View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Cronfa at Swansea University

Swansea University
Prifysgol Abertawe

Cronfa
Setting Research Free

Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in :
Journal of Philosophical Logic

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa21659

Paper:

Pelosse, Y. (2015). The Intrinsic Quantum Nature of Nash Equilibrium Mixtures. *Journal of Philosophical Logic*

http://dx.doi.org/10.1007/s10992-015-9349-7

This article is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Authors are personally responsible for adhering to publisher restrictions or conditions. When uploading content they are required to comply with their publisher agreement and the SHERPA RoMEO database to judge whether or not it is copyright safe to add this version of the paper to this repository.
http://www.swansea.ac.uk/iss/researchsupport/cronfa-support/

Noname manuscript No.
(will be inserted by the editor)

The Intrinsic Quantum Nature of Nash Equilibrium Mixtures

Yohan Pelosse

Received: date / Accepted: date

Abstract In classical game theory the idea that players randomize between their actions according to a particular optimal probability distribution has always been viewed as puzzling. In this paper, we establish a fundamental connection between n-person normal form games and quantum mechanics (QM), which eliminates the conceptual problems of these random strategies. While the two theories have been regarded as distinct, our main theorem proves that if we do not give any other piece of information to a player in a game, than the payoff matrix—the axiom of "no-supplementary data" holds—then the state of mind of a rational player is algebraically isomorphic to a pure quantum state. The "no supplementary data" axiom is captured in a Lukasiewicz's three-valued Kripke semantics wherein statements about whether a strategy or a belief of a player is rational are initially indeterminate i.e. neither true, nor false. As a corollary, we show that in a mixed Nash equilibrium, the knowledge structure of a player implies that probabilities must verify the standard "Born rule" postulate of QM. The puzzling "indifference condition" wherein each player must be rationally indifferent between all the pure actions of the support of his equilibrium strategy is resolved by his state of mind being described by a "quantum superposition" prior a player is asked to make a definite choice in a "measurement". Finally, these results demonstrate that there is an intrinsic limitation to the predictions of game theory, on a par with the "irreducible randomness" of quantum physics.
Keywords: Modal logic· Non-cooperative games· Quantum state· Born rule· Kripke models· Lukasiewicz's three-valued-logic

Y. Pelosse (✉).
School of Management, Swansea University, UK
E-mail: yohanpelosse@googlemail.com

1 Introduction

The notion of mixed strategy, as originally introduced by von Neumann and Morgenstern [1] is a basic ingredient of classical game theory. Yet, as pointed out by von Neumann and Morgerstern themselves, the idea that a rational player may have to use a randomizing device, such as a coin flip, to decide on his actions poses some insuperable conceptual difficulties:

"This is certainly no maximization problem, but a peculiar and disconcerting mixture of several conflicting maximum problems (...) we face here and now a really conceptual—and not merely technical—difficulty. And it is this problem which the theory of "games of strategy" is mainly devised to meet." (von Neumann and Morgenstern [1])

As acknowledged in the subsequent game-theoretic literature (see, e.g. Aumann [2]), the introduction of the Nash equilibrium (Nash, [3]) has just rendered this puzzle even more unpalatable. The very existence of a Nash equilibrium relies on the use of such "randomized strategies", which leads to the following conundrum: In equilibrium, each player has to be completely indifferent between the different actions of his mixed strategy. Moreover, the idea that human behavior, under some conditions, must appear indeterminate in order to be rational is rather troubling. Although many attempts to explain the underpinnings of mixed strategies have been made, none has been unanimously accepted as satisfactory (see e.g. Aumann, [2]). [1] The main goal of this paper is to give a complete characterization and a compelling interpretation of mixtures in games. Surprisingly, we shall prove that probabilities arising in any mixed Nash equilibrium are (necessarily) quantum-mechanical. From a purely game-theoretic perspective, one of the main consequence of this result is the elimination of all the aforementioned conceptual difficulties associated to such randomized strategies. So this raises a key issue: How is this foundational game-theoretic problem so connected to the apparently far removed field of quantum physics?

Unlike classical physics, the most troubling aspect when moving to the microscopic world is that quantum theory prescribes only the probabilities with which the various outcomes of an *individual* quantum system occur in an experiment, and is silent about the outcomes themselves.[2] However, apart from their probabilistic nature, conventional wisdom is that probabilities arising in game theory and QM have very different properties. In particular, there are by now ample experimental evidences supporting the view that we cannot interpret quantum probabilities as coming from certain statistical properties of an ensemble of similarly prepared systems, but each *individual* particle (e.g. a photon in the classic double-slit experiment) exhibits an irreducible random

[1] The alternative foundation—the so-called mass action approach—avoids becoming entangled in such philosophical issues. Rather than considering randomizations implemented by individual players, it addresses the question of how evolutionary selection processes or social learning allow to understand an equilibrium as an aggregate statistical behavior.

[2] A standard textbook on QM is Sakurai [4].

behavior. [3] The fact that quantum probabilities cannot be given a statistical interpretation is also the consequence of a series of mathematical results (see e.g. Kochen and Specker, [6]; Colbeck and Renner, [7, 8]).[4] By contrast, in the epistemic perspective of game theory, mixed strategies are usually interpreted as the expression of the players' ignorance about the actions of their opponents.[5] Hence it seems that we can definitively rule out any structural connection between the two theories: Unlike quantum probabilities, the standard view of mixtures in games is that probabilities are classical. Our main result proves that such an interpretation does not hold if one explicitly incorporates the fact that the payoff matrix alone does not generally contain enough information to decide which strategy constitutes a rational choice.

More precisely, our starting point is the observation that if we do not include some extra variables e.g. some payoff-irrelevant signals or hierarchies of beliefs, in the original description of the game, then a player has no other options than "putting himself into the shoes of the other players" i.e. adopts the decision problems of the other players, in order to determine his *own* rational choice. It then becomes a mere internal consistency requirement that the *same* player will have to determine an entire profile of beliefs (the choices for the others) and choice that are mutually rational to each others. In other words, *without any supplementary data added to the game, any rational choice of one player must therefore be part of a Nash equilibrium of the original game.* [6]

This paper aims at studying the implications of this *one-person* "ontological foundation" of the Nash equilibrium concept on the interpretation of mixed strategies. More specifically, the bulk of this paper consists in uncovering a structural connection—an isomorphism—between the knowledge structure of a single player facing a game without any other piece of information than the payoff matrix on one hand, and the algebraic structure of an individual quantum system, on the other hand.

