{A} - \hat{B} \in \mathscr{D}_{\underline{\Lambda}} \Leftrightarrow \big(\underline{\Lambda}(\hat{A} - \hat{B}) > 0 \text{ or } \hat{A} \gneq \hat{B}\big) \text{ for all } \hat{A}, \hat{B} \in \mathscr{H}. \tag{28}$$

To conclude this discussion of coherent lower and upper previsions, let's consider a few special cases. At one extreme end, we have the price functionals for the vacuous set of desirable measurements $\mathscr{D}_{\mathrm{vac}} = \mathscr{H}_{\gneq\hat{0}}$, which are the so-called *vacuous* lower and upper previsions, defined by

$$\underline{\Lambda}_{\mathrm{vac}}(\hat{A}) = \min \mathrm{spec}(\hat{A}) \text{ and } \overline{\Lambda}_{\mathrm{vac}}(\hat{A}) = \max \mathrm{spec}(\hat{A}) \text{ for all } \hat{A} \in \mathscr{H}.$$

Here, the *bid-ask spread* $\overline{\Lambda}_{\mathrm{vac}}(\hat{A}) - \underline{\Lambda}_{\mathrm{vac}}(\hat{A}) = \max \mathrm{spec}(\hat{A}) - \min \mathrm{spec}(\hat{A})$ between Your selling and buying prices for measurements $\hat{A}$ is as large as allowed by the coherence requirements [in particular LP1].

At the other extreme end, we recover the cases that are most often considered in more classical approaches to decision theory, where the bid-ask spread $\overline{\Lambda}(\hat{A}) - \underline{\Lambda}(\hat{A})$ between Your infimum selling and supremum buying prices for measurements $\hat{A}$ is as small as it's allowed to be by the coherence requirements [in particular LP4], namely 0 everywhere. This is the type of uncertainty model we'll now turn to in the coming section.

6. Coherent previsions and density operators

6.1. Bruno de Finetti's coherent previsions.

Coherent lower previsions $\underline{\Lambda}$ with the smallest allowable bid-ask spread are *self-conjugate* in the sense that $\underline{\Lambda} = \overline{\Lambda}$, so we can and will use the simpler notation Λ for $\underline{\Lambda} = \overline{\Lambda}$ in this case.

According to the discussion and definitions in the previous section, $\Lambda(\hat{A})$ is then at the same time Your supremum price for buying the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ and Your infimum price for selling it, and can therefore also be seen as *Your fair price* for the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$.

In conjunction with self-conjugacy, the coherence conditions LP1 to LP3 turn into

CP1. $\Lambda(\hat{A}) \geq \min \mathrm{spec}(\hat{A})$ for all $\hat{A} \in \mathscr{H}$; [bounds]

CP2. $\Lambda(\hat{A} + \hat{B}) = \Lambda(\hat{A}) + \Lambda(\hat{B})$ for all $\hat{A}, \hat{B} \in \mathscr{H}$. [additivity]

Homogeneity then follows from these two coherence conditions:

CP3. $\Lambda(\lambda\hat{A}) = \lambda\Lambda(\hat{A})$ for all $\hat{A} \in \mathscr{H}$ and $\lambda \in \mathbb{R}$. [homogeneity]

[19]… such as, for instance, being open.

Proof. First, CP2 implies that CP3 holds for all rational λ. It's clear that we can extend the equality to the reals, provided we can show that CP1 and CP2 imply the continuity of Λ. To this end, observe that CP2 implies in particular that $\Lambda(\hat{0}) = 0$ and $\Lambda(-\hat{A}) = -\Lambda(\hat{A})$ for all $\hat{A} \in \mathscr{H}$. But then also

$$\begin{aligned}\Lambda(\hat{A}) - \Lambda(\hat{B}) &= \Lambda(\hat{A}) + \Lambda(-\hat{B}) \\ &= \Lambda(\hat{A} - \hat{B}) = -\Lambda(-(\hat{A} - \hat{B})) \\ &\leq \max \operatorname{spec}(\hat{A} - \hat{B}) \leq \|\hat{A} - \hat{B}\|_\infty,\end{aligned}$$

where the second equality follows from CP2, the first inequality follows from CP1, and the second inequality from Eq. (12). Similarly, $\Lambda(\hat{B}) - \Lambda(\hat{A}) \leq \|\hat{A} - \hat{B}\|_\infty$, so

$$|\Lambda(\hat{A}) - \Lambda(\hat{B})| \leq \|\hat{A} - \hat{B}\|_\infty \text{ for all } \hat{A}, \hat{B} \in \mathscr{H},$$

which even proves the Lipschitz continuity of Λ.

It's now clearly enough to prove that we can derive CP1 and CP2 from the conditions LP1 to LP3 and self-conjugacy, which tells us that $\Lambda = \underline{\Lambda} = \overline{\Lambda}$. CP1 is clearly equivalent to LP1. For CP2, consider any $\hat{A}, \hat{B} \in \mathscr{H}$. Self-conjugacy implies that $\underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) = \overline{\Lambda}(\hat{A}) + \overline{\Lambda}(\hat{B})$. On the other hand, it follows from LP5 [which follows from LP2] that $\underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) \leq \underline{\Lambda}(\hat{A} + \hat{B}) \leq \overline{\Lambda}(\hat{A} + \hat{B}) \leq \overline{\Lambda}(\hat{A}) + \overline{\Lambda}(\hat{B})$ and therefore, indeed, that $\underline{\Lambda}(\hat{A}) + \underline{\Lambda}(\hat{B}) = \underline{\Lambda}(\hat{A} + \hat{B})$. □

It's easy to see that the self-conjugate coherent lower previsions can also be characterised exactly as all real functionals $\Lambda : \mathscr{H} \to \mathbb{R}$ that are (i) *linear* in the sense that

L1. $\Lambda(\lambda \hat{A} + \mu \hat{B}) = \lambda \Lambda(\hat{A}) + \mu \Lambda(\hat{B})$ for all $\hat{A}, \hat{B} \in \mathscr{H}$ and all $\lambda, \mu \in \mathbb{R}$;

(ii) *positive* in the sense that

L2. $\Lambda(\hat{A}) \geq 0$ for all $\hat{A} \geq \hat{0}$;

and (iii) *normalised* in the sense that

L3. $\Lambda(\hat{I}) = 1$.

Proof. Clearly, L1 is equivalent to CP2 and CP3. Condition L2 is an immediate consequence of CP1. Applying CP1 to $\hat{I}$ and $-\hat{I}$ leads to L3. It now only remains to prove that CP1 follows from L1 to L3. Consider, to this end, any measurement $\hat{A}$ and let $c := \min \operatorname{spec}(\hat{A})$, then $\hat{A} - c\hat{I} \geq \hat{0}$ and therefore $\Lambda(\hat{A} - c\hat{I}) \geq 0$, due to L2. L1 and L3 now ensure that $\Lambda(\hat{A} - c\hat{I}) = \Lambda(\hat{A}) - c\Lambda(\hat{I}) = \Lambda(\hat{A}) - c$. □

If we follow the standard functional-analytic approach in defining the following operator norm

$$\|\Gamma\|_{\text{op}} := \sup_{\hat{A} \neq \hat{0}} \frac{|\Gamma(\hat{A})|}{\|\hat{A}\|_\infty}$$

on (linear) real functionals $\Gamma \colon \mathscr{H} \to \mathbb{R}$, then it follows from Eq. (12), LP4 and L3 that

$$\|\Lambda\|_{\text{op}} = 1 \text{ for all self-conjugate coherent lower previsions } \Lambda; \tag{29}$$

they're bounded and therefore continuous, as also shown directly in the proof of CP3 above.

For these reasons, these self-conjugate coherent lower previsions can also be seen as the *coherent previsions* in the sense considered by Bruno de Finetti; see for instance Refs. [15] and [17, Ch. 3 and App.]. We denote the set of all such coherent previsions on $\mathscr{H}$ by $\mathbb{P}$. Coherent previsions on real Hilbert spaces are indeed the probability models envisaged by de Finetti, but we must caution against interpreting them without further ado as expectation operators associated with a (finitely additive) probability on some sample space, as this is an interpretational step that asks for further justification, which isn't provided by our argumentation above; see also the discussion further on in Section 6.3.

Interestingly, we can use the coherent previsions, or rather, sets of them, to characterise the coherent *lower* previsions. Taking the lower envelope of *any* non-empty set of coherent previsions $\mathcal{M} \subseteq \mathbb{P}$ leads to a coherent lower prevision $\underline{\Lambda}_{\mathcal{M}}(\bullet) := \inf\{\Lambda(\bullet) \colon \Lambda \in \mathcal{M}\}$. But, even stronger than this, it's a consequence of the Hahn–Banach theorem (or, in this finite-dimensional context, of Minkowski's hyperplane separation theorem) that any coherent lower prevision $\underline{\Lambda}$ is the lower envelope of a (non-empty) *convex and closed*[20] set of coherent previsions $\mathbb{P}(\underline{\Lambda})$, uniquely defined by

$$\mathbb{P}(\underline{\Lambda}) := \{\Lambda \in \mathbb{P} \colon (\forall \hat{A} \in \mathscr{H})\underline{\Lambda}(\hat{A}) \leq \Lambda(\hat{A})\}, \tag{30}$$

in the sense that

$$\underline{\Lambda}(\hat{A}) = \min\big\{\Lambda(\hat{A}) \colon \Lambda \in \mathbb{P}(\underline{\Lambda})\big\} \text{ for all } \hat{A} \in \mathscr{H}. \tag{31}$$

This result, also known as the *Lower Envelope Theorem*, guarantees that there's a one-to-one correspondence between coherent lower previsions and non-empty convex closed sets of coherent previsions. Following Levi's [34] terminology, we'll call the convex closed set $\mathbb{P}(\underline{\Lambda})$ the *credal set* that corresponds to, and characterises, the coherent lower prevision $\underline{\Lambda}$. We refer to Ref. [19, Cor. 13] for a more detailed argumentation, or alternatively to Refs. [50, Ch. 3] and [46, Ch. 8] for proofs in the context of gambles that can be readily transported to the present context. See also Ref. [7, Prop. IV.3], where Benavoli, Facchini and Zaffalon are the first to introduce credal sets in the quantum mechanical context, with a somewhat different interpretation.

Qubit running example: instalment 7

Let's return to our qubit system and consider an arbitrary coherent prevision Λ on $\mathscr{H}$. Since any Hermitian operator $\hat{C} \in \mathscr{H}$ can be written as $\hat{C} = w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z$ for some unique $(w,x,y,z) \in \mathbb{R}^4$, we find that, due to the linear character of Λ [coherence property L1],

$$\Lambda(\hat{C}) = w + x\Lambda(\hat{\sigma}_x) + y\Lambda(\hat{\sigma}_y) + z\Lambda(\hat{\sigma}_z), \tag{32}$$

where we also took into account that $\Lambda(\hat{I}) = 1$, by coherence property L3. We see that the coherent prevision Λ is therefore completely determined by its values $\Lambda_x := \Lambda(\hat{\sigma}_x)$, $\Lambda_y := \Lambda(\hat{\sigma}_y)$ and $\Lambda_z := \Lambda(\hat{\sigma}_z)$ in the Pauli operators. We can find the constraints imposed by coherence on these real numbers by looking at the coherence condition LP4, which, applied to any $\hat{C}$, tells us that $\min\operatorname{spec}(\hat{C}) \leq \Lambda(\hat{C}) \leq \max\operatorname{spec}(\hat{C})$. If we now recall Eq. (9) for the eigenvalues of $\hat{C}$, we see that this can be rewritten as

$$w - \sqrt{x^2+y^2+z^2} \leq w + x\Lambda_x + y\Lambda_y + z\Lambda_z \leq w + \sqrt{x^2+y^2+z^2},$$

or equivalently, as

$$(x\Lambda_x + y\Lambda_y + z\Lambda_z)^2 \leq x^2+y^2+z^2, \text{ for all } x,y,z \in \mathbb{R}.$$

Combining this with the Cauchy–Schwarz inequality, which can be rewritten as

$$\Lambda_x^2 + \Lambda_y^2 + \Lambda_z^2 = \max_{(x,y,z)\in\mathbb{R}^3\setminus\{(0,0,0)\}} \frac{(x\Lambda_x + y\Lambda_y + z\Lambda_z)^2}{x^2+y^2+z^2},$$

leads to

$$\Lambda_x^2 + \Lambda_y^2 + \Lambda_z^2 \leq 1, \tag{33}$$

which, together with $\Lambda(\hat{I}) = 1$, is all that coherence imposes on the values of a coherent prevision Λ in the Pauli operators, and therefore on its values in any $\hat{C} \in \mathscr{H}$ through Eq. (32).

[20]… in the (so-called weak*) topology of pointwise convergence on the set of all linear functionals on the linear space $\mathscr{H}$.

As we've seen in Instalment 6, the assessment $\mathscr{A} := \{\hat{\sigma}_x\}$ corresponds to a lower prevision $\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}$. Taking into account Eq. (22), we find that a coherent prevision Λ belongs to the credal set $\mathbb{P}(\underline{\Lambda}_{\mathscr{C}(\mathscr{A})})$ that corresponds to this lower prevision when

$$\begin{cases} w-\sqrt{y^2+z^2} & \text{if } x \geq 0 \\ w-\sqrt{x^2+y^2+z^2} & \text{if } x \leq 0 \end{cases} \leq w+x\Lambda_x+y\Lambda_y+z\Lambda_z \text{ for all } (x,y,y,z) \in \mathbb{R}^4,$$

and $\Lambda_x^2+\Lambda_y^2+\Lambda_z^2 \leq 1$ [these are the general coherence bounds established in Eq. (33)]. After some algebra, we find that the credal set $\mathbb{P}(\underline{\Lambda}_{\mathscr{C}(\mathscr{A})})$ is completely determined by the conditions

$$\Lambda_x \geq 0 \text{ and } \Lambda_x^2+\Lambda_y^2+\Lambda_z^2 \leq 1. \tag{34}$$

We then infer from Eq. (31) and conjugacy that, to find $\underline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C})$ and $\overline{\Lambda}_{\mathscr{C}(\mathscr{A})}(\hat{C})$, we have to minimise respectively maximise $w+x\Lambda_x+y\Lambda_y+z\Lambda_z$ subject to the constraints in Eq. (34). Some small effort shows that this leads to the same values as those given in Eq. (23) for the particular choices $\hat{\sigma}_x$, $\hat{\sigma}_y$ and $\hat{\sigma}_x-\hat{\sigma}_y$ for $\hat{C}$. These calculations correspond to solving *dual* semidefinite programming problems for the ones mentioned near the end of Instalment 6 of our running example. ♪

6.2. **Density operators.** In the more standard treatment of uncertainty in quantum mechanics, as discussed in Section 2.2, uncertainty about the state is often expressed through density operators. For a system in a mixed state, described by such a density operator $\hat{\rho}$, the *expected outcome* of the measurement $\hat{A}$ on the system was found to be $\mathrm{E}_{\hat{\rho}}(\hat{A}) = \mathrm{Tr}(\hat{\rho}\hat{A})$; see Eq. (6). This formalism doesn't come, however, with a clear view on how to interpret the probabilities and expected values that play a role in it.

As it's one of our aims in this paper to provide a decision-theoretic interpretation for such density operators and expected outcomes, by relating them to our coherent (lower) previsions and sets of desirable measurements, we'll now show that the coherent previsions on $\mathscr{H}$ provide us with an appropriate means of connecting them with our decision-theoretic set of desirable measurements approach. To do so, we'll again follow the path that was cleared by Benavoli, Facchini and Zaffalon in Ref. [7, Sec. IV], by focusing on the mathematical links between coherent (lower) previsions and density operators. We postpone a discussion of how to interpret these links to Section 6.3.