The implications of the quantum nature of equilibrium mixtures for our understanding of players' behavior in a game are manifold. First, the apparently exotic mathematical tools of QM e.g. "wave function"—a unit vector of a complex Hilbert space—solve the long-standing puzzling paradoxes of the classical game model, such as the famous "mixing problem" i.e. the fact that in equilibrium each player is indifferent between all the pure strategies of his support. In fact, the wave function of QM allows to give a clear-cut interpretation to the famous "indifference condition" as the sign that a particular pure choice has not yet been settled in the player's mind *prior* he makes an actual choice. In other words, a player's state of mind is described by a *quantum superposition*

---

[3] For an experimental evidence of the quantum nature of a single photon see e.g. Mandel [5].

[4] In short, Colbeck and Renner show that a subjective interpretation of the wave function is untenable since quantum theory is maximally informative, which means that the quantum-mechanical probability amplitudes of waves functions are *real physical properties.*

[5] For a general overview of the epistemic analysis of games see Brandenburger [9, 10].

[6] The intuition and the proof of this result—which has been thoroughly analyzed in a separate paper (Pelosse, 2011)—is reproduced in the Appendix of this paper.

before he makes an actual choice, because the "indifference condition" implies that the *linear combination* of any two states of mind corresponding to the choice of two pure actions must also be a possible state of mind. One of the most important consequence is then to offer a fundamental theoretical explanation for players to be indeterminate. This notably allows to understand why the usual equilibrium probabilities (mixtures) *cannot* be construed in terms of a randomization or as reflecting the ignorance of some outside observer.
Besides, it is well-known that these "wave functions" induce non-additive (quantum) probabilities (see e.g. Feynman et al. [11]). Hence, our result indicates that the classical game model generates *non-additive* (quantum) probabilities in its own right. In a nutshell, the existence of non-additive probabilities is the direct consequence of the player's "self-interaction": *before* the actual choice (of a pure strategy), the equilibrium state of mind of a player is not fixed on a particular pure action, and this "rational blend of states of mind" generates some "interference terms" as in QM. We note that this provides a sound theoretical justification to a growing number of probabilistic frameworks for modeling decision-making that make use of the quantum formalism in cognitive situations (see e.g. Khrennikov [12]), to capture incompatibility and interference effects that arise in human decisions.
A last series of results concern the structure of the informational states of a rational player making a choice. [7] First, we prove that probabilities arising in an equilibrium mixture automatically satisfy the Born rule of QM—one of the key postulate of QM. This rule allows to uncover the informational mental process of a rational player making a choice during concrete experiments. Second, the salient difference between classical and quantum mechanics is the transition from a classical event space, represented by the Boolean algebra of (Borel) subsets of a phase space, to a non-Boolean algebra. Our results indeed demonstrate that the algebra of epistemic states of a rational player making a choice (in a Nash equilibrium) generates exactly the same non-Boolean structure.
So these results raise a key question: Where does such a structural connection between classical game model and quantum mechanics come from?
As already hinted, the thread that connects the two theories can be traced back from the initial under-determination of the game model i.e. the matrix of a game does not generally contain enough information to make a rational choice. By incorporating this initial structural lack of information—using the Lukasiewicz's three-valued logic [13]—, we are able to analyze the origin and genesis of a rational choice of a player in a game. We remark that our main result is consistent with the aforementioned series of well-known results in QM—the so-called "no-go Theorems". Briefly stated, these results suggest that

[7] We use the term "informational state" rather than epistemic because under the quantum nature of probabilities, the usual distinction between the "ontic" and "epistemic" state of mind of a player breaks down. We discuss this conceptual distinction at length in Section 6.2.

measurements do *not* reveal the pre-existing properties of a quantum system.[8] For example, the Kochen-Specker Theorem [6] and the Conway-Kochen [14, 15] Theorems notably establish—modulo some technical requirements—that the knowledge of a particle's properties has nothing to do with measurement disturbances, as it is just that there is no way to assign pre-existing values to a particle prior a measurement in a consistent logical way.[9] These results are in line with the interpretations of Wheeler [16], or more recently Zurek [17] and Gisin (2010) that interpret individual process in quantum mechanics as an "*an elementary act of creation.*"[10] In this paper we do not attempt to make any claim about the interpretation of quantum mechanics itself. However we note that our findings are at least consistent with these results. So, from a purely game-theoretic perspective, the origin of the probabilities arising in the classical game model is *not* epistemic i.e. arising from a property unknown to the players or an outside observer—, but is *ontic*, on a par with the irreducible indeterminism of single events of QM.[11] In other words, mixed strategies are *not* the reflection of what the other players (or any outside observer) know about a player state of mind before his choice.[12]

Different streams of papers see e.g. Eisert et al. [19], have introduced formal tools of QM in decision/game theory in both the physics and decision theory communities where new strategies are "superpositions" of the underlying classical strategies. In short, strategies which are quantum-mechanical in nature are added to the original game. This paper argues instead that classical game theory is already quantum-mechanical in nature.

Only a few papers have tried to make a connection between classical game theory and quantum mechanics. Pietarinen [20, 21] is the first to explicitly raise the possibility of some fundamental connections between extensive form games of imperfect information, quantum logic and quantum mechanics. He notably points out the failure of the law of the excluded middle in games of imperfect information and links it to quantum mechanics and remarks that the truth-values of complex sentences make some games nondetermined. More

[8] Briefly stated, the Kochen-Specker (KS) Theorem is a mathematical result about the nature of Hilbert spaces (the special type of vector spaces that are the most general representation of the state space for a quantum system). It says that, if properties are represented as operators on a Hilbert space in a $1-1$ fashion (i.e. each property is represented by a unique operator), then these properties cannot all be said to simultaneously have values.

[9] For example, a logical possible implication of Bell inequality, is that either the inputs are not real, or the outcomes (or both). A loophole-free violation of Bell inequality would prove the impossibility of "*local realism*" i.e. the facts that measurements does not reveal pre-existing properties. On the other hand, the Kochen-Specker theorem [6] indicates that "*For any physical system, in any state, there exist a finite set of observables such that it is impossible to pre-assign them non contextual values respecting the predictions of QM*".

[10] As explained by Zurek, the wave function $|\Psi\rangle$ is an informational real entity: "*Thus,* $|\Psi\rangle$ *is in part information (as, indeed, Bohr thought), but also the obvious quantum object to explain existence*".

[11] To put it differently, "quantum events are not functions of space-time" (see Gisin (2010) and [18]).