We begin this brief mathematical exploration with the observation that expected outcomes make for perfectly acceptable coherent previsions.

Proposition 9
For every density operator $\hat{\rho} \in \mathscr{R}$, the real functional $\Lambda_{\hat{\rho}}(\bullet) := \mathrm{E}_{\hat{\rho}}(\bullet) := \mathrm{Tr}(\hat{\rho}\,\bullet)$ is a coherent prevision on $\mathscr{H}$.

Proof. The additivity condition CP2 follows at once from the linearity of the trace, so it only remains to prove the bounds condition CP1. Since $\hat{\rho}$ is a density operator, Proposition 3 tells us that there are kets $|\psi_1\rangle, \dots, |\psi_q\rangle$ in $\mathscr{X}$ and real numbers $\alpha_1, \alpha_2, \dots, \alpha_q \in [0,1]$ such that $\sum_{k=1}^q \alpha_k = 1$ and $\hat{\rho} = \sum_{k=1}^q \alpha_k |\psi_k\rangle\langle\psi_k|$. For every $\hat{A} \in \mathscr{H}$, we then find that, indeed,

$$\mathrm{Tr}(\hat{\rho}\hat{A}) = \sum_{k=1}^q \alpha_k \mathrm{Tr}(|\psi_k\rangle\langle\psi_k|\hat{A}) = \sum_{k=1}^q \alpha_k \langle\psi_k|\hat{A}|\psi_k\rangle \geq \min\mathrm{spec}(\hat{A}),$$

where the first equality follows from the linearity of the trace, the second equality from its cyclic character, and the inequality from Lemma 6 and the fact that $\sum_{k=1}^q \alpha_k = 1$. □

To prove a converse result, namely, that for every coherent prevision Λ there's a (unique) density operator $\hat{\rho}$ such that $\Lambda(\bullet) = \mathrm{Tr}(\hat{\rho}\,\bullet)$, we need a bit of mathematical background on Hilbert spaces. We've already had occasion to mention that we can provide the real linear space $\mathscr{H}$ with the Frobenius inner product, defined by $(\hat{A},\hat{B}) := \mathrm{Tr}(\hat{A}\hat{B})$ for all $\hat{A},\hat{B} \in \mathscr{H}$, which turns $\mathscr{H}$ into a real Hilbert space. We can then apply to this Hilbert space $\mathscr{H}$ the general Riesz Representation Theorem, which states that continuous — or equivalently, bounded — linear functionals on a Hilbert space are in a one-to-one relationship with its vectors.

Lemma 10 (Riesz Representation Theorem [40, Thm. II.4])
For any continuous linear functional Γ on $\mathscr{H}$, there's a unique $\hat{B}_\Gamma \in \mathscr{H}$ *such that* $\Gamma(\hat{A}) = \mathrm{Tr}(\hat{B}_\Gamma \hat{A})$ *for all* $\hat{A} \in \mathscr{H}$.

We now have the tools to prove the desired one-to-one relationship; see also the discussion in Section IV of [7], where Proposition IV.6 and Theorem IV.4 turn out to be closely related to what we have here.

Proposition 11
There's a one-to-one correspondence $\hat{R}$ between coherent previsions Λ on $\mathscr{H}$ and density operators $\hat{\rho}$ on $\mathscr{X}$, through $\Lambda(\bullet) = \mathrm{Tr}(\hat{\rho}\,\bullet)$ with $\hat{\rho} = \hat{R}(\Lambda)$. This bijection preserves convex combinations, in the sense that $\hat{R}(\alpha\Lambda_1 + (1-\alpha)\Lambda_2) = \alpha\hat{R}(\Lambda_1) + (1-\alpha)\hat{R}(\Lambda_2)$, for all real $\alpha \in [0,1]$ and all $\Lambda_1, \Lambda_2 \in \mathbb{P}$.

Proof. We begin with the first statement. By Proposition 9, it suffices to prove that with any coherent prevision $\Lambda \in \mathbb{P}$ there corresponds a unique density operator $\hat{\rho}$ such that $\Lambda(\bullet) = \mathrm{Tr}(\hat{\rho}\,\bullet)$. We can then denote this $\hat{\rho}$ by $\hat{R}(\Lambda)$. We know that Λ is bounded and therefore continuous, from the discussion leading up to Eq. (29), so we can infer from Lemma 10 that there's a unique $\hat{B}_\Lambda \in \mathscr{H}$ such that $\Lambda(\bullet) = \mathrm{Tr}(\hat{B}_\Lambda \bullet)$. It's enough, therefore, to prove that the Hermitian operator $\hat{B}_\Lambda$ is a density operator, or in other words, that $\hat{B}_\Lambda \geq 0$ and that $\mathrm{Tr}(\hat{B}_\Lambda) = 1$.

First, observe that $\mathrm{Tr}\big(\hat{B}_\Lambda\big) = \mathrm{Tr}\big(\hat{B}_\Lambda \hat{I}\big) = \Lambda(\hat{I}) = 1$, where the last equality follows from L3.

Next, consider any $|\psi\rangle \in \bar{\mathscr{X}}$. Then for the projection $\hat{P}_{|\psi\rangle} := |\psi\rangle\langle\psi| \in \mathscr{H}$, we find that $\mathrm{spec}(\hat{P}_{|\psi\rangle}) \subseteq \{0,1\}$ and therefore, by the bounds condition CP1, that $\Lambda(\hat{P}_{|\psi\rangle}) \geq 0$. As the trace is cyclic, this leads to $0 \leq \Lambda\big(\hat{P}_{|\psi\rangle}\big) = \mathrm{Tr}\big(\hat{B}_\Lambda \hat{P}_{|\psi\rangle}\big) = \mathrm{Tr}\big(\hat{B}_\Lambda |\psi\rangle\langle\psi|\big) = \langle\psi|\hat{B}_\Lambda|\psi\rangle$, which implies that, indeed, $\hat{B}_\Lambda \geq 0$.

To prove that the second statement, consider any real $\alpha \in [0,1]$ and any $\Lambda_1, \Lambda_2 \in \mathbb{P}$, and let $\hat{\rho}_1 := \hat{R}(\Lambda_1)$ and $\hat{\rho}_2 := \hat{R}(\Lambda_2)$. And, by the linearity of the trace,

$$\begin{aligned}\mathrm{Tr}\big((\alpha\hat{\rho}_1 + (1-\alpha)\hat{\rho}_2)\bullet\big) &= \alpha\,\mathrm{Tr}\big(\hat{\rho}_1\bullet\big) + (1-\alpha)\,\mathrm{Tr}\big(\hat{\rho}_2\bullet\big)\\ &= \alpha\Lambda_1(\bullet) + (1-\alpha)\Lambda_2(\bullet) = \big(\alpha\Lambda_1 + (1-\alpha)\Lambda_2\big)(\bullet). \qquad \square\end{aligned}$$

More generally, when Your beliefs about $|\Psi\rangle$ are modelled by a coherent lower prevision $\underline{\Lambda}$ that isn't self-conjugate, then we've already established that this $\underline{\Lambda}$ can be written as the lower envelope of the convex closed set of coherent previsions, or credal set, $\mathbb{P}(\underline{\Lambda})$. Because it preserves convex combinations, the map $\hat{R}$ turns this set into the *convex and closed*[21] *set of density operators*

$$\begin{aligned}\mathscr{R}(\underline{\Lambda}) &:= \hat{R}(\mathbb{P}(\underline{\Lambda})) = \{\hat{R}(\Lambda) : \Lambda \in \mathbb{P}(\underline{\Lambda})\}\\ &= \{\hat{\rho} \in \mathscr{R} : (\forall \hat{A} \in \mathscr{H})\,\mathrm{Tr}(\hat{\rho}\hat{A}) \geq \underline{\Lambda}(\hat{A})\},\end{aligned} \tag{35}$$

and, clearly, this set of density operators completely determines the coherent lower prevision $\underline{\Lambda}$ in the sense that

$$\underline{\Lambda}(\hat{A}) = \min\{\mathrm{Tr}(\hat{\rho}\hat{A}) : \hat{\rho} \in \mathscr{R}(\underline{\Lambda})\} \text{ for all } \hat{A} \in \mathscr{H}. \tag{36}$$

[21]... in the topology on $\mathscr{R}$ induced by the map $\hat{R}$, characterised by the fact that a sequence of densities $\hat{\rho}_n$ converges to the density $\hat{\rho}$ if and only if corresponding sequence of linear functionals $\mathrm{Tr}(\hat{\rho}_n\bullet)$ converges (pointwise, or equivalently, in norm, due to the finite dimension of the underlying Hilbert space $\mathscr{X}$) to the linear functional $\mathrm{Tr}(\hat{\rho}\,\bullet)$.

This is the generalisation of Born's Rule, in its most general version (6), that we derive from our decision-theoretic argumentation, where the single density operator $\hat{\rho}$ is replaced by a convex closed set $\mathscr{R}(\underline{\Lambda})$ of them, and where the expected outcome $\mathrm{E}_{\hat{\rho}}(\hat{A})$ of a measurement $\hat{A}$ is replaced its lower prevision $\underline{\Lambda}(\hat{A})$.

One meaning that our decision-theoretic approach attributes to this lower prevision $\underline{\Lambda}$ is, on the one hand, that it characterises (up to border behaviour, see the previous section) Your preferences between the acts $\mathrm{act}_{\hat{A}}$ associated with measurements $\hat{A}$. On the other hand, we also recall that $\underline{\Lambda}(\hat{A})$ was defined as Your supremum price for buying the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$ associated with performing the measurement $\hat{A}$ on the quantum system in its unknown state $|\Psi\rangle$. And finally, if we were to allow ourselves to follow a more standard approach to probability in quantum mechanics, $\underline{\Lambda}(\hat{A})$ could be seen as a tight lower bound on the expected values of the measurement $\hat{A}$, as is indicated by Eq. (36).

Qubit running example: instalment 8
Let's now find out how to apply these ideas in our running example.

First, let's consider a density operator $\hat{\rho}$. Since $\hat{\rho}$ is in particular Hermitian, we know that there are $a,b,c,d \in \mathbb{R}$ such that $\hat{\rho} = d\hat{I} + a\hat{\sigma}_x + b\hat{\sigma}_y + c\hat{\sigma}_z$; its eigenvalues are then $d \pm \sqrt{a^2+b^2+c^2}$, due to Eq. (9). Expressing that $\mathrm{Tr}(\hat{\rho}) = 1$ and $\hat{\rho} \geq 0$ therefore leads to the conditions:

$$d = \frac{1}{2} \text{ and } a^2+b^2+c^2 \leq \frac{1}{4}. \tag{37}$$

Now consider any measurement $\hat{C} \in \mathscr{H}$ and let, as before, $\hat{C} = w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z$, with $x,y,z,w \in \mathbb{R}$. Another application of the argumentation behind Eq. (9) tells us that $\mathrm{Tr}(\hat{\rho}\hat{C})$ is twice the $\hat{I}$-component of $\hat{\rho}\hat{C}$. Moreover, by the properties of Pauli operators [9, p. 418], we find that

$$\begin{aligned}\hat{\rho}\hat{C} &= \Big(\frac{1}{2}\hat{I} + a\hat{\sigma}_x + b\hat{\sigma}_y + c\hat{\sigma}_z\Big)(w\hat{I} + x\hat{\sigma}_x + y\hat{\sigma}_y + z\hat{\sigma}_z)\\ &= \Big(\frac{1}{2}w + ax + by + cz\Big)\hat{I} + (\dots)\hat{\sigma}_x + (\dots)\hat{\sigma}_y + (\dots)\hat{\sigma}_z\end{aligned}$$

and therefore $\mathrm{Tr}(\hat{\rho}\hat{C}) = w + 2(ax+by+cz)$. If we now consider the coherent prevision Λ that corresponds to $\hat{\rho}$ in the sense that $\Lambda(\bullet) = \mathrm{Tr}(\hat{\rho}\,\bullet)$, and invoke Eq. (32), we find that

$$w + 2(ax+by+cz) = w + x\Lambda_x + y\Lambda_y + z\Lambda_z \text{ for all } w,x,y,z \in \mathbb{R}.$$

This tells us that $\Lambda_x = 2a$, $\Lambda_y = 2b$ and $\Lambda_z = 2c$, and therefore

$$\hat{\rho} = \frac{1}{2}\hat{I} + \frac{\Lambda_x\hat{\sigma}_x + \Lambda_y\hat{\sigma}_y + \Lambda_z\hat{\sigma}_z}{2}, \tag{38}$$

while the constraints (37) turn into $\Lambda_x^2 + \Lambda_y^2 + \Lambda_z^2 \leq 1$, which is in complete accordance with Eq. (33).

The elements $\hat{\rho}$ of the closed convex set $\mathscr{R}(\underline{\Lambda}_{\mathscr{C}(\mathscr{A})})$ are therefore characterised by

$$\hat{\rho} = \frac{1}{2}\hat{I} + \frac{\Lambda_x\hat{\sigma}_x + \Lambda_y\hat{\sigma}_y + \Lambda_z\hat{\sigma}_z}{2} \text{ with } \Lambda_x \geq 0 \text{ and } \Lambda_x^2 + \Lambda_y^2 + \Lambda_z^2 \leq 1.$$

6.3. **A few words of caution ... and a curious observation.** It's important to stress again that the version of Born's rule (36) that we've thus derived from our assumptions, doesn't start from a probability argument, but rather from a preference argument, which then leads to coherent (lower) previsions. We feel this distinction to be quite important, and we'll use this and the next section to clarify our point.

The lower previsions $\underline{\Lambda}$ that we've introduced here, are functionals that serve as an alternative characterisation for Your preferences, in the sense that, up to boundary behaviour, $\hat{A} \rhd \hat{B}$ if and only if $\underline{\Lambda}(\hat{A} - \hat{B}) > 0$ or $\hat{A} \gneq \hat{B}$; see Eq. (28) and the discussion leading to it.

We'll see further on that this will definitely be the case if a coherent set of desirable measurement is *maximal*, that is, not included in any other such set, for which an equivalent condition is that

M. $\{\hat{A}, -\hat{A}\} \cap \mathscr{D} \neq \emptyset$ for all non-null $\hat{A} \in \mathscr{H}$. [maximality]

Proof of the equivalence. Assume that the coherent set of desirable measurements $\mathscr{D}$ is maximal and consider any non-null $\hat{A} \in \mathscr{H}$ such that $-\hat{A} \notin \mathscr{D}$, then we need to prove that $\hat{A} \in \mathscr{D}$. We're done if we can show that $\mathscr{D} \cup \{\hat{A}\}$ is consistent, because then $\mathscr{C}(\mathscr{D} \cup \{\hat{A}\})$ will be coherent and therefore equal to $\mathscr{D}$ by its maximality; and clearly $\hat{A} \in \mathscr{C}(\mathscr{D} \cup \{\hat{A}\})$. Suppose towards contradiction that $\mathscr{D} \cup \{\hat{A}\}$ isn't consistent, then we infer from Eq. (17) that $\hat{0} \in \operatorname{posi}(\mathscr{D} \cup \{\hat{A}\}) + \mathscr{H}_{\geq \hat{0}}$. Taking into account that

$$\begin{aligned}
\operatorname{posi}(\mathscr{D} \cup \{\hat{A}\}) + \mathscr{H}_{\geq \hat{0}} &= \Big(\operatorname{posi}(\mathscr{D}) \cup \operatorname{posi}(\hat{A}) \cup \big(\operatorname{posi}(\mathscr{D}) + \operatorname{posi}(\{\hat{A}\})\big)\Big) + \mathscr{H}_{\geq \hat{0}} \\
&= \big(\mathscr{D} \cup \operatorname{posi}(\{\hat{A}\}) \cup (\mathscr{D} + \operatorname{posi}(\{\hat{A}\}))\big) + \mathscr{H}_{\geq \hat{0}} \\
&= (\mathscr{D} + \mathscr{H}_{\geq \hat{0}}) \cup (\operatorname{posi}(\{\hat{A}\}) + \mathscr{H}_{\geq \hat{0}}) \cup (\mathscr{D} + \operatorname{posi}(\{\hat{A}\}) + \mathscr{H}_{\geq \hat{0}}) \\
&= \mathscr{D} \cup (\operatorname{posi}(\{\hat{A}\}) + \mathscr{H}_{\geq \hat{0}}) \cup (\mathscr{D} + \operatorname{posi}(\{\hat{A}\})),
\end{aligned}$$

where the second equality follows from $\mathscr{D} = \operatorname{posi}(\mathscr{D})$ [use the coherence of $\mathscr{D}$ and D2 and D3] and the last equality follows from $\mathscr{D} + \mathscr{H}_{\geq \hat{0}} = \mathscr{D}$ [use the coherence of $\mathscr{D}$ and D6], we then find that, necessarily, $-\hat{A} \in \mathscr{H}_{\gneq \hat{0}}$ [because $\hat{A} \neq \hat{0}$] or $-\hat{A} \in \mathscr{D}$, contradicting our assumption that $-\hat{A} \notin \mathscr{D}$ [use the coherence of $\mathscr{D}$ and D4].