[12] In particular, as argued by Gisin ((2010) and [18]) when discussing of quantum probabilities: "*One can't simulate these probabilities on a classical computer*".

recently, Brandenburger [22] establishes a formal connection between classical game theory with the non-local correlations arising in QM: He proves that adding quantum signals does not necessarily differ from the addition of classical signals. [13] This paper therefore responds and complements these papers in several ways by proving that classical game theory is structurally quantum-mechanical. In particular, in a separate paper (2014) we show how the quantum nature of Nash equilibrium leads to an intrinsic one-person representation of a game.

QM is not based on a generally accepted conceptual foundation (see e.g. Zeilinger [23]). The field of quantum foundations seeks to determine the principles that underlie quantum theory (see Hardy and Spekkens [24]). Thus, at a broad level, there is a connection between this paper and the possibility of deriving or interpreting some of the quantum formalism from decision or information-theoretic axioms as explored in the physics community (see e.g. Deutsch [25], Zeilinger [23]).[14] A derivation of quantum theory is outside the scope of this paper. We have done so in a separate paper (2012) by uncovering the intrinsic game-theoretic structure of any individual quantum system.

The paper is organized as follows. The next section is an informal discussion of the issues and results to follow. The formal treatment is in Sections 3–6. Section 7 concludes. All the proofs have been relegated to an Appendix.

2 Previews of the main result

To fix ideas, consider the game of Figure 1. In this game, Ann chooses the row, Bob chooses the column, Charlie chooses the matrix. Each player must (simultaneously and independently) choose to go to the North Pole or the South Pole. For Ann and Bob, this is simply the "Battle of the Sexes": their payoffs are no affected by Charlie's choice. Our main postulate is that Ann, Bob and Charlie are all assumed to have *no* other piece of information than the matrix game itself i.e. the game satisfies the axiom of "`no-supplementary data`". We shall give a formal definition of this postulate in the formal treatment. For the sake of discussion, let us focus on the following Nash equilibrium of this game $(\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}, \frac{1}{3}\mathtt{n} \oplus \frac{2}{3}\mathtt{s}, N)$. Consider the implication of formally incorporating the aforementioned axiom of `no-supplementary data` on the behavior of each rational player. Let us focus our discussion on Ann.

First, note that if Ann is rational, but has no supplementary data to determine her choice than the matrix game, then she really has to adopt Bob and Charlie's decision problems in order to form her beliefs and compute her own best response. In other words, she concretely has to "put herself into the shoes" of Bob and Charlie and at the same time, consider her own decision problem.

[13] Brandenburger (2010) offers an insightful discussion on the relations between the epistemic program in game theory and the under-determination of the classical game model as initially envisioned by von Neumann.

[14] In particular, note that Deutsch is the first to propose a decision-theoretic derivation of the Born rule in a QM setup.

Since Ann is rational, she must be rational in all the decision problems she is simultaneously considering: In her own perspective (A_A), as well as when she is considering Bob and Charlie's decision problems at perspective $A_{B\otimes C}$. So Ann must *simultaneously* put herself, "in the shoes" of Bob and Charlie (the meta-perspective $A_{B\otimes C}$) *and* in her own shoes (her own perspective A_A). Here is a visual way to grasp this situation more clearly.
Imagine that Ann is sitting simultaneously in two different *transparent* "cubicles" (the term is taken from Kohlberg and Mertens [27, p.1005]) A_A and $A_{B\otimes C}$. Below, we illustrate the formal mental process occurring in the mind of Ann when her beliefs over Bob and Charlie's (independent) destinations are initially non existent. Start by identifying the directed graph

$$\longleftarrow\!\!\circ \;\; A_A$$

as describing the informational process "input-output" by which Ann can determine her rational strategy (the output) given that she has determined her beliefs over Bob and Charlie (the input). For short, let $\multimap$ denote this mental process. The converse process is Ann determining her beliefs i.e. a profile of a mutually rational profile of strategies for Bob and Charlie. Hence, the "dual" directed graph

$$A_{B\otimes C}\circ\!\!\longrightarrow$$

corresponds to Ann determining the truth value of the *input* of the first process. For short, we represent this "dual" mental process by $(\multimap)^{\dagger}$ i.e. the "dual" directed graph, representing the converse direction, output-input. Note that $\dagger$ is the operation which changes the direction of the mental process. At such, it corresponds to an involution in the sense that $(\multimap)^{\dagger\dagger} = \multimap$. As shown in Theorem 1, this "duality property" of the graph is not accidental; it is directly linked to the dual algebraic structure between vector and co-vectors (or linear mappings) of the Hilbert space structure of QM. Now note that the above processes allow Ann to determine that the statement "$\frac{1}{3}\mathtt{n} \oplus \frac{2}{3}\mathtt{s}$ *is rational for Bob and the South Pole is rational for Charlie*" is true in her cubicle $A_{B\otimes C}$, because Ann can check—by taking a look at cubicle A_A—that the statement "$\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ *is rational for Ann*" is indeed true in the corresponding cubicle, A_A, relative to the statement that "$\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ *is rational for Bob and the South Pole is rational for Charlie*" is true in cubicle $A_{B\otimes C}$, whose she knows to be true by looking from her cubicle A_A, and so on. This simple observation leads us to conclude that the only possible process for Ann to determine a rational strategy in her mind without any supplementary data is to combine the two processes *simultaneously*,

$$A_{B\otimes C}\circ\!\!\longrightarrow \qquad \bigwedge \qquad \longleftarrow\!\!\circ \;\; A_A \equiv \qquad A_{B\otimes C}\circ\!\!-\!\!\!-\!\!\circ \;\; A_A.$$

Moreover, since Ann is a single rational entity, she has to be rational in these two processes, which immediately implies that the above process corresponds to Ann determining a Nash equilibrium of the game of Fig. 1. As a result, the determination of a rational strategy is equivalent to Ann "self-interacting" in a

Table 1 Figure 1.

	n	s
N	2,1,0	0,0,0
S	0,0,0	1,2,0

N

	n	s
N	2,1,0	0,0,0
S	0,0,0	1,2,0

S

mixed Nash equilibrium of the game of Fig.1. The general proof of this result is established in the appendix.[15] Notice that Ann could arrive at a similar result by sitting in the three distinct glass cubicles, A_A, A_B and A_C, instead of the big cubicle $A_{B\otimes C}$. However, in our main theorem we prove that models of "introspection" with more than two perspectives are in fact not well-defined. The reason is that these models would not induce Ann to play a well-defined probability measure—her equilibrium mixture $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$—in an experiment. Alternatively, we can also examine the situation of Ann if she adopts (simultaneously) the perspectives of all players, $A_{A\otimes B\otimes C}$. In this case, Ann is sitting in her "big cubicle" in which she simultaneously assigns truth-values to statements for all players. By definition, in this big glass cubicle she can only make *absolute* statements about all players. But Ann is unable to check—by looking through the windows of her cubicle—that a given meta-statement, i.e. a profile of strategies, is indeed rational relative to itself. For this she needs the perspective of another cubicle. Formally, Ann cannot check that a given profile is indeed a fixed point of the combined best-response mapping.

To recap: Without the aid of any extra variables, Ann can only determine a rational strategy by interacting with herself (self-interacting) in a Nash equilibrium. What are the implications of this "self-interactive" rational process in the interpretation of a mixed Nash equilibrium?

Consider the mixed-strategy equilibria in the game of Figure 1. In equilibrium, Ann chooses the North Pole ($\mathtt{N}$) with a probability $\frac{2}{3}$ and the South Pole ($\mathtt{S}$) with probability $\frac{1}{3}$. Suppose that an outside observer can read in the mind of Ann. What would he observe *before* Ann chooses a destination (before she makes her actual choice)?

As argued above, without any supplementary data in her mind, Ann's introspection corresponds to the determination of a Nash equilibrium. By looking simultaneously through the transparent walls of her two cubicles—during her introspection—Ann therefore has been able to determine that:

(1)"*Strategy* $\frac{2}{3}N \oplus \frac{1}{3}S$ *is rational for Ann*" is true at A_A and;

(2)"*Strategy* $\frac{2}{3}N \oplus \frac{1}{3}S$ *is rational for Ann*" is true at $A_{B\otimes C}$.

In other words, Ann thinks that strategy $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ is rational for her in *each* of her two cubicles ((meta)-perspectives). Of course, this is just the usual definition of an equilibrium: The common belief of Bob and Charlie at $A_{B\otimes C}$ meets the mixed strategy of Ann. However, the above introspection now entails that $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ is in fact Ann's *self-referential belief.* As shown in the rest of this

[15] A complete analysis of the foundation of Nash equilibrium has been carried out in a separate paper (see Pelosse, 2011).

paper, *this is precisely the very existence of such self-referential beliefs that are at the origin of players' quantum states of mind* in a game. To see this here notice that Ann deems equally rational $\mathtt{N}$ and $\mathtt{S}$ *if and only if* the (common) belief she holds when determining the others' player strategies matches his own mixed strategy. Thus, by construction, Ann can hold $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ as a self-referential belief *if and only if* she has not already made-up her mind on $\mathtt{N}$ or $\mathtt{S}$. We shall now show that the algebraic characterization of this indifference is given by a pure quantum state. To see this, let us first point out that Ann's state of mind should respect the following minimal reasonable properties:
[**Indivisibility**] Ann is a *single* player (she is in the flesh!);
[**Independence**] A_A and $A_{B\otimes C}$ are mutually *independent* perspectives (in her mind, Ann determines simultaneously each statement in two *different* "cubicles"). Moreover, at A_A, Ann determines her *mixed strategy*, while she determines the *common belief* of Bob and Charlie at $A_{B\otimes C}$;
[**Faithful**] Operationally, if an outside observer runs an experiment with many copies of Ann, the empirical frequencies of outcomes $\mathtt{N}$ and $\mathtt{S}$ must agree with what each copy of Ann has determined in her mind as a rational strategy.
The axiom of Indivisibility is just the fact that in a measurement Ann has to choose a *unique* action. Transported *in the mind* of Ann, this means that a choice occurs if and only if the pole chosen by Ann at one of her two cubicles matches the other. So we can visually describe a choice as the situation where Ann picks a choice in each of her two cubicles, A_A and $A_{B\otimes C}$, while looking simultaneously through the windows of her cubicles.
The axiom of independence simply says that Ann remains seated in her two cubicles, A_A and $A_{B\otimes C}$. The last axiom requires that in a measurement i.e. a pure choice, the empirical probability measure obtained by an observer reflects exactly Ann's equilibrium mixed strategy, $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$.
Now we return to the description of the state of mind of Ann. First, recall that in (any) mixed equilibrium, Ann must be indifferent between going either to the North Pole or to the South Pole. This a general result: *If a strategy profile, $(\sigma^i)_{i\in N}$, is a Nash equilibrium, then every pure strategy in the support of each strategy σ^i has to be a best reply to i's belief σ^{-i}*. (see e.g. Osborne and Rubinstein, [26]). Taken together, the axioms entail that Ann may think in her two cubicles that the choice of the North Pole and the South Pole are rational. The upshot is that the state of mind of Ann will be given by a 2×2 matrix (see Figure 2), which is the simplest example of the so-called density matrices of QM characterizing a pure quantum state (qubit). What are the entries of this matrix?
The faithful property implies that the diagonal elements, $(\mathtt{N},\mathtt{N})$ and $(\mathtt{S},\mathtt{S})$, take on values $\frac{2}{3}$ and $\frac{1}{3}$, respectively, together with the property that the vector representing Ann when she is in cubicle A_A can be distinguished from her vector representing her cubicle $A_{B\otimes C}$. Thus, by the axiom of independence, we conclude that complex vectors like, e.g. $(\frac{1}{\sqrt{6}} \pm i\frac{1}{\sqrt{2}})\mathtt{N} \oplus (\frac{1}{2\sqrt{3}} \pm i\frac{1}{2})\mathtt{S} := z_1\mathtt{N} \oplus z_2\mathtt{S}$ characterize the state of mind of Ann (her mixed strategy) when she is in perspective A_A, while its conjugate, $(\frac{1}{\sqrt{6}} \mp i\frac{1}{\sqrt{2}})\mathtt{N} \oplus (\frac{1}{2\sqrt{3}} \mp i\frac{1}{2})\mathtt{S} := \overline{z}_1\mathtt{N} \oplus \overline{z}_2\mathtt{S}$

Table 2 Figure 2. Ann's equilibrium state of mind

	N	S
N	$\frac{2}{3}$	$z_1\bar{z}_2$
S	$z_2\bar{z}_1$	$\frac{1}{3}$

reflects her state of mind (the common belief of Bob and Charlie) at $A_{B\otimes C}$. This emergence of complex numbers is a general result: As shown in the formal treatment, a faithful algebraic characterization —an isomorphism—of a pair of a player's self-referential belief involves the complex field. How do we interpret these vectors?