Conversely, assume that $\mathscr{D}$ satisfies the condition M, then we must show that $\mathscr{D}$ is maximal. Let $\mathscr{D}'$ be any coherent set of desirable measurements that includes $\mathscr{D}$, assume towards contradiction that the inclusion is strict and consider any $\hat{A} \in \mathscr{D}' \setminus \mathscr{D}$. But then $-\hat{A} \in \mathscr{D}$, and therefore also $-\hat{A} \in \mathscr{D}'$. Since, by assumption, also $\hat{A} \in \mathscr{D}'$, we infer from the coherence of $\mathscr{D}'$ [namely, D2] that $\hat{0} = -\hat{A} + \hat{A} \in \mathscr{D}'$, contradicting the coherence of $\mathscr{D}'$ [namely, D1]. □

Indeed, the self-conjugacy of $\underline{\Lambda}_{\mathscr{D}}$ is equivalent to $\mathscr{D}$ satisfying the so-called *weak maximality* condition, which is formally similar to, but somewhat weaker than, the condition M for maximality; see also Refs. [13, 19] for related discussion.

WM. $\{\hat{A} + \varepsilon\hat{I}, -\hat{A} + \varepsilon\hat{I}\} \cap \mathscr{D} \neq \emptyset$ for all $\hat{A} \in \mathscr{H}$ and all real $\varepsilon > 0$. [weak maximality]

Weak maximality means that, if we include Your marginal preferences to come to a notion of weak preference, You always have a(t least a weak) preference between any $\hat{A}$ and its additive inverse $-\hat{A}$. This statement is clarified and made more precise by the following proposition, which established a direct connection between weakly maximal coherent sets of desirable measurements and coherent previsions.

Proposition 12

A coherent set of desirable measurements $\mathscr{D}$ is weakly maximal if and only if the corresponding lower prevision $\underline{\Lambda}_{\mathscr{D}}$ is a coherent prevision, so $\underline{\Lambda}_{\mathscr{D}} = \overline{\Lambda}_{\mathscr{D}} = \Lambda_{\mathscr{D}}$. In that case, the corresponding credal set $\mathbb{P}(\Lambda_{\mathscr{D}})$ is the singleton $\{\Lambda_{\mathscr{D}}\}$.

Proof. First, assume that $\mathscr{D}$ is weakly maximal. Consider any $\hat{A} \in \mathscr{H}$ and any real $\varepsilon > 0$, then it follows from Eq. (21) that $(\overline{\Lambda}_{\mathscr{D}}(\hat{A}) - \varepsilon) - \hat{A} \notin \mathscr{D}$. So it follows from WM that $\hat{A} - (\overline{\Lambda}_{\mathscr{D}}(\hat{A}) - \varepsilon) \in \mathscr{D}$, and therefore Eq. (20) implies that $\overline{\Lambda}_{\mathscr{D}}(\hat{A}) - \varepsilon \leq \underline{\Lambda}_{\mathscr{D}}(\hat{A})$, implying that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}) \geq \overline{\Lambda}_{\mathscr{D}}(\hat{A})$ and therefore also that $\underline{\Lambda}_{\mathscr{D}}(\hat{A}) = \overline{\Lambda}_{\mathscr{D}}(\hat{A})$, taking into account LP4.

Conversely, assume that $\overline{\Lambda}_{\mathscr{D}} = \underline{\Lambda}_{\mathscr{D}}$ is a coherent prevision $\Lambda_{\mathscr{D}}$. Assume towards contradiction that there are $\hat{A} \in \mathscr{H}$ and real $\varepsilon > 0$ such that both $\hat{A} + \varepsilon\hat{I} \notin \mathscr{D}$ and $-\hat{A} + \varepsilon\hat{I} \notin \mathscr{D}$. Hence, both $\Lambda_{\mathscr{D}}(\hat{A} + \varepsilon\hat{I}) = \underline{\Lambda}_{\mathscr{D}}(\hat{A} + \varepsilon\hat{I}) \leq 0$ and $\Lambda_{\mathscr{D}}(-\hat{A} + \varepsilon\hat{I}) = \underline{\Lambda}_{\mathscr{D}}(-\hat{A} + \varepsilon\hat{I}) \leq 0$, by Lemma 8. Using L1 and L3, this implies that both $\Lambda_{\mathscr{D}}(\hat{A}) \leq -\varepsilon$ and $-\Lambda_{\mathscr{D}}(\hat{A}) \leq -\varepsilon$, a contradiction.

The rest of the proof is now immediate. □

In the special, so-called *precise*, case that is characterised by Proposition 12, we find that the uncertainty is described similarly as on a more standard account of uncertainty in quantum mechanics, because we've seen in Proposition 11 that the coherent previsions Λ on $\mathscr{H}$ are in a one-to-one correspondence with the density operators $\hat{\rho}$ in $\mathscr{R}$. In fact, we then retrieve a version of Born's rule in the presence of *epistemic uncertainty* about the state of the system, as Proposition 11 tells us that

$$\Lambda(\hat{A}) = \mathrm{Tr}(\hat{\rho}\hat{A}) = \mathrm{E}_{\hat{\rho}}(\hat{A}) \text{ for all } \hat{A} \in \mathscr{H},$$

where the last equality follows from Eq. (6). We see that when we let $\hat{\rho} := \hat{R}(\Lambda)$, Your fair price $\Lambda(\hat{A}) = \mathrm{Tr}(\hat{\rho}\hat{A})$ for the measurement $\hat{A}$ can be made to correspond to the expected value $\mathrm{E}_{\hat{\rho}}(\hat{A})$ of the outcome of a measurement as considered in the standard approach to quantum mechanics, and discussed in Section 2.2.

But, we have to be cautious in making the connection between our approach and the more standard one. In the way we've set up the argument in this paper, the decision problem comes first — is *primary* — and the linear price functional $\Lambda_{\mathscr{D}}$ is a mathematical tool that, under certain rather restrictive conditions, allows us to characterise Your preferences as captured in $\mathscr{D}$, in the sense that

$$\hat{A} \in \mathscr{D} \Leftrightarrow \big(\Lambda_{\mathscr{D}}(\hat{A}) > 0 \text{ or } \hat{A} \gneq \hat{0}\big).$$

It turns out that this $\Lambda_{\mathscr{D}}$ can be characterised mathematically by a density operator $\hat{\rho}_{\mathscr{D}} = \hat{R}(\Lambda_{\mathscr{D}})$, which, through Proposition 3, is often given a probabilistic interpretation.[22] But this probabilistic interpretation, and the corresponding probabilistic interpretation of the trace $\mathrm{Tr}(\hat{\rho}_{\mathscr{D}}\hat{A})$ as the expected value of the outcome of the measurement $\hat{A}$, is only *secondary*, or *derivative*. Indeed, what we have shown above is that $\Lambda(\hat{A}) = \mathrm{Tr}(\hat{\rho}\hat{A})$ is Your fair price for the uncertain reward $u_{\hat{A}}(|\Psi\rangle)$, which *isn't necessarily the same thing as Your expected value for the outcome of the measurement* $\hat{A}$.

6.4. **An interesting special case.** To shed still more light on this issue, let's now consider the special case that You know with certainty that the state is some specific $|\phi\rangle$ in $\mathscr{X}$ — so You know that $|\Psi\rangle = |\phi\rangle$.[23] This could for instance be the case after You've just performed a measurement $\hat{B}$ on the quantum system and observed an outcome — eigenvalue — with a one-dimensional eigenspace; see QM5. Through what coherent set of desirable measurements $\mathscr{D}$ can this knowledge be represented?

As we mentioned before right after announcing DTB1 in Section 3.2, there's now no longer any uncertainty about the reward You'll get from performing any measurement $\hat{A}$: it's the real number $u_{\hat{A}}(|\phi\rangle) = \langle\phi|\hat{A}|\phi\rangle$, and You'll clearly strictly prefer that to the status quo provided that $\langle\phi|\hat{A}|\phi\rangle > 0$. This leads to the assessment $\mathscr{A} := \{\hat{A} \in \mathscr{H} \colon \langle\phi|\hat{A}|\phi\rangle > 0\}$. This assessment is consistent [see Eq. (17)], and [using Eq. (16)] we find that the smallest coherent set of desirable measurements that includes it is given by

$$\mathscr{D}_{|\phi\rangle} := \{\hat{A} \in \mathscr{H} \colon \langle\phi|\hat{A}|\phi\rangle > 0 \text{ or } \hat{A} \gneq 0\};$$

which is, by the way, weakly maximal. The corresponding *coherent prevision* $\Lambda_{|\phi\rangle} := \Lambda_{\mathscr{D}_{|\phi\rangle}}$ is then given by

$$\begin{aligned}
\Lambda_{|\phi\rangle}(\hat{A}) :=& \sup\{\alpha \in \mathbb{R} \colon \hat{A} - \alpha\hat{I} \in \mathscr{D}_{|\phi\rangle}\} = \sup\{\alpha \in \mathbb{R} \colon \langle\phi|A|\phi\rangle > \alpha \text{ or } \hat{A} \gneq \alpha\hat{I}\} \\
=& \sup\{\alpha \in \mathbb{R} \colon \langle\phi|A|\phi\rangle > \alpha \text{ or } (\min\operatorname{spec}(\hat{A}) \geq \alpha \text{ and } \hat{A} \neq \alpha\hat{I})\} \\
=& \langle\phi|A|\phi\rangle = u_{\hat{A}}(|\phi\rangle), \text{ for all } \hat{A} \in \mathscr{H},
\end{aligned}$$

[22] This interpretation is not without its problems, because the 'decomposition' in Proposition 3 typically isn't unique.

[23] Recall that the state of the system is actually identified by a ray in Hilbert space $\mathscr{X}$, of which $|\phi\rangle$ is only one of the elements, which completely determines it. By 'You know that $|\Psi\rangle = |\phi\rangle$', we mean that You know that the system is in the state $|\phi\rangle$, up to a phase factor.

where the penultimate equality follows from the fact that $\langle\phi|A|\phi\rangle \geq \min\operatorname{spec}(\hat{A})$ [see Eq. (14)]. In other words, if You know that $|\Psi\rangle = |\phi\rangle$, then $u_{\hat{A}}(|\phi\rangle) = \langle\phi|A|\phi\rangle$ is necessarily *Your fair price for performing the measurement* $\hat{A}$. Again, this isn't necessarily the same thing as Your expected value for the outcome of the measurement $\hat{A}$.

Stating that Your fair price $u_{\hat{A}}(|\phi\rangle) = \langle\phi|A|\phi\rangle$ for performing the measurement $\hat{A}$ is the same thing as Your expected value for the outcome of the measurement $\hat{A}$, is tantamount to making an *extra assumption*, namely that there are probabilities for each of the possible outcomes $\lambda \in \operatorname{spec}(\hat{A})$ of the measurement $\hat{A}$ and that these probabilities are given by Eq. (2). Making this extra assumption therefore amounts to accepting QM7, which is something we've been wanting to avoid all along in this paper.

We'll now explain why it might make sense to make this extra assumption, all the while realising that it *is* indeed an additional step to take. In DTB1 we assumed the existence of a reward function $w_{\hat{A}}$ such that $w_{\hat{A}}(|\phi\rangle)$ is the reward for performing measurement $\hat{A}$ when the system is in the state $|\phi\rangle$ and there intended to be a one-number summary of the possible outcomes $\lambda \in \operatorname{spec}(\hat{A})$ that the measurement $\hat{A}$ may yield in the state $|\phi\rangle$. The exact form of the reward function $w_{\hat{A}}$ was then fixed by the later postulates RF1–RF4 to be $w_{\hat{A}} = u_{\hat{A}} = \langle\bullet|\hat{A}|\bullet\rangle$. Interestingly, these postulates make sure that the one-number summary $u_{\hat{A}}(|\phi\rangle)$ behaves linearly in $\hat{A}$ and therefore *acts as if it were an expectation*:

$$u_{\hat{A}}(|\phi\rangle) = \langle\phi|\hat{A}|\phi\rangle = \left\langle \phi \middle| \sum_{\lambda\in\operatorname{spec}(\hat{A})} \lambda \hat{P}_\lambda \middle| \phi \right\rangle = \sum_{\lambda\in\operatorname{spec}(\hat{A})} \lambda \langle\phi|\hat{P}_\lambda|\phi\rangle = \sum_{\lambda\in\operatorname{spec}(\hat{A})} \lambda u_{\hat{P}_\lambda}(|\phi\rangle),$$

where the $u_{\hat{P}_\lambda}(|\phi\rangle) = \langle\phi|\hat{P}_\lambda|\phi\rangle$ then act as if they were the expected outcomes of the measurements $\hat{P}_\lambda$. Since the only possible outcomes of a projection $\hat{P}_\lambda$ are 0 and 1, the one-number summary $u_{\hat{P}_\lambda}(|\phi\rangle)$ then also acts as if it were the probability of its outcome being 1, or in other words, the probability of the measurement $\hat{A}$ producing the outcome λ, in accordance with Born's rule QM7.

This special case of our decision-theoretic argument brings us quite close to a result by Deutsch, which has a similar interpretation, but follows a different argumentation, as we'll see next.

7. DEUTSCH'S DECISION-THEORETIC APPROACH

In an important and seminal paper [23], David Deutsch presented a different, and earlier, approach to what we're attempting here.

Deutsch aimed at deriving the probabilistic postulate QM7 of quantum mechanics — or Born's rule in its simplest version (2) — from the non-probabilistic ones and straightforward 'non-probabilistic' decision-making principles. While our argumentation in this paper is inspired by Deutsch's take on this issue, it's also rather different in quite a number of respects, in that we're not necessarily convinced that the 'non-probabilistic' decision-making principles he suggested are all that straightforward. This is why we made quite some effort in this paper to clearly state exactly what decision-making approach we follow, which assumptions we make and which we don't make. We believe our derivation can be defended on any interpretation of quantum mechanics that accepts the postulates QM1–QM4 (and QM5) and the reward function postulates RF1–RF4, which make sense against the decision-making background (DTB1 and DTB2) sketched in Section 3.2.

To allow readers to better see the similarities and differences between Deutsch's argument and ours, we now give a brief account of it, which should also be compared to and contrasted with the approach we've set up in Sections 3 and 4.