The matrix of Figure 2 represents the global structure of knowledge of Ann. For example, the complex weight $z_1\bar{z}_2$ assigned to entry $(\mathtt{N}, \mathtt{S})$ represents the fact that Ann has not yet fixed her mind on `N` or `S`: When sitting in cubicle A_A she knows that going to the North Pole is rational but she simultaneously knows at the other cubicle, $A_{B\otimes C}$, that the South pole is also rational. Thus, the matrix of Figure 2 is the modal expression of Ann's *rational indifference* and hence "self-rational ignorance" of `N` or `S` prior she is asked to make a choice. So the matrix of Figure 2 is nothing but the algebraic characterization of the (self)-knowledge of Ann satisfying the (necessary and sufficient) "indifference condition" required in a mixed Nash equilibrium. We can now answer the question: What is the nature of equilibrium mixtures in games?

At a broad level, Ann's mixed strategy, $\frac{2}{3}\mathtt{N} \oplus \frac{1}{3}\mathtt{S}$ must be understood as the link between the *mental* reality of Ann's optimal strategy and the *empirical* realization of this state of mind in an experiment. Since Ann is a single player, there must be a perfect (physical) correlation across her perspectives during a choice. Thus, these probabilities represent the weights with which both perspectives simultaneously think of the same (rational) destination.[16]

The upshot is that the epistemic characterization of the "indifference condition" given in the matrix of Figure 2 rules out the "classicality" of probabilities arising in mixed Nash equilibria: Those probabilities can*not* be interpreted as reflecting a randomization from the part of the players, but as a quantum superpositions of states of mind resulting from each player's holding his beliefs about his own future choice or equivalently his own rational indifference amongst different future alternatives Note also that the way a player chooses a particular pure strategy is left outside the scope of the classical game model. This is so since any notion of a "measurement" would require that we extend the game model by incorporating the details of the interaction between a player and its environment (or an outside observer).

The above Gedankenexperiment is the illustration of our main finding–that each player is characterized by a quantum state of mind in a Nash equilib-

[16] Note that a conceptual consequence of this story is that the equilibrium intrinsic state of Ann is an epistemic state that is ontic as it describes complete knowledge i.e. the distinction between an ontic state and an epistemic state breaks down.

rium. Loosely stated, our result is as follows: (Theorem 1)[17]: *Suppose we have a finite n-person game in strategic form where each player is rational and has no extra piece of information than the payoff matrix itself i.e. the axiom of* `no-supplementary data` *holds. Then, in any mixed Nash equilibrium profile* $(\sigma^i)_{i\in N}$ *of the game being played, each player i is described by a superposition of states of mind over the pure strategies* s^i *in the support of* σ^i *i.e. a pure quantum state representing the global state of mind of player i is given by an orthogonal projector (or dyad) with complex off-diagonal terms and the empirical probability of having a pure strategy* s^i *in an experiment is given by the Born rule, as in QM.*

This result offers a completely new way of looking at mixed Nash equilibria. Recall that according to the classical view, a mixed strategy represents deliberate randomizations on the part of players. In this scenario a player commits to a randomization device and delegates the play to a trustworthy party and the *actual* choice is made by the random device. Hence, in the above example, an observer would realize that Ann believes she will go to the North Pole with probability $\frac{2}{3}$ and to the South Pole with probability $\frac{1}{3}$.

The Bayesian approach would suggest a different answer. There, a mixed-strategy equilibrium is interpreted as an expression of what each player believes his opponent will do. So, in either of these interpretations, the "indifference condition" amongst the different actions breaks down: Each player's actual choice has already been taken, which means that the state of reality is already determined prior a player has been asked to make a choice.[18]

The presence of wave probability amplitudes allow to understand why equilibrium mixtures have been so difficult to interpret; these probabilities reflect the necessary physical "act of creation" of the future states of mind of a player (with itself), while classical probabilities proceed from the mere ignorance of the actual state of mind of the player.

Some readers may think that we could bypass the use of complex vectors. However, the algebraic description of the mental states of the player during his introspection requires that we distinguish between the two "flows of knowledge" of the player with itself, which with the property of involution lead to the complex algebra. One may also think of describing the dual mental state of a player with the notion of "bi vector"from geometric algebra instead of complex vectors. However, if possible, this mathematical alternative description of a player's mental state would blur our connection with the standard formulation of QM.[19]

Finally, notice that our connection between the mental state of a player and

[17] This result could be extended to Euclidean games and some other classes of games with infinite strategy spaces.

[18] If a player does have the option of making a randomized choice, this can be added to the (pure) strategy set. Of course a similar interpretation follows in the "mass-action" interpretation of Nash [3], as a mixed-strategy profile is formally identical with a population distribution over the pure strategies.

[19] A bi-vector can be interpreted as a directed number which describes an oriented plane segment, with the direction of the bi-vector representing the oriented plane and the magnitude of the bi vector measuring the area of the plane segment.

the QM state-space is structural i.e. our use of a Kripke structure is sufficiently general to guaranty the robustness of Theorem 1; the self-interaction of a single player leads inevitably to a structure of knowledge—the graph-theoretic representation of the Kripke framework modeling this structure of knowledge—that is *isomorphic* to the dual algebraic structure of a complex Hilbert space.

2.1 A game-theoretic double-slit experiment

Let us now conclude the preview of our results by illustrating how non-additive probabilities enter the picture of the classical game model. We consider the following extension of the previous thought experiment. As before, Ann is still playing the game illustrated by Table 1. But now we append an hypothetical experimental setup in which an experimentalist can read in the mind of Ann *before* she makes a definite choice. More precisely, this "mind reading machine" translates Ann's states of mind onto a fluorescent sphere representing the Earth, so that each time Ann thinks she will go to one of the two poles, her thought is registered as a flash of light on the sphere at one of the continents of the Hemisphere containing the pole chosen by Ann. For instance, a flash light on the Europe is the signature that Ann has fixed her mind on $\mathtt{N}$.[20] Note that if we force Ann to go to say, the South Pole, then we oblige Ann to fix her mind on one of the continents of the Southern Hemisphere, and we see one flash on one of the continents of this Hemisphere. That is, we cannot observe a flash on the Europe. We suppose that the occurrence of a flash on a particular continent lying in the Hemisphere is random. This randomness has nothing to do with the choice of Ann in the game, it simply reflects some additional characteristics of the players, like the nationality, or the culture of Ann, not under the control of the experimentalist. From an *operational* viewpoint, we can construct an empirical probability distribution by counting up the number of flashes in a continent to obtain the fraction of flashes that occur on this continent, when a large number of copies of Ann plays the game. [21] Let $P_{\mathtt{N}}(x)$ (resp. $P_{\mathtt{S}}(y)$) be the probability that we observe a flash of light on a continent $x \in \{\mathtt{Europe}, \mathtt{Africa}, \mathtt{Asia}, \mathtt{America}\}$ (resp. $y \in \{\mathtt{Antartica}, \mathtt{Australia}, \mathtt{Africa}, \mathtt{Asia}, \mathtt{America}\}$) of the Hemisphere when Ann's state of mind is fixed on the North Pole (resp. South Pole). Note that we *cannot directly observe* the state of mind of Ann, because if we knew that Ann's state of mind is fixed on say, the North Pole, then this would imply that Ann knows it either, which would mean that we have forced her to make-up her mind in some way. Otherwise stated, the state of mind of Ann must be treated as a *latent variable*. Now, note that if Ann had fixed her mind on one of the two poles prior making a choice, we would observe for each copy of Ann