Deutsch considers a quantum system whose states belong to some complex Hilbert space $\mathscr{X}$. He focuses on a measurement $\hat{A} \in \mathscr{H}$ on this quantum system, which can be in any of several possible states $|\psi\rangle \in \bar{\mathscr{X}}$ right before the measurement is carried out.

His decision-making set-up considers a subject, whom we'll also call *You* and who can choose between playing several games of the type

$$\mathrm{game}^{\hat{A}}_{|\psi\rangle} := \text{“perform the measurement } \hat{A} \text{ on the quantum system in state } |\psi\rangle\text{”},$$

one for each $|\psi\rangle \in \bar{\mathscr{X}}$ and $\hat{A} \in \mathscr{H}$. The outcome of $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ is unknown to You before playing the game — doing the measurement — but, after the measurement, You'll receive as a pay-off the actual outcome of the measurement, expressed in units of some linear utility scale. While the latter is an assumption we also make in our set-up in Section 3.1, our general approach is different in that Deutsch focuses on the case that You know what the system's state is, while we allow You to be uncertain about it from the outset.

The *first assumption* that Deutsch makes,[24] is that each game $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ has some real *value* $\mathscr{V}_{\hat{A}}(|\psi\rangle)$, intended as a one-number summary of the possible pay-offs that the game $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ may yield. He seems to be inspired by de Finetti's [15] definition of a prevision for a random quantity as a fair price for it when he defines this value as

> the utility of a hypothetical pay-off such that [You are] indifferent to playing the game and receiving that pay-off unconditionally.

Since it follows from the non-probabilistic postulates of quantum mechanics [in particular QM4] that performing the measurement $\alpha\hat{I}$ yields the real outcome α unconditionally — meaning independently of $|\psi\rangle$ —, this implies in effect that $\mathscr{V}_{\alpha\hat{I}} = \alpha$. Hence, there is — You are assumed to have — some *valuation* $\mathscr{V}\colon \mathscr{H} \times \bar{\mathscr{X}} \to \mathbb{R}\colon (\hat{A}, |\phi\rangle) \mapsto \mathscr{V}_{\hat{A}}(|\phi\rangle)$ that assigns a *value* $\mathscr{V}_{\hat{A}}(|\psi\rangle)$ to each $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ in such a way that

$$\mathscr{V}_{\hat{A}}(|\psi\rangle) = \mathscr{V}_{\alpha\hat{I}}(|\psi\rangle) \Leftrightarrow \mathrm{game}^{\hat{A}}_{|\psi\rangle} \equiv \mathrm{game}^{\hat{\alpha}}_{|\psi\rangle}, \tag{39}$$

where we define the statement '$\mathrm{game}^{\hat{A}}_{|\psi\rangle} \equiv \mathrm{game}^{\hat{\alpha}}_{|\psi\rangle}$' to mean that You're *indifferent* between the game $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ and the game $\mathrm{game}^{\hat{\alpha}}_{|\psi\rangle}$ with unconditional pay-off α. That there should be such a real number for all such $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$, is an assumption that is very close to our decision-theoretic background assumption DTB1, where we assume that there's some reward function[25] $w_{\hat{A}}\colon \bar{\mathscr{X}} \to \mathbb{R}$ such that $w_{\hat{A}}(|\psi\rangle)$ is the reward, expressed in utiles, for $\mathrm{act}_{\hat{A}}$ when the quantum system under consideration system is in state $|\psi\rangle$, intended as a one-number summary of the possible pay-offs that the measurement $\hat{A}$ may yield on a system in state $|\psi\rangle$.

Deutsch's further assumptions deal with the nature of Your 'value of a game', and he argues that they're strong enough to fix its form. The *third assumption* is that You adhere to the *principle of substitutability* in composite games, which are games that involve subgames. It means that if any of the subgames is replaced by some game of equal value, the value of the composite game remains the same to You.

The *fourth and final assumption* that Deutsch makes, is what he calls the *zero-sum rule*, and it concerns Your attitude towards games where You can take one of two possible roles — say, R and S — with the following property: whenever You were to receive a pay-off x in role R, You would receive $-x$ in role S. The rule then states that if $\mathscr{V}_R$ and $\mathscr{V}_S$ are Your respective values for playing the game in roles R and S, then it must be that $\mathscr{V}_R + \mathscr{V}_S = 0$.

Using these decision-theoretic assumptions, together with the non-probabilistic quantum mechanical postulates, Deutsch argues that Your valuation must then necessarily be given by

$$\mathscr{V}_{\hat{A}}(|\phi\rangle) = \langle\phi|\hat{A}|\phi\rangle \text{ for all } \hat{A} \in \mathscr{H} \text{ and } |\phi\rangle \in \bar{\mathscr{X}}. \tag{40}$$

Recall that, according to Deutsch's argumentation, this $\mathscr{V}_{\hat{A}}(|\psi\rangle$ is the fair price You're willing to pay for playing the game $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$, or in other words, for performing the measurement $\hat{A}$ when the quantum system's state is $|\psi\rangle$. It's in this specific sense that Deutsch's

[24]The order in which we list these assumptions doesn't follow the order in which Deutsch presents them, but rather reflects our attempt at structuring his argumentation to allow for a comparison with our approach.

[25]Recall that, initially, we use the notation $w_{\hat{A}}$ for any such reward function, and we reserve the notation $u_{\hat{A}}$ for the unique reward function that turns out to satisfy our reward function postulates.

argument leads to a justification of, and interpretation for, Born's rule as put forth in Postulate QM7, and in particular Eq. (3). This is comparable to our justification for it using the postulates RF1–RF4 and the background assumptions DTB1 and DTB2, and caveats apply here that are similar to the ones we formulated for our derivation near the end of Sections 6.3 and 6.4.

On Deutsch's approach, the system state $|\psi\rangle$ is assumed to be known, or given, and the value $\mathscr{V}_{\hat{A}}(|\psi\rangle)$ of $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ is Your fair price for performing the measurement $\hat{A}$ upon the system in that state $|\psi\rangle$, in the sense that You deem performing the measurement and getting as a reward its outcome in units of a linear utility to be equivalent to getting the fixed amount $\mathscr{V}_{\hat{A}}(|\psi\rangle)$ of that linear utility. His way of dealing with the problem is, like ours, very much de Finetti-like, in that he starts with a pre-determined linear utility and implicitly assumes that all games $\mathrm{game}^{\hat{A}}_{|\psi\rangle}$ with $\hat{A}\in\mathscr{H}$, for a fixed $|\psi\rangle\in\bar{\mathscr{X}}$ can be compared to one another through the intervention of the values that their valuation $\mathscr{V}_{\hat{A}}(|\psi\rangle)$ assumes on the real line;[26] see also our discussion on decision-theoretic foundations in Section 3.2.

It should be clear that Deutsch's decision problem, and the context he places it in, is in its conception a bit different from ours. We start from the assumption that You are (or at least may be) uncertain about the state $|\Psi\rangle$ the system is in, and that You express Your beliefs about what that state is by expressing preferences between the uncertain rewards $w_{\hat{A}}(|\Psi\rangle)$ for different measurements $\hat{A}$. In addition, our postulates, as announced in Section 3.3, are quite different from Deutsch's. But, as we point out in Section 6.4, in the special case that You know that the state $|\Psi\rangle$ of the system is actually $|\psi\rangle$, we're still led to the same conclusion as Deutsch: You must have a *fair price* for the uncertain reward associated with performing the measurement $\hat{A}$ in that — now certain — state, *and* that fair price must be given by $u_{\hat{A}}(|\psi\rangle)=\langle\psi|A|\psi\rangle$.

To be sure, while Deutsch's discussion in Ref. [23] deals directly only with the case that You know the system to be in some pure state $|\psi\rangle$, he does consider that You might be uncertain about the quantum state, by mentioning the possibility that the system might be in a mixed state. Even if he leaves the discussion of mixed states implicit, he does hint, at the end of Section 3 there, at ways to deal with them:

> Generalizing these results to cases where [the quantum system] is not in a pure state is trivial if [the system] is part of a larger system that is in a pure state, for then every measurement on [the system] is also a measurement on the larger system. Further generalisation to exotic situations in which the universe as a whole may be in a mixed state [...] is left as an exercise for the readers.

It appears from the first sentence of this quote that Deutsch wants to treat epistemic uncertainty about the system's state by using the well-known observation that a mixed state can be treated as a pure state for a larger system, which would make his earlier arguments about how to derive Born's rule for pure states amenable to mixed states as well. But, since he assumes a linear ordering of games/measurements for such pure states, it seems fair to infer that he'd be willing to order games/measurements linearly also when the system's state isn't perfectly known to You.[27]

Our willingness to work, in that case, with *partial* preference orderings — which aren't reducible to mixed states nor to pure states in higher-dimensional state spaces — when dealing with Your epistemic uncertainty about the system's state, can therefore only be seen

[26]Bruno de Finetti [15, 17] likewise assumes that all uncertain quantities X can be compared to one another through the intervention of the values that their prevision, or fair price, $P(X)$ assume on the real line.

[27]This of course assumes that You're able to identify the pure state that the larger system is in, which is quite a strong assumption to make. The strength of this assumption is a good indicator of the strength of the *totality* assumption for Your preference ordering. In addition, it's not entirely clear to us why Deutsch would consider a situation where You have epistemic uncertainty about the 'state of the universe as a whole' — that the universe as a whole is in a mixed state — to be 'exotic'.

as a fundamental departure from Deutsch's way of thinking. Interestingly, as we've seen in Section 6.2, such partial preference orderings are mathematically (almost-)equivalent to convex closed sets of mixed states — density operators — and therefore to *sets of pure states* for a higher-dimensional system.

8. WALLACE'S DECISION-THEORETIC APPROACH

Soon after Deutsch published his argument, Barnum et al. [5] pointed to a flaw in it: in order for Deutsch's proof to work, another critical *assumption* is necessary, namely that Your value for the game $\text{game}^{\hat{A}}_{|\psi\rangle}$ should equal Your value for game $\text{game}^{\hat{B}}_{|\phi\rangle}$, where $\hat{B} = \hat{U}\hat{A}\hat{U}^{\dagger}$ and $|\phi\rangle = \hat{U}|\psi\rangle$, and where $\hat{U}$ is some unitary operator on $\mathscr{X}$. In words, a unitary transformation of the Hilbert space and the corresponding measurements mustn't have any influence on Your valuation of the game.[28] Wallace [48] has argued that the above condition can be derived in the context of the Everettian interpretation and has continued to improve on this work [47, 49].

However, in contradistinction with Deutsch, who we've already argued relies on de Finetti's approach in representing Your preferences between games involving measurements, Wallace uses Savage's approach; see Section 3.2 for a short summary of Savage's ideas. Because Wallace's argument is much more detailed and intricate, we'll try to hint at its decision-theoretic essence by using what we believe to be relevant quotations from his more recent work [47] and providing them with our interpretations,[29] where we've preserved the wording, but adapted the mathematical notation to accord with ours. We'll discuss the differences with our approach as we go along.

> A quantum state is to be prepared in some superposition; the system is measured in some basis; a bet is made by the agent on the outcome of that measurement. Our agent knows (we assume) that the Everett interpretation is correct; he is also assumed to know the universal quantum state, or at least the state of his branch. [...] His preferences can be represented by an ordering relation on these bets. Since (in Everettian quantum mechanics, at any rate) preparations, measurements and payments made to agents are all physical processes, there is a certain simplification available: any preparation-followed-by-measurement-followed-by-payments can be represented by a single unitary transformation. So our agent's rational preference is actually representable by an ordering on unitary transformations.

So, we see that You are to express preferences, in a quantum state $|\psi\rangle$, between acts, which are the unitary transformations $\hat{U}$ that are available to You in that state.

> We now need to represent the agent's preferences between acts. Since those preferences may well depend on the state, we write it as follows: if the agent prefers (at $|\psi\rangle$) act $\hat{U}$ to act $\hat{U}'$, we write
>
> $$\hat{U} \succ^{|\psi\rangle} \hat{U}'.$$
>
> To be meaningful, of course, this requires that $\hat{U}$ and $\hat{U}'$ are both available at $|\psi\rangle$'s macrostate. So $\succ^{|\psi\rangle}$ is to be a two-place relation on the set of acts available at that macrostate.

So, like Deutsch, and unlike us, Wallace assumes the quantum system to be in a known (macro)state. Unlike Deutsch, and like us, he starts with a strict preference ordering that You have on a set of acts and where each act has an uncertain reward. But as we'll see, unlike both Deutsch and us, he doesn't assume that the rewards can be expressed in some

[28] This invariance under unitary transformations is in particular implied by our postulate RF2.

[29] We're referring to the arXiv version, which is dated 'October 22, 2018', of a chapter published in 2010 as *How to prove the Born Rule* in Ref. [41].

predetermined linear utility scale; like Savage, he 'constructs' the utilities from the properties of the preference ordering and the richness of the act (and reward) space.

Indeed, Wallace then imposes a number of (so-called *richness*) axioms, or assumptions, that guarantee that the set of acts (and rewards) is rich enough, and a number of (so-called *rationality*) axioms on Your preferences on such acts, the most important of which for our present discussion is the following:

> **Ordering:** The relation $\succeq^{|\psi\rangle}$ is a total ordering for each ψ on the set of acts available at ψ, for each ψ (that is: it's transitive, irreflexive and asymmetric, and if we define $\hat{U} \sim^{|\psi\rangle} \hat{V}$ as holding whenever $\hat{U} \succ^{|\psi\rangle} \hat{V}$ and $\hat{V} \succ^{|\psi\rangle} \hat{U}$ fail to hold, then $\sim^{|\psi\rangle}$ is an equivalence relation).

First off, we believe the symbol '$\succeq^{|\psi\rangle}$' to be a typo for '$\succ^{|\psi\rangle}$', as the former symbol is nowhere defined in Wallace's text and the 'weak' aspect that the extra horizontal bar might suggest is contradicted by the irreflexivity requirement. So, this axiom, fully in line with Savage's approach, requires that Your preference ordering $\succ^{|\psi\rangle}$ should be a strict partial ordering[30] (as is also assumed on our approach) that has the additional *totality* property: the associated relation $\sim^{|\psi\rangle}$, which generally speaking represents both indifference and incomparability, must actually be an equivalence relation and can therefore only represent indifference; incomparability in a given state is excluded on Wallace's approach (as it is on ours as well).

Based on his richness and rationality axioms and following the path blazed by Savage, Wallace is able to provide an ingenious argument to show that there's a utility function on the rewards, such that the strict weak ordering $\succ^{|\psi\rangle}$ on the acts $\hat{U}$ can be represented by the strict total ordering of their expected utilities, where the underlying probabilities are provided by Born's rule.

9. What are the advantages to our approach?

It's clear that the decision problem that Wallace considers, is quite similar to (if more involved than) Deutsch's, in that acts are compared for a quantum system in a given, perfectly known, quantum state $|\psi\rangle$. Contrary to our approach, there's no uncertainty about the quantum system's state.

In our more general decision problem — whose solution coincides with Deutsch's and Wallace's solutions in the special case that the system state is indeed known — we allow for strict partial (preference) orderings *that needn't be total*, so we allow for incomparability between acts: $\hat{A} \not\succ \hat{B}$ and $\hat{B} \not\succ \hat{A}$ needn't imply that You are indifferent between the measurements $\hat{A}$ and $\hat{B}$: You may also hold them to be incomparable, and then You won't be able to compare them on a linear scale.

Indeed, much of what we do in this paper amounts to showing that this total ordering requirement isn't necessary *in our more general decision problem*: we can still derive interesting results, and in particular also recover Born's rule as a special case, without it, using our decision-theoretic assumptions. At the same time, we've shown in the previous sections that in doing so, we uncover an interesting mathematical toolbox for representing, and making conservative inferences about, uncertainty about quantum systems. We start out from the assumption that You are (or may be) uncertain about the state $|\Psi\rangle$ of the quantum system under consideration and that performing measurements $\hat{A}$ on the system, where the rewards are expressed in units of some predetermined linear utility, will therefore yield uncertain rewards, which You can order via a strict ordering that is allowed to be merely *partial when You don't know the system's state*. This freedom has at least two advantages, besides its making the resulting theory more general: (i) it's more realistic because specifying only a partial order is a lot easier for You to do, and all the more so if You only have limited — finite — time and resources at Your disposal; and (ii) it uncovers

[30] A strict partial ordering is an irreflexive and transitive binary relation; the asymmetry is then implied.

the constructive conservative inference mechanism that underlies doing inferences about the state of a quantum system, that tends to get hidden by a focus on total orderings.[31]

Born's rule is therefore still recovered, even if we allow for partial orderings to represent Your preferences in our more general decision problem where the system state is unknown to You.

10. Formulation and proof of the main theorem

Let's now prove the main result in this paper, which is, essentially, that the postulates RF1–RF4 guarantee that the shape of the reward functions $w_{\hat{A}}$ is completely fixed, as in Eq. (8).

The exact formulation in Theorem 13 below takes some care and preparation, however. We'll need to be able to switch freely between various finite-dimensional Hilbert spaces, which is why we'll need to consider the set $\mathbb{X}$ of all of them.

We'll use the notation $\mathscr{X}, \mathscr{X}', \mathscr{X}''$ for elements of $\mathbb{X}$. Every such Hilbert space comes with its own inner product and its own real linear space of Hermitian operators, and we'll always use the same respective notations $\langle\bullet|\bullet\rangle$ and $\mathscr{H}$ for these, as it will always be clear from the context which Hilbert space they're associated with.

In each such finite-dimensional Hilbert space $\mathscr{X} \in \mathbb{X}$, there are of course infinitely many ways to associate a reward function $w_{\hat{A}} \in \mathscr{L}(\bar{\mathscr{X}})$ with each measurement $\hat{A} \in \mathscr{H}$, and each of them corresponds to a so-called *reward assignation* $w_\bullet \colon \mathscr{H} \to \mathscr{L}(\bar{\mathscr{X}}) \colon \hat{A} \mapsto w_{\hat{A}}$ *for* $\mathscr{X}$. Recall that we denote by $\mathscr{L}(\bar{\mathscr{X}})$ the set of all real-valued maps on the state space $\bar{\mathscr{X}}$. We'll denote by $\mathbb{W}_{\mathscr{X}}$ the set of all such possible reward assignations $w_\bullet$ for the Hilbert space $\mathscr{X}$.

Each of the many ways of doing this for all possible finite-dimensional Hilbert spaces is captured in a specific map W on the set $\mathbb{X}$ that selects a reward assignation $W(\mathscr{X}) \in \mathbb{W}_{\mathscr{X}}$ for each Hilbert space $\mathscr{X} \in \mathbb{X}$. We'll call any such map $W \colon \mathbb{X} \to \mathbb{W}_{\mathscr{X}}$ a *reward assignation system*. We'll use the notation $w_\bullet$ for the reward assignation $W(\mathscr{X})$, as it will always be clear from the context what Hilbert space $\mathscr{X}$ we're working in.

The task before us, now, is to show that the postulates RF1–RF4 allow for only one specific reward assignation system, namely the one specified by Eq. (8). We'll say that a reward assignation system W *obeys* a specific postulate if the statement in the postulate is true for the reward assignations $w_\bullet = W(\mathscr{X})$ of all the possible instantiations $\mathscr{X}$ of all the various Hilbert spaces mentioned in the postulate.

Theorem 13
There's a unique reward assignation system U that obeys Postulates RF1–RF4. *For all Hilbert spaces $\mathscr{X} \in \mathbb{X}$ and all corresponding $\hat{A} \in \mathscr{H}$, its corresponding reward function $u_{\hat{A}} \colon \bar{\mathscr{X}} \to \mathbb{R}$ is given by*

$$u_{\hat{A}}(|\psi\rangle) = \langle\psi|\hat{A}|\psi\rangle \text{ for all } |\psi\rangle \in \bar{\mathscr{X}}. \tag{41}$$

We want to stress here that Eq. (41) determines the reward assignation system U fully: for any Hilbert space $\mathscr{X}$, it specifies the values $u_{\hat{A}}$ of the reward assignation $u_\bullet = U(\mathscr{X})$ in all Hermitian operators $\hat{A} \in \mathscr{H}$ on $\mathscr{X}$.

In our proof for this result further on, we'll rely heavily on the additivity property of the uncertain rewards for commuting measurements. Before we can formulate this property in Proposition 15 and prove it as a special consequence of Postulate RF3, it's useful to recall the following essential property of commuting Hermitian operators.

Proposition 14 ([37, Thm. 2.2])
Two Hermitian operators $\hat{A}, \hat{B}$ on a Hilbert space $\mathscr{X}$ commute, meaning that $[\hat{A}, \hat{B}] :=$

---

[31] The analogy with propositional logic is telling: (coherent) partial orderings correspond to deductively closed sets and the total orderings to maximal deductively closed sets, or complete theories; focussing on complete theories alone hides the deductive character of propositional logic and it surfaces only when other, non-maximal, deductively closed sets are considered.

$\hat{A}\hat{B} - \hat{B}\hat{A} = \hat{0}$*, if and only if there's some orthonormal basis* $\{|a_1\rangle, \dots, |a_n\rangle\}$ *of* $\mathscr{X}$ *such that each of its elements* $|a_\ell\rangle$ *is an eigenstate for both operators.*

Proposition 15
Consider any reward assignation system W that obeys Postulate RF3*, and any two commuting Hermitian operators* $\hat{A}, \hat{B}$ *on a Hilbert space* $\mathscr{X}$*. Then* $w_{\hat{A}+\hat{B}}(|\phi\rangle) = w_{\hat{A}}(|\phi\rangle) + w_{\hat{B}}(|\phi\rangle)$ *for all* $|\phi\rangle \in \bar{\mathscr{X}}$*.*

Proof. Proposition 14 guarantees that there's some orthogonal basis of states $\{|a_1\rangle, \dots, |a_n\rangle\}$ that are eigenstates for both $\hat{A}$ and $\hat{B}$. Let $\lambda_1, \dots, \lambda_n$ and $\mu_1, \dots, \mu_n$ be the (not necessarily distinct) corresponding eigenvalues of $\hat{A}$ and $\hat{B}$ respectively, then $\hat{A} = \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k|$ and $\hat{B} = \sum_{k=1}^n \mu_k |a_k\rangle\langle a_k|$, by Proposition 1. Let $\lambda_{\max} := \max_{k=1}^n |\lambda_k| = \|\hat{A}\|_\infty$, $\mu_{\max} := \max_{k=1}^n |\mu_k| = \|\hat{B}\|_\infty$ and $\alpha := 2(\lambda_{\max} + \mu_{\max}) + \varepsilon$ with $\varepsilon > 0$. Also let $\lambda'_k := \lambda_k + k\alpha$, $\lambda''_k := -k\alpha$, $\mu'_k := \mu_k + k\alpha$ and $\mu''_k := -k\alpha$ for all $k \in \{1, \dots, n\}$. Then, on the one hand,

$$\lambda_k = \lambda'_k + \lambda''_k \text{ and } \mu_k = \mu'_k + \mu''_k \text{ for all } k \in \{1, \dots, n\}, \tag{42}$$

while, on the other hand,

$$\left.\begin{array}{l} \lambda'_k + \mu'_k \neq \lambda'_\ell + \mu'_\ell \text{ and } \lambda''_k + \mu''_k \neq \lambda''_\ell + \mu''_\ell \\ \lambda'_k \neq \lambda'_\ell \text{ and } \lambda''_k \neq \lambda''_\ell \\ \mu'_k \neq \mu'_\ell \text{ and } \mu''_k \neq \mu''_\ell \end{array}\right\} \text{ for all } k, \ell \in \{1, \dots, n\} \text{ such that } k \neq \ell. \tag{43}$$

Let $\hat{A}' := \sum_{k=1}^n \lambda'_k |a_k\rangle\langle a_k|$ and $\hat{A}'' := \sum_{k=1}^n \lambda''_k |a_k\rangle\langle a_k|$, and similarly $\hat{B}' := \sum_{k=1}^n \mu'_k |a_k\rangle\langle a_k|$ and $\hat{B}'' := \sum_{k=1}^n \mu''_k |a_k\rangle\langle a_k|$, then it follows from Eq. (42) that $\hat{A} = \hat{A}' + \hat{A}''$ and $\hat{B} = \hat{B}' + \hat{B}''$. Moreover, Eq. (43) guarantees that the measurements $\hat{A}'$ and $\hat{A}''$, as well as the measurements $\hat{B}'$ and $\hat{B}''$, have distinct eigenvalues. Postulate RF3 then ensures that

$$w_{\hat{A}} = w_{\hat{A}'} + w_{\hat{A}''} \text{ and } w_{\hat{B}} = w_{\hat{B}'} + w_{\hat{B}''}. \tag{44}$$

But Eq. (43) also guarantees that the measurements $\hat{A}'$ and $\hat{B}'$, as well as the measurements $\hat{A}''$ and $\hat{B}''$, have distinct eigenvalues, so we can equally well apply RF3 to find that

$$w_{\hat{A}'+\hat{B}'} = w_{\hat{A}'} + w_{\hat{B}'} \text{ and } w_{\hat{A}''+\hat{B}''} = w_{\hat{A}''} + w_{\hat{B}''}. \tag{45}$$

Finally, Eq. (43) guarantees that the measurements $\hat{A}' + \hat{B}'$ and $\hat{A}'' + \hat{B}''$ have distinct eigenvalues, and since $\hat{A} + \hat{B} = (\hat{A}' + \hat{B}') + (\hat{A}'' + \hat{B}'')$, we can again use Postulate RF3 to get

$$w_{\hat{A}+\hat{B}} = w_{\hat{A}'+\hat{B}'} + w_{\hat{A}''+\hat{B}''}. \tag{46}$$

But then, summarising,

$$\begin{aligned} w_{\hat{A}+\hat{B}} &\overset{(46)}{=} w_{\hat{A}'+\hat{B}'} + w_{\hat{A}''+\hat{B}''} \overset{(45)}{=} (w_{\hat{A}'} + w_{\hat{B}'}) + (w_{\hat{A}''} + w_{\hat{B}''}) \\ &= (w_{\hat{A}'} + w_{\hat{A}''}) + (w_{\hat{B}} + w_{\hat{B}''}) \overset{(44)}{=} w_{\hat{A}} + w_{\hat{B}}. \end{aligned}$$

□

We'll also rely on the following consequence of RF2, in a form that doesn't require the eigenvalues considered to be different.

Proposition 16
Consider any reward assignation system W that obeys Postulate RF2*, any Hermitian operator* $\hat{A}$ *on a Hilbert space* $\mathscr{X}_1$*, and any orthogonal basis of its eigenstates* $\{|a_1\rangle, \dots, |a_n\rangle\}$ *with corresponding (not necessarily distinct) real eigenvalues* $\lambda_1, \dots, \lambda_n$*, implying that* $\hat{A} = \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k|$*. Also consider a Hilbert space* $\mathscr{X}_2$ *and an Hermitian operator* $\hat{B}$ *on* $\mathscr{X}_2$ *with the same eigenvalues* $\lambda_1, \dots, \lambda_n$ *and with orthogonal eigenstates* $|b_1\rangle, \dots, |b_n\rangle$ *corresponding to these respective eigenvalues. Choose any* $\alpha_k \in \mathbb{C}$ *such that* $\sum_{k=1}^n |\alpha_k|^2 = 1$*, and consider the states* $|\phi_{\hat{A}}\rangle := \sum_{k=1}^n \alpha_k |a_k\rangle \in \mathscr{X}_1$ *and* $|\phi_{\hat{B}}\rangle := \sum_{k=1}^n \alpha_k |b_k\rangle \in \mathscr{X}_2$*. Then* $w_{\hat{A}}(|\phi_{\hat{A}}\rangle) = w_{\hat{B}}(|\phi_{\hat{B}}\rangle)$*.*

Proof. We may assume without loss of generality that there are $r \in \mathbb{N}$ distinct eigenvalues $\mu_k \in \mathbb{R}$ with corresponding multiplicities $m_k \in \mathbb{N}$, so we have for the eigenvalues of $\hat{A}$ that $\lambda_\ell = \mu_k$ for all $\ell \in \{M_{k-1}+1, \dots, M_k\}$ and $k \in \{1, \dots, r\}$, $M_0 := 0$ and $M_k := \sum_{\ell=1}^{k} m_\ell$. The eigenspace for $\hat{A}$ that corresponds to the eigenvalue μ_k is then given by $\mathscr{E}_k := \operatorname{span}(\{|a_\ell\rangle : \ell \in \{M_{k-1}+1, \dots, M_k\}\})$. Correspondingly, we see that the linear space $\operatorname{span}(\{|b_\ell\rangle : \ell \in \{M_{k-1}+1, \dots, M_k\}\})$ is a subspace of the eigenspace $\mathscr{F}_k$ for $\hat{B}$ corresponding to the eigenvalue μ_k. The r orthogonal eigenspaces $\mathscr{E}_k$ span $\mathscr{X}_1$, and the r orthogonal eigenspaces $\mathscr{F}_k$ span $\mathscr{X}_2$, by Corollary 2.

For any $k \in \{1, \dots, r\}$, choose the complex number β_k in such a way that $|\beta_k|^2 = \sum_{\ell=M_{k-1}+1}^{M_k} |\alpha_\ell|^2$. If $\beta_k \neq 0$, then let

$$|c_k\rangle := \frac{1}{\beta_k} \sum_{\ell=M_{k-1}+1}^{M_k} \alpha_\ell |a_\ell\rangle \text{ and } |d_k\rangle := \frac{1}{\beta_k} \sum_{\ell=M_{k-1}+1}^{M_k} \alpha_\ell |b_\ell\rangle.$$

If $\beta_k = 0$, then let $|c_k\rangle$ be any state in $\mathscr{E}_k$ and $|d_k\rangle$ any state in $\mathscr{F}_k$. It follows that $|c_k\rangle$ is always a state in $\mathscr{E}_k$ and that $|d_k\rangle$ is always a state in $\mathscr{F}_k$, and also that $\sum_{k=1}^{r} |\beta_k|^2 = \sum_{\ell=1}^{n} |\alpha_\ell|^2 = 1$, $|\phi_{\hat{A}}\rangle = \sum_{\ell=1}^{n} \alpha_\ell |a_\ell\rangle = \sum_{k=1}^{r} \beta_k |c_k\rangle$ and $|\phi_{\hat{B}}\rangle = \sum_{\ell=1}^{n} \alpha_\ell |b_\ell\rangle = \sum_{k=1}^{r} \beta_k |d_k\rangle$. Now apply RF2 [with $\lambda_k \rightsquigarrow \mu_k$, $\alpha_k \rightsquigarrow \beta_k$, $|a_k\rangle \rightsquigarrow |c_k\rangle$ and $|b_k\rangle \rightsquigarrow |d_k\rangle$] to conclude that $w_{\hat{A}}(|\phi_{\hat{A}}\rangle) = w_{\hat{B}}(|\phi_{\hat{B}}\rangle)$. □

To make the proof of Theorem 13 easier to digest, we'll split the argument into several successive propositions that are stepping stones on our way to the main result, and which allow us to uncover the form of the reward function $w_{\hat{A}}(|\psi\rangle)$ in increasingly more general types of arguments $|\psi\rangle \in \bar{\mathscr{X}}$.