[20] It might be worthwhile pointing out that this yet largely hypothetical experiment could be performed in the future by using technologies like brain imaging.

[21] Again, as already discussed, as in QM the use of copies of Ann is purely operational, and the induced empirical distribution cannot be construed as reflecting the various characteristics of players as in evolutionary game theory

a single flash of light either on a region of the continent of the Northern Hemisphere or the Southern Hemisphere. In this case, we would then expect that the probability of observing a flash on a continent z overlapping the two Hemispheres is, $P_{\mathtt{NS}}(z) = P_{\mathtt{N}}(z) + P_{\mathtt{S}}(z)$, for $z \in \{\mathtt{Africa}, \mathtt{Asia}, \mathtt{South\ America}\}$, in conformity with the classical probability calculus. So let us first determine the formulae of these probabilities when Ann is asked to make a choice i.e. an apparatus records the choice of Ann.

If Ann has fixed her mind on one of the poles during an experiment, this is tantamount to saying that she has to pick the *same* destination d in her two perspectives A_A and $A_{B\otimes C}$. Let $\Psi_d(z)$ for $d = \mathtt{N}, \mathtt{S}$, (resp. $\overline{\Psi_{\mathtt{d}}(z)}$) be the complex weight (resp. its conjugate) assigned to the directed graph representing the knowledge structure of Ann at perspective A_A (resp. $A_{B\otimes C}$) when she picks continent z, given that she knows that destination d is rational. The case wherein Ann views the same destination d as rational, simultaneously, in her two perspectives and picks a continent z is illustrated below.

$$\xleftarrow{\Psi_{\mathtt{N}}(z)}\!\circ \;\; A_A \bigwedge \;\; A_{B\otimes C} \circ\!\xrightarrow{\overline{\Psi_{\mathtt{N}}(z)}}$$

From the above illustration we can therefore conclude that if these weights induce the probabilities that given a choice of d, Ann picks a continent z, then $P_{\mathtt{d}}(z) = \Psi_{\mathtt{N}}(z)\overline{\Psi_{\mathtt{N}}(z)}$, which is the Born rule of QM, and the total probability is such that

$$P_{\mathtt{NS}}(z) = |\Psi_{\mathtt{d}}(z)|^2 + |\Psi_{\mathtt{d}}(z)|^2 .$$

Now, we ask: What is the probability of a flash light on a continent z both in the Southern and Northern Hemispheres, if we *do not ask* each copy of Ann to choose one of the two poles in the game, so that we *only* observe the mental state of mind of Ann determining her rational mixed strategy in the game?

In this case, we can only indirectly observe the state of mind of Ann via a flash light on the sphere. According to Theorem 1, the state of mind of Ann prior a measurement has to be described by a wave, as a consequence of the "indifference condition", since Ann's state of mind is not yet settled. Formally, this means that Ann knows *simultaneously* that $\mathtt{N}$ *or* $\mathtt{S}$ (the strategies of Ann, when viewed from her own perspective) are rational at A_A, on one hand, *and* that $\mathtt{N}$ *or* $\mathtt{S}$ (recall that the strategies of Ann coincide with the beliefs of Bob and Charlie when she contemplates the perspective of these players) are rational at $A_{A\otimes B}$, on the other hand. Here is a visual way to derive the probability formula in this situation, in the light of the knowledge structure of Ann.

$$\xleftarrow{\Psi_{\mathtt{S}}(z)}\!\circ \; A_A \;\; \bigvee \;\; \xleftarrow{\Psi_{\mathtt{N}}(z)}\!\circ \; A_A \bigwedge \quad A_{B\otimes C} \circ\!\xrightarrow{\overline{\Psi_{\mathtt{N}}(z)}} A_A \;\; \bigvee \;\; A_{B\otimes C} \circ\!\xrightarrow{\overline{\Psi_{\mathtt{S}}(z)}} A_A .$$

This set of weighted directed graphs depicts the compositions of the different parts of the knowledge structure of Ann. Here, the operation $\bigvee$ corresponds to the usual logical connective "or"; given a perspective, Ann can consider $\mathtt{N}$ *or* $\mathtt{S}$. The operation $\bigwedge$ coincides with the logical connective "and", since Ann has to adopt the two perspectives A_A and $A_{A\otimes B}$, simultaneously, as stated in Theorem 1 (The details of the proof are given in Appendix). As already discussed, we have to assign a complex number $\Psi_d(z)$, for $d = \mathtt{N}, \mathtt{S}$ to each directed graph in order to have a faithful algebraic description of Ann's knowledge structure (an isomorphism). From this, we conclude that the total probability $P(z)$ of a flash light on a continent z both in the Southern and Northern Hemispheres is given by $P(z) = |\Psi_{\mathtt{N}}(z) + \Psi_{\mathtt{S}}(z)|^2$, which coincides with the theory of interference of waves. In fact, we have just derived the formulae used in the classical double-slit experiment in physics. The first formula, $P_{\mathtt{NS}}(z)$, corresponds to the situation where the experimental set-up allows to monitor which slit the electron passed through (we force Ann to make a choice), while $P(z)$ is the correct formula when the experimentalist does not devise an experiment to determine which slit an electron passes through (we do not oblige Ann to "make-up her mind" on one of the two poles). The upshot is thus that as in this classical physics experiment, the predictions on the future behavior of Ann, require the use of non-additive probabilities since $P(z) \neq P_{\mathtt{N}}(z) + P_{\mathtt{S}}(z)$. This inequality is due to the interference term $2\overline{\Psi_{\mathtt{S}}(z)}\Psi_{\mathtt{N}}(z)$. This non-zero additional term is easily derived from the above picture and its meaning is crystal clear: In addition to the cases of perfect correlation between the two perspectives of Ann i.e. given that Ann considers the same pole, she picks the same continent z, we must also add the cases where Ann picks the same continent z, while she is considering $\mathtt{N}$ at one perspectives and $\mathtt{S}$ at the other. The twist is thus that the seemingly exotic non-classical probability calculus of QM arises in the classical game model as the *consequence* of the rationality of a player, via his self-interaction in a Nash equilibrium.