In a first step, we manage to fix the value of the reward function in equal-amplitude superpositions of eigenstates of a measurement operator.

Proposition 17
Consider any reward assignation system W that obeys Postulates RF1–RF3*; any Hilbert space $\mathscr{X} \in \mathbb{X}$, with corresponding reward assignation $w_\bullet := W(\mathscr{X})$; and any measurement $\hat{A} \in \mathscr{H}$, with any orthogonal basis of eigenstates $\{|a_1\rangle, \dots, |a_n\rangle\}$. Then*

$$w_{\hat{A}}(|\psi_m\rangle) = \langle \psi_m | \hat{A} | \psi_m \rangle \textit{ for all } |\psi_m\rangle := \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |a_k\rangle, \textit{ with } 1 \leq m \leq n.$$

Proof. Before we start, infer from Lemma 4 that all $|\psi_m\rangle \in \bar{\mathscr{X}}$. Also, denote by λ_k the eigenvalue of $\hat{A}$ corresponding to the eigenstate $|a_k\rangle$, for all $k \in \{1, \dots, n\}$.

We begin with the simplest case that $m = 1$, so with $|\psi_1\rangle = |a_1\rangle$. RF1 then guarantees that, on the one hand, $w_{\hat{A}}(|\psi_1\rangle) = w_{\hat{A}}(|a_1\rangle) = \lambda_1$. On the other hand, we infer from Lemma 6 that $\lambda_1 = \langle a_1 | \hat{A} | a_1 \rangle = \langle \psi_1 | \hat{A} | \psi_1 \rangle$. Hence, indeed, $w_{\hat{A}}(|\psi_1\rangle) = \langle \psi_1 | \hat{A} | \psi_1 \rangle$.

Next, we turn to the more involved case that $1 < m \leq n$. We denote by $\pi_m : \{1, \dots, n\} \to \{1, \dots, n\}$ the cyclic permutation of the first m indices, defined by

$$\pi_m(k) := \begin{cases} k+1 & \text{if } k+1 \leq m \\ 1 & \text{if } k = m \\ k & \text{if } m < k \leq n \end{cases} \quad \text{for all } k \in \{1, 2, \dots, n\}.$$

For any $\ell \in \{0, 1, \dots, m\}$, we let π_m^ℓ denote the result of applying π_m ℓ times, so π_m^ℓ satisfies

$$\pi_m^\ell(k) = \begin{cases} k+\ell & \text{if } k+\ell \leq m \\ k+\ell-m & \text{if } m < k+\ell \leq m+\ell \\ k & \text{if } m < k \leq n \end{cases} \text{ for all } k \in \{1, 2, \dots, n\}.$$

Also observe that

$$(\pi_m^\ell)^{-1} = \pi_m^{m-\ell} \text{ for all } \ell \in \{0, 1, 2, \dots, m\}. \tag{47}$$

Using these permutations, we can now construct m Hermitian operators $\hat{A}_\ell \in \mathscr{H}$ as follows [see Proposition 1]:

$$\hat{A}_\ell := \sum_{k=1}^{n} \lambda_k |a_{\pi_m^\ell(k)}\rangle\langle a_{\pi_m^\ell(k)}|, \text{ for } \ell \in \{0,\dots,m-1\},$$

is the Hermitian operator with eigenstates $|a_{\pi_m^\ell(k)}\rangle$ corresponding to the respective real eigenvalues λ_k for $k \in \{1,\dots,n\}$. To use Proposition 16, with $\mathscr{X}_1 \rightsquigarrow \mathscr{X}$, $\mathscr{X}_2 \rightsquigarrow \mathscr{X}$, $\hat{A} \rightsquigarrow \hat{A}$, $\hat{B} \rightsquigarrow \hat{A}_\ell$ and $|\phi_{\hat{A}}\rangle \rightsquigarrow |\psi_m\rangle$ for all $\ell \in \{0,1,\dots,m-1\}$, we define

$$|\phi_{\hat{B}}\rangle = |\phi_{\hat{A}_\ell}\rangle := \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |a_{\pi_m^\ell(k)}\rangle = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |a_k\rangle = |\psi_m\rangle \text{ for all } \ell \in \{0,1,\dots,m-1\},$$

and then Proposition 16 implies that

$$w_{\hat{A}}(|\psi_m\rangle) = w_{\hat{A}}(|\phi_{\hat{A}}\rangle) = w_{\hat{B}}(|\phi_{\hat{B}}\rangle) = w_{\hat{A}_\ell}(|\psi_m\rangle) \text{ for all } \ell \in \{0,1,\dots,m-1\},$$

and therefore also that

$$\sum_{\ell=0}^{m-1} w_{\hat{A}_\ell}(|\psi_m\rangle) = m w_{\hat{A}}(|\psi_m\rangle). \tag{48}$$

Furthermore, because we infer from Proposition 1 that the elements of the orthonormal basis $\{|a_1\rangle,\dots,|a_n\rangle\}$ are the eigenstates for each of the Hermitian operators $\hat{A}_0$, $\hat{A}_1$, …, $\hat{A}_{m-1}$, we infer from Proposition 14 that these operators commute: $[\hat{A}_r,\hat{A}_s] = 0$ for all $r,s \in \{0,1,\dots,m-1\}$. But then clearly also $[\hat{A}_r, \sum_{k=0}^{r-1} \hat{A}_k] = 0$ for all $r \in \{1,\dots,m-1\}$. Successively applying Proposition 15 $m-1$ times, and letting $\hat{C} := \sum_{\ell=0}^{m-1} \hat{A}_\ell$, we then find for the left-hand side of Eq. (48) that

$$\sum_{\ell=0}^{m-1} w_{\hat{A}_\ell}(|\psi_m\rangle) = w_{\hat{C}}(|\psi_m\rangle). \tag{49}$$

Moreover, we recall from the definition of $\hat{A}_0$, …, $\hat{A}_{m-1}$ that

$$\begin{aligned}\hat{C} &= \sum_{\ell=0}^{m-1} \hat{A}_\ell = \sum_{\ell=0}^{m-1} \sum_{k=1}^{n} \lambda_k |a_{\pi_m^\ell(k)}\rangle\langle a_{\pi_m^\ell(k)}| \overset{(47)}{=} \sum_{\ell=0}^{m-1} \sum_{r=1}^{n} \lambda_{\pi_m^{m-\ell}(r)} |a_r\rangle\langle a_r| = \sum_{s=1}^{m} \sum_{r=1}^{n} \lambda_{\pi_m^s(r)} |a_r\rangle\langle a_r| \\ &= \sum_{r=1}^{m} \sum_{s=1}^{m} \lambda_{\pi_m^s(r)} |a_r\rangle\langle a_r| + \sum_{r=m+1}^{n} \sum_{s=1}^{m} \lambda_{\pi_m^s(r)} |a_r\rangle\langle a_r| = \sum_{r=1}^{m} \Big(\sum_{\ell=1}^{m} \lambda_\ell \Big) |a_r\rangle\langle a_r| + \sum_{r=m+1}^{n} (m\lambda_r) |a_r\rangle\langle a_r|.\end{aligned}$$

This, together with Proposition 1, tells us that the Hermitian operator $\hat{C}$ has the same eigenvalue $\sum_{\ell=1}^{m} \lambda_\ell$ corresponding to each of the eigenstates $|a_k\rangle$ for $k \in \{1,2,\dots,m\}$. Therefore, the linear combination $|\psi_m\rangle = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} |a_k\rangle$ of these m eigenstates will also be an eigenstate of $\hat{C}$ with this same eigenvalue $\sum_{\ell=1}^{m} \lambda_\ell$. RF1 then guarantees that $w_{\hat{C}}(|\psi_m\rangle) = \sum_{\ell=1}^{m} \lambda_\ell$, so Eqs. (48) and (49) then yield that $w_{\hat{A}}(|\psi_m\rangle) = \frac{1}{m} \sum_{\ell=1}^{m} \lambda_\ell$. Since, on the other hand, Lemma 6 guarantees that $\langle \psi_m | \hat{A} | \psi_m \rangle = \sum_{\ell=1}^{m} \frac{1}{m} \lambda_\ell$, we're done. □

In a second step, we succeed in fixing the value of the reward function in real linear combinations $\sum_{k=1}^{n} \alpha_k |a_k\rangle$ of the eigenstates $|a_k\rangle$ of a measurement operator $\hat{A}$, provided that the squares α_k^2 of their real amplitudes α_k are rational.

Proposition 18

Consider any reward assignation system W that obeys Postulates RF1–RF3*, any n-dimensional Hilbert space $\mathscr{X} \in \mathbb{X}$ with corresponding reward assignation $w_\bullet := W(\mathscr{X})$ and any measurement $\hat{A} \in \mathscr{H}$ with any orthogonal basis of eigenstates $\{|a_1\rangle,\dots,|a_n\rangle\}$. Then*

$$w_{\hat{A}}(|\psi_m\rangle) = \langle \psi_m | \hat{A} | \psi_m \rangle \textit{ for all } |\psi_m\rangle := \sum_{k=1}^{m} \sqrt{q_k} |a_k\rangle,$$
$$\textit{with } 1 \le m \le n \textit{ and } q_1,\dots,q_m \in \mathbb{Q}_{>0} \textit{ such that } \sum_{k=1}^{m} q_k = 1.$$

Proof. Before we start, infer from Lemma 4 that all $|\psi_m\rangle \in \bar{\mathscr{X}}$. Due to Proposition 1, we know that $\hat{A} = \sum_{k=1}^n \lambda_n |a_k\rangle\langle a_k|$, with λ_k the eigenvalue corresponding to eigenstate $|a_k\rangle$.

First, since $q_1,\dots,q_m$ are positive rational numbers, there are natural numbers r and $p_1,\dots,p_m$ such that $p_k = q_k r$ for all $k \in \{1,\dots,m\}$. Since $\sum_{k=1}^m q_k = 1$, we also have that $\sum_{k=1}^m p_k = r$, and then clearly $|\psi_m\rangle = \frac{1}{\sqrt{r}}\sum_{k=1}^m \sqrt{p_k}|a_k\rangle$. We'll also consider an arbitrary r-dimensional Hilbert space, which we denote by $\mathscr{X}'$.

Define the Hermitian operator $\hat{B} := \hat{A}\otimes\hat{I} = \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k| \otimes \hat{I}$ on the tensor product space $\mathscr{X}\otimes\mathscr{X}'$. For the remainder of the argument, we'll now consider an arbitrary but fixed $|\phi\rangle \in \bar{\mathscr{X}}'$. Then $|a_k\rangle\otimes|\phi\rangle$ is an eigenstate of $\hat{B} = \hat{A}\otimes\hat{I}$ with corresponding eigenvalue λ_k, and all these n eigenstates $|a_k\rangle\otimes|\phi\rangle$ are mutually orthogonal.

We now want to make use of Proposition 16, with $\mathscr{X}_1 = \mathscr{X}$, $\mathscr{X}_2 = \mathscr{X}\otimes\mathscr{X}'$, $\hat{A} \rightsquigarrow \hat{A}$, $\hat{B} \rightsquigarrow \hat{B}$, $|\phi_{\hat{A}}\rangle \rightsquigarrow |\psi_m\rangle$. We then also let $|\phi_{\hat{B}}\rangle \rightsquigarrow |\psi_m\rangle\otimes|\phi\rangle = \frac{1}{\sqrt{r}}\sum_{k=1}^m \sqrt{p_k}|a_k\rangle\otimes|\phi\rangle$. Since each $|a_k\rangle\otimes|\phi\rangle$ is an eigenstate of $\hat{B}$ corresponding to the eigenvalue λ_k, it follows from Proposition 16 that

$$w_{\hat{A}}(|\psi_m\rangle) = w_{\hat{B}}(|\psi_m\rangle\otimes|\phi\rangle). \tag{50}$$

Now, use Lemma 5 to construct from the chosen and fixed $|\phi\rangle \in \bar{\mathscr{X}}'$, for any $k \in \{1,\dots,m\}$, the p_k mutually orthogonal states $|\phi_1^k\rangle,\dots,|\phi_{p_k}^k\rangle \in \bar{\mathscr{X}}'$ such that $|\phi\rangle = \frac{1}{\sqrt{p_k}}\sum_{\ell=1}^{p_k}|\phi_\ell^k\rangle$ [which is possible since $0 < p_k \leq r$]. Due to the bi-linearity of the tensor product, we then find that

$$|\psi_m\rangle\otimes|\phi\rangle = \left(\frac{1}{\sqrt{r}}\sum_{k=1}^m \sqrt{p_k}|a_k\rangle\right)\otimes\left(\frac{1}{\sqrt{p_k}}\sum_{\ell=1}^{p_k}|\phi_\ell^k\rangle\right) = \frac{1}{\sqrt{r}}\sum_{k=1}^m\sum_{\ell=1}^{p_k}|a_k\rangle\otimes|\phi_\ell^k\rangle. \tag{51}$$

This is now an equal-amplitude superposition of the $r = \sum_{k=1}^m p_k$ mutually orthogonal eigenstates $|a_k\rangle\otimes|\phi_\ell^k\rangle$ of $\hat{B}$ with corresponding eigenvalue λ_k, where $\ell \in \{1,\dots,p_k\}$ for $k \in \{1,\dots,m\}$. For all $k \in \{1,2,\dots,n\}$, we define the one-dimensional linear subspaces $\mathscr{V}_k := \operatorname{span}(\{|a_k\rangle\})$. For any fixed $k \in \{1,2,\dots,m\}$, we can use the Gram–Schmidt procedure to extend the orthonormal collection $|a_k\rangle\otimes|\phi_1^k\rangle, |a_k\rangle\otimes|\phi_2^k\rangle,\dots,|a_k\rangle\otimes|\phi_{p_k}^k\rangle$ — with p_k elements — to an orthonormal basis $\mathscr{B}_k$ for $\mathscr{V}_k\otimes\mathscr{X}'$ — with r elements. Similarly, for any $k \in \{m+1,\dots,n\}$, we define an orthonormal basis $\mathscr{B}_k$ for $\mathscr{V}_k\otimes\mathscr{X}'$ — with r elements. As the direct sum of these orthogonal linear subspaces $\mathscr{V}_k\otimes\mathscr{X}'$ is the complete Hilbert space $\mathscr{X}\otimes\mathscr{X}'$, the union of the bases $\bigcup_{k=1}^n \mathscr{B}_k$ is a basis of eigenstates for the Hilbert space $\mathscr{X}\otimes\mathscr{X}'$ that contains all the orthogonal eigenstates $|a_k\rangle\otimes|\phi_\ell^k\rangle$ in Eq. (51). Since $|\psi_m\rangle\otimes|\phi\rangle = \frac{1}{\sqrt{r}}\sum_{k=1}^m\sum_{\ell=1}^{p_k}|a_k\rangle\otimes|\phi_\ell^k\rangle$, we're therefore in a position to apply Proposition 17 and find that

$$\begin{aligned} w_{\hat{B}}(|\psi_m\rangle\otimes|\phi\rangle) &= (\langle\psi_m|\otimes\langle\phi|)\hat{B}(|\psi_m\rangle\otimes|\phi\rangle) = (\langle\psi_m|\otimes\langle\phi|)\hat{A}\otimes\hat{I}(|\psi_m\rangle\otimes|\phi\rangle) \\ &= \langle\psi_m|\hat{A}|\psi_m\rangle\langle\phi|\hat{I}|\phi\rangle = \langle\psi_m|\hat{A}|\psi_m\rangle\langle\phi|\phi\rangle \\ &= \langle\psi_m|\hat{A}|\psi_m\rangle. \end{aligned}$$

This, combined with Eq. (50), shows that $w_{\hat{A}}(|\psi_m\rangle) = \langle\psi_m|\hat{A}|\psi_m\rangle$, so we're done. □

In the next and penultimate step, we move from amplitudes that are the square roots of rationals to all non-negative real amplitudes.