3 Epistemic models of self-projection in games

3.1 Finite games

Let $N = \{1, ..., n\}$ be a finite set of players with $n \geq 2$. An n-person finite normal-form game of complete information, interpreted as a one-shot game is given by $G = \left\langle S^1, ..., S^n; \pi^1, ..., \pi^n \right\rangle$ where each set S^i consists of m_i pure strategies, with typical element, s^i, available to player i, and $\pi^i : \Pi_{i=1}^n S^i \to \mathbb{R}$ is i's utility function. We adopt the convention that $S = \Pi_{i=1}^n S^i$, $S^{-i} = \Pi_{j\neq i} S^j$ and $S^J = \Pi_{j\in J} S^j$ where $J \subset N$. The set of mixed strategies of player i is thus the $(m_i - 1)-$dimensional unit simplex $\Delta^i = \left\{\sigma^i \in \mathbb{R}_+^{m_i} : \sum_{s^i\in S^i} \sigma^i(s^i) = 1\right\}$ and $\Delta = \times_{i=1}^N \Delta^i$ is the polyhedron of mixed-strategy combinations $\sigma = (\sigma^1, ..., \sigma^n)$ in the game. We identify each pure strategy $s^i \in S^i$ with the corresponding unit vector $a_{s^i} \in \Delta^i$. When $J \subset N$, we set $\Delta^J = \times_{j\in J} \Delta^j$ with $\sigma^{-J} := (\sigma^k)_{k\notin J} \in \Delta^J$ and as usual $\Delta^{-i} := \Delta^{N\setminus\{i\}}$. We extend π^i to Δ in the usual way: $\pi^i(\sigma^i, \sigma^{-i}) = \sum_{s^i\in S^i} \sum_{s^{-i}\in S^{-i}} \sigma^i(s^i)\sigma^{-i}(s^{-i})\pi^i(s^i, s^{-i})$ with $\sigma^{-i} \in \Delta^{-i}$. The mapping $\pi^J : S^J \to \mathbb{R}^J$ gives the payoffs of pure strategy

combinations of players $j \in J$, and its extension $\pi^J : \Delta^J \to \mathbb{R}^J$ is defined in the obvious manner. The support of some mixed strategy $\sigma^i \in \Delta^i$ is denoted by $\mathtt{supp}(\sigma^i) = \left\{s^i \in S^i : \sigma^i(s^i) > 0\right\}$.
Last, for each strategy combination $\sigma \in \Delta$,

$$\mathtt{BR}_i(\sigma) = \left\{\sigma^i \in \Delta^i : \pi^i(\sigma^i, \sigma^{-i}) \geq \pi^i(\sigma'^i, \sigma^{-i})\ \forall \sigma'^i \in \Delta^i\right\},$$

denotes the mixed best replies of player $i \in N$. In the rest of the paper we always consider the mixed-strategy extension of G.

3.2 Self-projective Kripke frames

In game theory, the idea that players can shape their beliefs by putting themselves "into the shoes of others" is not new; Luce and Raiffa [28, p.306] were among the earliest to suggest such a process. The Kripke structure defined below aims at giving a formal content to the idea that a player in a game can mentally simulate the choices of the other players by "self-projecting" himself into their decision problems.
Given a game $G = \left\langle S^1, ..., S^n; \pi^1, ..., \pi^n \right\rangle$, a **self-projective frame** for player i in G is the structure, $\mathcal{F}_G^i = \left\langle W, \mathcal{W}, \mathcal{R}^i \right\rangle$, where $W = \left\{w^{i_1}, w^{i_2}, ..., w^{i_n}\right\}$ represents the (non-empty) set of the n **atomic (informational) perspectives** corresponding to the n decision problems that can be considered by player i in the game G. Each w^{i_j} represents player i when he is mentally considering the decision problem—the perspective—of player j.
As already stressed, we study a structure encompassing all the Kripke models that a player could possibly adopt in a game. So, we do not impose any requirement on $\mathcal{W}$ i.e. player i may consider some "non-partitional" perspectives and $\mathcal{W}$ can be any class of m (nonempty) subsets of W whose union is W.[22] Thus, if player i has a frame $\mathcal{F}_G^i$ with m different perspectives, then $M = \{S_l \subseteq N : l = 1, ..., m, \bigcup_{l=1}^m S_l = N\}$ represents the induced class of subsets of N. It is therefore natural to refer to each non-singleton cell $w^{i_J} := \left\{w^{i_j} : j \in J \subseteq M\right\} \in \mathcal{W}$ as the **meta-perspective** J of player i in G. Each cell $w^{i_J} \in \mathcal{W}$ must be thought of as the viewpoint of player i when he is simultaneously treating,(all) the decision problem(s) of player(s) $j \in J$ independently so that they are mutually rational. Hereafter, the generic term "**perspective**" will be employed for an atomic or meta-perspective.[23] The above abstract structure captures *all* the possible combinations of viewpoints that can be taken by a player in a game. Given a particular class M, we refer to the resulting structure $\mathcal{F}_G^i$ as a M-**frame**. As a particular case, a M-frame where $M = \{i, -i\}$ formalizes the above Luce and Raiffa's suggestion.

[22] That M-frames must necessarily verify that $\bigcup_{l=1}^m S_l = N$ for each player i is obvious for otherwise, some players' decision problem(s) are not even considered by i.

[23] Note however that in a pure rationalistic world it would be more natural to confine the analysis to partitions of W so as to exclude frames wherein players hold some redundant (overlapping) perspectives.