Proposition 19
Consider any reward assignation system W that obeys Postulates RF1–RF4*, any Hilbert space $\mathscr{X} \in \mathbb{X}$ with corresponding reward assignation $w_\bullet := W(\mathscr{X})$, and any measurement $\hat{A} \in \mathscr{H}$ with any orthogonal basis of eigenstates $\{|a_1\rangle,\dots,|a_n\rangle\}$. Then*

$$w_{\hat{A}}(|\psi\rangle) = \langle\psi|\hat{A}|\psi\rangle \textit{ for all } |\psi\rangle := \sum_{k=1}^n \alpha_k|a_k\rangle \in \bar{\mathscr{X}}, \qquad \textit{with } \alpha_1,\dots,\alpha_n \in \mathbb{R}_{\geq 0} \textit{ and } \sum_{k=1}^n \alpha_k^2 = 1.$$

Proof. We'll denote by λ_k the eigenvalue of $\hat{A}$ corresponding to the eigenstate $|a_k\rangle$, for all $k \in \{1,\dots,n\}$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there are n sequences ${}_mq_k \in \mathbb{Q}$, $k \in \{1,\dots,n\}$ such that for all $m \in \mathbb{N}$: (i) $\sum_{k=1}^{n} {}_mq_k = 1$; (ii) ${}_mq_k > 0$ for all $k \in \{1,\dots,n\}$; and (iii) $\lim_{m\to+\infty} {}_mq_k = \alpha_k^2$ for all $k \in \{1,\dots,n\}$. We define a sequence of states as follows: let $|\psi_m\rangle := \sum_{k=1}^{n} \sqrt{{}_mq_k}|a_k\rangle$ for all $m \in \mathbb{N}$. Then $\lim_{m\to+\infty} |\psi_m\rangle = |\psi\rangle$, because it follows from our assumptions that $\lim_{m\to+\infty} \sqrt{{}_mq_k} = \alpha_k$ for all $k \in \{1,\dots,n\}$. Now invoke RF4 to find that, on the one hand,

$$w_{\hat{A}}(|\psi\rangle) = \lim_{m\to+\infty} w_{\hat{A}}(|\psi_m\rangle) = \lim_{m\to+\infty} w_{\hat{A}}\left(\sum_{k=1}^{n} \sqrt{{}_mq_k}|a_k\rangle\right) = \lim_{m\to+\infty} \sum_{k=1}^{n} {}_mq_k\lambda_k = \sum_{k=1}^{n} \alpha_k^2\lambda_k,$$

where for the third equality, we used Proposition 18 and Lemma 6. On the other hand, we also infer from Lemma 6 that $\langle\psi|\hat{A}|\psi\rangle = \sum_{k=1}^{n} |\alpha_k|^2\lambda_k = \sum_{k=1}^{n} \alpha_k^2\lambda_k$, so we're done. $\square$

With all these stepping stones in place, it's now fairly straightforward to prove the existence and uniqueness Theorem 13.

Proof of Theorem 13. We begin with the existence part: it's enough to show that the reward assignation system U obeys all four postulates RF1–RF4.

DT1. Fix any Hilbert space $\mathscr{X}$ and any Hermitian operator $\hat{A}$ on $\mathscr{X}$. If $|a\rangle$ is an eigenstate associated with an eigenvalue λ of $\hat{A}$, then, indeed, $u_{\hat{A}}(|a\rangle) = \langle a|\hat{A}|a\rangle = \langle a|\lambda a\rangle = \lambda\langle a|a\rangle = \lambda$.

DT2. Fix any two Hilbert spaces $\mathscr{X}_1$ and $\mathscr{X}_2$, and consider the Hermitian operator $\hat{A} := \sum_{k=1}^{r} \lambda_k \hat{P}_{\mathscr{E}_k}$ on $\mathscr{X}_1$, with (distinct) real eigenvalues $\lambda_1,\dots,\lambda_r$ corresponding to the respective mutually orthogonal eigenspaces $\mathscr{E}_1,\dots,\mathscr{E}_r$ that span $\mathscr{X}_1$. Similarly, consider the Hermitian operator $\hat{B} := \sum_{k=1}^{r} \lambda_k \hat{P}_{\mathscr{F}_k}$ on $\mathscr{X}_2$, with the same (distinct) eigenvalues $\lambda_1,\dots,\lambda_r$, corresponding to the respective mutually orthogonal eigenspaces $\mathscr{F}_1,\dots,\mathscr{F}_r$ that span $\mathscr{X}_2$. Fix any states $|a_k\rangle \in \mathscr{E}_k$ and $|b_k\rangle \in \mathscr{F}_k$, and any $\alpha_k \in \mathbb{C}$ such that $\sum_{k=1}^{r} |\alpha_k|^2 = 1$. Consider the states $|\phi_{\hat{A}}\rangle := \sum_{k=1}^{r} \alpha_k|a_k\rangle \in \mathscr{X}_1$ and $|\phi_{\hat{B}}\rangle := \sum_{k=1}^{r} \alpha_k|b_k\rangle \in \mathscr{X}_2$. Then it follows from Lemma 6 that $u_{\hat{A}}(|\phi_{\hat{A}}\rangle) = \langle\phi_{\hat{A}}|\hat{A}|\phi_{\hat{A}}\rangle = \sum_{k=1}^{r} |\alpha_k|^2\lambda_k$, and similarly, that $u_{\hat{B}}(|\phi_{\hat{B}}\rangle) = \langle\phi_{\hat{B}}|\hat{B}|\phi_{\hat{B}}\rangle = \sum_{k=1}^{r} |\alpha_k|^2\lambda_k$. Hence, indeed, $u_{\hat{A}}(|\phi_{\hat{A}}\rangle) = u_{\hat{B}}(|\phi_{\hat{B}}\rangle)$.

DT3. Fix any Hilbert space $\mathscr{X}$, then we find for any two Hermitian operators $\hat{A}$ and $\hat{B}$ on $\mathscr{X}$ [regardless of whether they have the same eigenspaces or not] that

$$u_{\hat{A}+\hat{B}}(|\phi\rangle) = \langle\phi|(\hat{A}+\hat{B})|\phi\rangle = \langle\phi|\hat{A}|\phi\rangle + \langle\phi|\hat{B}|\phi\rangle = u_{\hat{A}}(|\phi\rangle) + u_{\hat{B}}(|\phi\rangle).$$

DT4. Fix any Hilbert space $\mathscr{X}$ and any Hermitian operator $\hat{A}$ on $\mathscr{X}$. The argument is a standard one, but we include it here for the sake of completeness. Consider any sequence of states $|\phi_n\rangle$ with $|\phi\rangle = \lim_{n\to+\infty} |\phi_n\rangle$, or in other words, $\||\phi\rangle - |\phi_n\rangle\| \to 0$. If we also consider the sequence of kets $|\psi_n\rangle := \hat{A}|\phi_n\rangle$, then

$$\|\hat{A}|\phi\rangle - |\psi_n\rangle\| = \|\hat{A}|\phi\rangle - \hat{A}|\phi_n\rangle\| = \|\hat{A}(|\phi\rangle - |\phi_n\rangle)\| \overset{(13)}{\leq} \|\hat{A}\|_{\text{op}} \cdot \||\phi\rangle - |\phi_n\rangle\| \to 0, \quad (52)$$

so $\hat{A}|\phi_n\rangle \to \hat{A}|\phi\rangle$: any Hermitian operator is continuous. Moreover,

$$\begin{aligned}
|u_{\hat{A}}(|\phi\rangle) - u_{\hat{A}}(|\phi_n\rangle)| &= |\langle\phi|\hat{A}|\phi\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle| \\
&= |\langle\phi|\hat{A}|\phi\rangle - \langle\phi|\hat{A}|\phi_n\rangle + \langle\phi|\hat{A}|\phi_n\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle| \\
&\overset{(1)}{=} |\langle\phi|\hat{A}|\phi\rangle - \langle\phi|\hat{A}|\phi_n\rangle + \langle\phi_n|\hat{A}|\phi\rangle^* - \langle\phi_n|\hat{A}|\phi_n\rangle^*| \\
&\leq |\langle\phi|\hat{A}|\phi\rangle - \langle\phi|\hat{A}|\phi_n\rangle| + |\langle\phi_n|\hat{A}|\phi\rangle - \langle\phi_n|\hat{A}|\phi_n\rangle| \\
&= |\langle\phi|\hat{A}(|\phi\rangle - |\phi_n\rangle)| + |\langle\phi_n|\hat{A}(|\phi\rangle - |\phi_n\rangle)| \\
&\leq 2\|\hat{A}(|\phi\rangle - |\phi_n\rangle)\| \overset{(52)}{\to} 0,
\end{aligned}$$

where the second inequality follows from the Cauchy–Schwartz inequality and the fact that $\||\phi\rangle\| = \||\phi_n\rangle\| = 1$. This latest result can also be seen as a direct consequence of the continuity of the inner product in the topology generated by the associated norm. Hence, indeed $u_{\hat{A}}(|\phi\rangle) = \lim_{n\to\infty} u_{\hat{A}}(|\phi_n\rangle)$.

We can now move to the unicity part. Consider any reward assignation system W and assume that it satisfies postulates RF1–RF4, then we're going to show that then necessarily $W = U$. To do this, we're going to consider any Hilbert space $\mathscr{X}$ and any Hermitian operator $\hat{A}$ on $\mathscr{X}$, and prove that $w_{\hat{A}} = u_{\hat{A}}$, where, of course, $w_\bullet = W(\mathscr{X})$.

By Proposition 1, we know that there's an orthonormal basis for $\mathscr{X}$ that consists of eigenstates $|a_1\rangle,\dots,|a_n\rangle$ of $\hat{A}$. If we denote by λ_k the eigenvalue of $\hat{A}$ that corresponds with the eigenstate $|a_k\rangle$, for $k \in \{1,\dots,n\}$, then $\hat{A} = \sum_{k=1}^n \lambda_k |a_k\rangle\langle a_k|$.

Fix any state $|\psi\rangle \in \bar{\mathscr{X}}$. Since $\{|a_1\rangle,\dots,|a_n\rangle\}$ is an orthonormal basis for $\mathscr{X}$, we know from Lemma 4 that there are complex numbers $\alpha_1,\dots,\alpha_n$ such that $\sum_{k=1}^n |\alpha_k|^2 = 1$ and $|\psi\rangle = \sum_{k=1}^n \alpha_k |a_k\rangle$. For all $k \in \{1,2,\dots,n\}$, let $\beta_k := |\alpha_k| \in \mathbb{R}_{\geq 0}$ and $\theta_k = \arg\alpha_k \in [0,2\pi)$. Then $\alpha_k = \beta_k e^{i\theta_k}$ and $|\phi_k\rangle := e^{i\theta_k}|a_k\rangle$ is then obviously also an eigenstate of $\hat{A}$ with eigenvalue λ_k. It's immediate that the $|\phi_1\rangle,\dots,|\phi_n\rangle$ also constitute an orthonormal basis for $\mathscr{X}$ and that

$$|\psi\rangle = \sum_{k=1}^n \alpha_k |a_k\rangle = \sum_{k=1}^n \beta_k e^{i\theta_k} |a_k\rangle = \sum_{k=1}^n \beta_k |\phi_k\rangle \text{ with } \sum_{k=1}^n \beta_k^2 = 1.$$

We then infer from Proposition 19 that, indeed, $w_{\hat{A}}(|\psi\rangle) = \langle\psi|\hat{A}|\psi\rangle = u_{\hat{A}}(|\psi\rangle)$. □

11. Conclusion

We've formulated the problem of uncertainty in a quantum mechanical system as a decision problem between acts that are measurements You can perform on the system. We've tried to address this decision problem under uncertainty using a number of postulates that fix the form of the reward function $u_{\hat{A}}$ associated with each act (or measurement) $\hat{A}$ as $u_{\hat{A}} = \langle\bullet|\hat{A}|\bullet\rangle$.

The upshot of this is that Born's rule turns out to be already incorporated in the reward or the utility aspect of the decision problem and is thereby to some extent freed from its purely probabilistic connotations: we can, in a de Finetti-like approach, separate the utilities and the probabilities, and mathematically express Your preferences as coherent preference orderings on the uncertain rewards, or equivalently, on the measurements. Such coherent preference orderings, as we have seen, can be represented using coherent lower previsions on the real Hilbert space of all measurements. Coherent lower previsions can also be represented mathematically by closed convex sets of density operators; only in special cases do these sets reduce to single density operators and to the standard probabilistic models used in quantum mechanics. In this sense, the decision problem is primary and probabilities are derivative.

To try and drive home this point, let's backtrack a bit and look at *general* decision problems and the role that coherent previsions play there, as described in Section 3.2. Generally speaking, any coherent prevision on the real linear space $\mathscr{G}(\Omega)$ of *all bounded real-valued functions* on a state space Ω can always be seen as an expectation operator E with respect to a (finitely additive) probability measure on that space: the restrictions of the coherent prevision to the indicator functions associated with the subsets of the state space; see for instance Refs. [2, 46, 50]. This is the way that such probability measures can be made to help in characterising preferences, as is made clear in Eq. (7). But in this characterisation of preferences, the role of probabilities is — to repeat what we said above — merely derivative.

This derivative character of probabilities becomes all the more apparent in the *more specific* present context, where the coherent previsions are defined on the real linear space $\mathscr{H}$ of *all measurements* $\hat{A}$, or equivalently, on the real linear space $\mathscr{U}$ of all bounded real-valued uncertain rewards $u_{\hat{A}}$ on the state space $\bar{\mathscr{X}}$. This $\mathscr{U}$ is only a linear subspace of the linear space $\mathscr{G}(\bar{\mathscr{X}})$ of all *bounded* real-valued maps on $\bar{\mathscr{X}}$: the uncertain rewards $u_{\hat{A}}$ are 'quadratic' functions of the system state and can in no way be associated with indicators of subsets of the state space $\bar{\mathscr{X}}$, as $\mathscr{U}$ contains no indicator functions. The coherence conditions make sure that there are probability measures on the state space $\bar{\mathscr{X}}$ such that a given coherent prevision Λ on $\mathscr{H}$ coincides on the quadratic functions in $\mathscr{U}$ with the expectation operator associated with that probability measure, but these probability measures are in no way necessary to deal with the decision problem: the coherent prevision Λ on $\mathscr{H}$ suffices and the role of probabilities is, it bears repeating, derivative. Due to the nature of quantum-mechanical decision problems, probabilities have no foundational part in their treatment — unless a confusing one — but coherent (lower) previsions on the real Hilbert space of all measurements, and equivalently, (sets of) density operators, do. On this view, probabilities are mere artefacts that turn out to appear and be useful in some decision problems, but not in others.

Incidentally, some of the stranger problems associated with using classical probabilities in quantum mechanics, such as the violation of Bell's inequalities, or their incompatibility with the Tsirelson bound, can be simply explained (away) by taking into account the consequences of the geometry of the Hilbert space of measurements, as is elaborately discussed and argued in Ref. [31, Secs. 2.6 and 3.4]. It turns out that we've been doing this here by considering specific price functionals — coherent (lower) previsions — *on this Hilbert space*. Working with coherent (lower) previsions on measurements, rather than classical probabilities on the state space, also fits in perfectly with the geometric focus in the work of de Finetti [14] and Fisher [29] and solves the above-mentioned problems. It seems to us one of the simplest ways of justifying, and working with, 'quantum probabilities'.

Finally, we want to point out that when we use coherent sets of desirable *measurements* to express Your uncertainty, we're essentially relying on the Heisenberg picture of quantum mechanics. And, since density operators are mixed quantum *states*, we see that sets of density operators fit well within the Schrödinger picture. The (almost-)equivalence between working with *sets of measurements* on the one hand in Section 4.2, and working with *sets of density operators* on the other in Sections 5 and 6, allows us to recover the well-known duality between the Heisenberg and Schrödinger pictures in our decision-theoretic approach as well.

Our argumentation also enables us to get rid of the totality requirement on Your preference ordering whenever You're uncertain about the state of the quantum system. It therefore allows for a more general treatment of quantum-mechanical uncertainty, that still gets back to Born's rule under the right (and usual) circumstances; a treatment of uncertainty that fits in perfectly with the more recent developments in the field of imprecise probabilities, that allows us to deal with partial probability assessments and that looks at probabilistic inference as a special case of conservative (deductive) inference. To put it very succinctly, on our way of looking at it in this paper, quantum mechanics doesn't enforce a probability model for the uncertainty, as all the peculiar quantum mechanical aspects of the associated decision problem are captured in the reward functions and Your preference orderings can therefore be allowed to be only partial. This leaves us with a lot of freedom to accommodate for other uncertainty models than precise probability models, which could (but needn't) be seen as partial and less perfect approximations to these precise models.

We've already taken the opportunity to point out reasons why such more general models can be useful: they can be more realistic for a You with limited time, resources and information, and who therefore can't be expected to come up with a total ordering of the infinity of acts associated with the decision problem; and the conservative inference mechanism associated with them allows You to mathematically draw all the inferences that can be drawn from the preferences You've actually expressed; see, for instance, the various instalments of our qubit running example.

But, there are, next to these more fundamental reasons for using more general models, also practical ones. Letting go of precision (or totality, or completeness) can make a problem that is computationally too complex or expensive more manageable. One interesting case in point is lumping, where a reduction of the dimension of the state space reduces the computational complexity of making inferences about a quantum system, but introduces imprecision. Working with imprecise probability models and techniques will then allow us to turn a computationally hard problem into a more manageable one, at the cost of our only being able to find conservative bounds on quantities of interest whose exact calculation is too expensive. In this way, our discussion here provides the foundation for practical applications in quantum mechanical inference and for importing into quantum mechanics ideas that have proved useful in classical dynamical systems theory, such as for instance, lumping in Markov chains [25–27]. Before an attempt can be made at using our approach for solving such problems, we must be able to import the essentially static considerations in this paper into a dynamical context, where the quantum state of a system evolves in time. This is the subject of our current research.

There are a number of issues and questions that remain, and which we feel to be worthy of attention. First, it might be argued that our approach still isn't general enough, in that we still, in our decision-theoretic background assumption DTB1, assume the existence of a reward function $u_{\hat{A}}$, which *in effect* still imposes a *total* strict ordering on measurements when the system is in a known state $|\Psi\rangle = |\psi\rangle$, because then $\hat{A} \rhd \hat{B} \Leftrightarrow u_{\hat{A}}(|\psi\rangle) > u_{\hat{B}}(|\psi\rangle)$. Would it be possible to let go of even this *a priori* assumption and to then try and recover it *a posteriori* on the basis of other decision-theoretic assumptions, in the spirit of our postulates RF1–RF4? A second question is whether similar conclusions to ours, and similar justifications for using imprecise probabilities in quantum mechanics can be reached when working, as Savage (and Wallace) did, in a decision-theoretic context where the utilities aren't supposed to exist as extraneous to the decision problem, but have to be constructed from the (partial) preference relation on acts.

To conclude, we've already pointed out on several occasions that our work here can be used to provide a decision-theoretic foundation to ideas about introducing imprecise probability models (desirable gambles, lower and upper previsions, sets of density operators) in quantum mechanics, as first proposed by Benavoli, Facchini and Zaffalon in Ref. [7]. As pointed out in a number of relevant places in Sections 4 to 6, we recover a number of their results, even if our interpretation of them may differ. They have developed their ideas further in a different direction in Ref. [6], where they provide an alternative way of thinking about, and justifying, imprecise probability models in quantum mechanics. In contrast to what we do here, they don't follow the practice, standard in quantum mechanics, of using the tensor product space $\otimes_{k=1}^{m}\mathscr{X}_k$ of the particle state spaces $\mathscr{X}_k$ to represent the system state. Instead, they essentially use only a subset of this space, namely the Cartesian product $\times_{k=1}^{m}\mathscr{X}_k$. On this alternative but smaller set of quantum states of the type $x := (|\phi_1\rangle, \dots, |\phi_m\rangle) \in \times_{k=1}^{m}\mathscr{X}_k$ they consider all quadratic gambles, which are defined as functions of the form

$$g_{\hat{A}}(|\phi_1\rangle, \dots, |\phi_m\rangle) := \big(\otimes_{k=1}^{m}\langle\phi_k|\big)\hat{A}\big(\otimes_{k=1}^{m}|\phi_k\rangle\big),$$

also symbolically written as $g_{\hat{A}}(x) = x^{\dagger}\hat{A}x$, corresponding to the Hermitian operators $\hat{A}$ on $\otimes_{k=1}^{m}\mathscr{X}_k$. They then introduce a concept of algorithmic rationality that leads to coherence axioms and a framework of desirability that closely resembles ours in spirit, but is rather different in the mathematical details. Their gambles $g_{\hat{A}}$ correspond to our reward functions, but essentially restricted to the smaller Cartesian product; it's easy to see that there's a one-to-one correspondence between them, as they're both isomorphic to the space of Hermitian operators.[32] They furthermore take the quadratic shape of their gambles as a

[32] When there's only one particle, the state space will be identical in both frameworks.

given, whereas we derive it from a set of postulates. We therefore feel justified in claiming that our approach here is rather different and deserving of separate consideration.

AUTHOR CONTRIBUTIONS

This paper originated in many intensive discussions amongst the four of us and Natan T'Joens in the context of Keano De Vos's master dissertation, in the late autumn and winter of 2020 and early spring of 2021, during the second Covid-19 lockdown. Keano wrote a first draft of the paper, which was later extensively revised, expanded, reorganised and added to by Gert, while Alexander and Jasper focused on thoroughly reviewing and commenting on the intermediate and final stages and showering Keano and Gert with constructive criticism.

ACKNOWLEDGMENTS

Gert's research on this paper was partly supported by a sabbatical grant from Ghent University, and from the FWO, reference number K801523N. He is grateful to Jason Konek for funding several short research stays, as well as a one-month sabbatical stay, at University of Bristol's Department of Philosophy, which allowed him to work on this topic and present results to various people in the department, in the context of Jason Konek's ERC Starting Grant "Epistemic Utility for Imprecise Probability" under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852677). He also wishes to express his gratitude to Teddy Seidenfeld, whose generous funding helped realise a sabbatical stay at Carnegie Mellon University's Department of Philosophy, during which some of the revision work for this paper was done.

REFERENCES

[1] Francis J. Anscombe and Robert J. Aumann. 'A definition of subjective probability'. In: *The Annals of Mathematical Statistics* 34 (1963), pp. 199–2005.

[2] Thomas Augustin et al., eds. *Introduction to Imprecise Probabilities*. John Wiley & Sons, 2014.

[3] Robert J. Aumann. 'Utility theory without the completeness axiom'. In: *Econometrica* 30 (1962), pp. 445–462.

[4] Robert J. Aumann. 'Utility theory without the completeness axiom: a correction'. In: *Econometrica* 32 (1964), pp. 210–212.

[5] Howard Barnum et al. 'Quantum probability from decision theory?' In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 456 (2000), pp. 1175–1182.

[6] Alessio Benavoli, Alessandro Facchini and Marco Zaffalon. 'Computational complexity and the nature of quantum mechanics'. arXiv: 1902.04569. 2019.

[7] Alessio Benavoli, Alessandro Facchini and Marco Zaffalon. 'Quantum mechanics: the Bayesian theory generalised to the space of Hermitian matrices'. In: *Physical Review A* 94 (2016). arXiv: `1605.08177 [quant-ph]`.

[8] Stephen Boyd and Lieven Vandenberghe. 'Convex optimization'. English. In: *Cambridge University Press* (2004).

[9] Claude Cohen-Tannoudji, Frank Laloe and Bernard Diu. *Quantum Mechanics*. Vol. 1. Wiley, 1977.

[10] Ines Couso and Serafín Moral. 'Sets of desirable gambles: Conditioning, representation, and precise probabilities'. In: *International Journal of Approximate Reasoning* 52 (2011).

[11] Jasper De Bock. 'Credal Networks under Epistemic Irrelevance: Theory and Algorithms'. PhD thesis. Ghent University, Faculty of Engineering and Architecture, 2015.

[12] Jasper De Bock and Gert de Cooman. 'A desirability-based axiomatisation for coherent choice functions'. In: *Uncertainty Modelling in Data Science (Proceedings of SMPS 2018)*. 2018, pp. 46–53. Extended arXiv version: `1806.01044`.

[13] Jasper De Bock and Gert de Cooman. 'Interpreting, axiomatising and representing coherent choice functions in terms of desirability'. In: *Proceedings of Machine Learning Research* 103 (2019), pp. 125–134. Extended arXiv version: `1903.00336`.

[14] Bruno de Finetti. 'A proposito di correlazione'. In: *Economia* 3 (1937).

[15] Bruno de Finetti. 'La prévision: ses lois logiques, ses sources subjectives'. In: *Annales de l'Institut Henri Poincaré* 7 (1937). English translation in [33], pp. 1–68.

[16] Bruno de Finetti. *Teoria delle Probabilità*. Turin: Einaudi, 1970.

[17] Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*. English translation of [16], two volumes. Chichester: John Wiley & Sons, 1974–1975.

[18] Gert de Cooman. 'Belief models: an order-theoretic investigation'. In: *Annals of Mathematics and Artificial Intelligence* 45 (2005), pp. 5–34.

[19] Gert de Cooman. 'Coherent and Archimedean choice in general Banach spaces'. In: *International Journal of Approximate Reasoning* 140 (2022), pp. 255–281.

[20] Gert de Cooman, Jasper De Bock and Márcio Alves Diniz. 'Coherent predictive inference under exchangeability with imprecise probabilities'. In: *Journal of Artificial Intelligence Research* 52 (2015), pp. 1–95.

[21] Gert de Cooman and Erik Quaeghebeur. 'Exchangeability and sets of desirable gambles'. In: *International Journal of Approximate Reasoning* 53 (2012). Special issue in honour of Henry E. Kyburg, Jr., pp. 363–395.

[22] Gert de Cooman, Arthur Van Camp and Jasper De Bock. 'Desirable sets of things and their logic'. In: *International Journal of Approximate Reasoning* 172 (2024), p. 109241.

[23] David Deutsch. 'Quantum theory of probability and decisions'. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 455 (1999), pp. 3129–3137.

[24] Paul A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 1981.

[25] Alexander Erreygers. 'Markovian Imprecise Jump Processes: Foundations, Algorithms and Applications'. Doctoral dissertation. Ghent University, 2021. Available at `https://users.ugent.be/~aerreyge/`.

[26] Alexander Erreygers and Jasper De Bock. 'Bounding inferences for large-scale continuous-time Markov chains: A new approach based on lumping and imprecise Markov chains'. In: *International Journal of Approximate Reasoning* 115 (2019), pp. 96–133.

[27] Alexander Erreygers et al. 'Imprecise Markov Models for Scalable and Robust Performance Evaluation of Flexi-Grid Spectrum Allocation Policies'. In: *IEEE Transactions on Communications* 66 (2018), pp. 5401–5414.

[28] Jan Faye. 'Copenhagen interpretation of quantum mechanics'. In: *The Stanford Encyclopedia of Philosophy*. Winter 2019. Metaphysics Research Lab, Stanford University, 2019.

[29] Ronald A. Fisher. 'The distribution of the partial correlation coefficient'. In: *Metron* 3 (1924), pp. 329–332.

[30] Christopher A. Fuchs, N. David Mermin and Rüdiger Schack. 'An introduction to QBism with an application to the locality of quantum mechanics'. In: *American Journal of Physics* 82 (2014), pp. 749–754.

[31] Michael Janas, Michael E. Cuffaro and Michel Janssen. *Understanding Quantum Raffles: Quantum Mechanics on an Informational Approach: Structure and Interpretation*. Boston Studies in the Philosophy and History of Science. Springer, 2021.

[32] Alexei Yu Kitaev et al. *Classical and Quantum Computation*. American Mathematical Society, 2002.

[33] Henry E. Kyburg Jr. and Henry E. Smokler, eds. *Studies in Subjective Probability*. Second edition (with new material) 1980. New York: Wiley, 1964.

[34] Isaac Levi. *The Enterprise of Knowledge*. London: MIT Press, 1980.

[35] Serafín Moral. 'Epistemic irrelevance on sets of desirable gambles'. In: *Annals of Mathematics and Artificial Intelligence* 45 (2005), pp. 197–214.

[36] Robert Nau. 'The shape of incomplete preferences'. In: *Annals Of Statistics* 34 (2006), pp. 2430–2448.

[37] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[38] Erik Quaeghebeur. 'Desirability'. In: *Introduction to Imprecise Probabilities*. Ed. by Thomas Augustin et al. John Wiley & Sons, 2014.

[39] Erik Quaeghebeur, Gert de Cooman and Filip Hermans. 'Accept & reject statement-based uncertainty models'. In: *International Journal of Approximate Reasoning* 57 (2015), pp. 69–102.

[40] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics: Functional analysis*. Vol. I. Academic Press, 2003.

[41] Simon Saunders et al. *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford University Press, 2010.

[42] Leonard J. Savage. *The Foundations of Statistics*. Second revised edition, first published 1954. New York: Dover, 1972.

[43] Eric Schechter. *Handbook of Analysis and Its Foundations*. San Diego, CA: Academic Press, 1997.

[44] Teddy Seidenfeld, Mark J. Schervish and Jay B. Kadane. 'A representation of partially ordered preferences'. In: *The Annals of Statistics* 23 (1995). Reprinted in [45], pp. 69–129, pp. 2168–2217.

[45] Teddy Seidenfeld, Mark J. Schervish and Jay B. Kadane. *Rethinking the Foundations of Statistics*. Cambridge: Cambridge University Press, 1999.

[46] Matthias C. M. Troffaes and Gert de Cooman. *Lower Previsions*. Wiley, 2014.

[47] David Wallace. 'A formal proof of the Born rule from decision-theoretic assumptions'. In: *arXiv* (2009). eprint: `0906.2718`.

[48] David Wallace. 'Everettian rationality: defending Deutsch's approach to probability in the Everett interpretation'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 34 (2003), pp. 415–439.

[49] David Wallace. 'Quantum probability from subjective likelihood: Improving on Deutsch's proof of the probability rule'. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 38 (2007), pp. 311–332.

[50] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall, 1991.

[51] Peter Walley. 'Towards a unified theory of imprecise probability'. In: *International Journal of Approximate Reasoning* 24 (2000), pp. 125–148.

[52] Marco Zaffalon and Enrique Miranda. 'Axiomatising Incomplete Preferences through Sets of Desirable Gambles'. In: *Journal Of Artificial Intelligence Research* 60 (2017), pp. 1057–1126.

GHENT UNIVERSITY, FOUNDATIONS LAB FOR IMPRECISE PROBABILITIES, TECHNOLOGIEPARK-ZWIJNAARDE 125, 9052 ZWIJNAARDE, BELGIUM