The interpretation of a M-frame for a player i is straightforward: This formalizes the self-referential mentalizing of player i. That is, the fact that a player might imagine himself determining what constitutes a rational strategy when adopting the others' decision problems in order to form some initially nonexistent beliefs in a consistent way, with itself. So a M-frame has a purely self-referential interpretation, which captures the aforementioned Luce and Raiffa's idea of introspection of a player in a game. Finally, note the implications of such a self-referential nature of a M-frame. When a (rational) player $i \notin J$ adopts a M-frame with a perspective w^{i_J}, he continues to be a rational player. This means that player i has to be rational *in the role of* other players J. This entails that the formation of the beliefs of player i at w^{i_J}, coincides with the determination of a rational strategy profile σ^J, whose all components will be *mutually rational* for players J or equivalently, for player i when he adopts the decision problems of players J.
As usual, the "knowledge" of player i at perspective w^{i_J}—we shall give a formal definition of this term in due course—is represented by the binary accessibility relation, $R^{i_J} \in \mathcal{R}^i$ over the cells of $\mathcal{W}$ i.e. $R^{i_J} \subseteq \mathcal{W} \times \mathcal{W}$. Each R^{i_J} is assumed to be a reflexive and antisymmetric accessibility relation.[24] Note that the antisymmetric property is consistent with the notion of perspective: If player i is in perspective w^{i_J}, then he cannot R^{i_J}-access w^{i_J} from another perspective $w^{i_K} \neq w^{i_J}$. [25] Hereafter, each structure $\langle w^{i_J}, R^{i_J} \rangle$, with $w^{i_J} \in \mathcal{W}$ is referred to as the $J-$**frame** of player i. Given the importance of the notion of a self-projective frame in this paper, we record this notion in the following definition.

Definition 1 We say that $\mathcal{F}_G^i = \langle \mathcal{W}, \mathcal{R}^i \rangle$ is the M-**self-projective frame** of player i in game G if each J-frame of i for $J \in M$, $\langle w^{i_J}, R^{i_J} \rangle$, with $w^{i_J} \in \mathcal{W}$ is such that $R^{i_J} \in \mathcal{R}^i$ is a reflexive and antisymmetric binary accessibility relation over $\mathcal{W}$.

We refer to the self-projective M-frame $\mathcal{F}_G^i = \langle \mathcal{W}, \mathcal{R}^i \rangle$ with $\mathcal{W} = \{w^{i_i}, w^{i_{-i}}\}$ as the **canonical self-projective frame** of player i. Hereafter, this frame will be thought of as whole class of M-frames with two perspectives i.e. the equivalence class of all frames with $|M| = 2$ such that $S_1 \bigcup S_2 = N$.

4 The axiom of "no-supplementary data" and its logical-relational semantics

This section aims at giving a formal content to the idea that a player has no other information than the one contained in the payoff matrix of a game. i.e. the axiom of `no-supplementary data` holds. With this notion in mind will emerge the notion of *relational* truth-values attached to the several rational alternatives of a rational player in a game.

[24] A relation $R \subseteq S \times S$ on a set S is reflexive if $\forall x \in S, xRx$ and antisymmetric in that xRy and yRx imply $x = y$.

[25] This is only a consistency requirement. It does not play a role in any of our results.

4.1 The axiom of non-supplementary data

Given a game G, the decision problems solved by player i at perspective $w^{i_J} \in \mathcal{W}$ pertain to "rationalistic" statements on what constitutes a rational strategy for each player $j \in J$ in G. Hence, for each player $j \in N$, we define a set of statements, $\mathbb{A}^j$. Each element $A^j \in \mathbb{A}^j$ represents a statement like $A^j :=$ "*strategy σ^j is optimal in G*", without any reference to the strategies of players $k \neq j$.[26] We shall refer to statements A^j as the **atomic statements** of player j in G. In the sequel, we set $\mathbb{A} = \mathbb{A}^1 \bigcup \mathbb{A}^2, ..., \bigcup \mathbb{A}^n$. We shall use the usual metalinguistic abbreviation: $A \wedge A^{'}$ for $\neg(\neg A \vee \neg A^{'})$ with $A, A^{'} \in \mathbb{A}$.

`[Axiom of no-supplementary data]` *Let A^j :="strategy σ^j is rational for j in G" where σ^j is neither a strongly dominant action, nor a never best reply. We say that player i has no* `no-supplementary data` *if A^j is indeterminate in the sense of the Lukasiewicz's three value-logic i.e. if $w(A^j) = \mathtt{i} \notin \{0, 1\}$ where* `i` *denotes the absence of a truth value 0 or 1 (see Table 3 below).*

The absence of a truth-value `i` formalizes the fact that the matrix game is *maximally informative*: No other extra information added to the game would induce a sharp-truth value to statement A^j. As illustrated in the Gedankenexperiment, the axiom of no supplementary entails that the truth-value at one perspective of a player will depend on the truth-value at another perspective i.e. the truth-values have only a *relational* meaning.

4.2 Kripke relational valuations

A self-projective frame is a Kripke modal structure [29, 30] formalizing the self-referential reasoning carried out by a player in a game. More specifically, in such structures, player i *simultaneously* and *independently* handles the decision problems of players $j \neq i$ by considering the viewpoints of those players with formulae A^{-i} like "*strategies σ^j, with $j \neq i$ are mutually rational between players $j \neq i$*", without any reference to strategy $\sigma^i \in \Delta^i$. Under the axiom of `no-supplementary data`, we need to extend the usual notion of an **absolute valuation** w by incorporating the possibility that its truth-values are under-determinate i.e. $w : \mathbb{A}^j \to \{0, \mathtt{i}, 1\}$. As a consequence, we are led to formalize the idea that the truth-value of a statement made at one perspective depends upon the the other truth-values assigned at the other perspectives. The upshot is thus that instead of having a usual "absolute" valuation mapping, that assigns a sharp truth-value to any given statement, we have some *contextual relational valuations*. The term "relational" captures the fact the sharp-truth values at one perspective will depend on the truth-value assigned by the player at his other perspectives. Contextuality refers to the fact that the valuation mappings depend on the particular M-frame used by the player i.e. on the number and the particular class of perspectives $\mathcal{W}$ adopted by a player.

[26] We use the term "optimal" and "rational" interchangeably.

Table 3 [Lukasiewicz' three-valued semantics]

$\vee$	1	i	0	$\wedge$	1	i	0	$\to_{\mathrm{L}}$	1	i	0
1	1	1	1	1	1	i	0	1	1	i	0
i	1	i	i	i	i	i	0	i	1	1	i
0	1	i	0	0	0	0	0	0	1	1	1

Thereafter, given a M-self-projective frame, we denote the set of formulae $\mathbb{A}^L$ considered at (meta)-perspective $w^{i_L} \in \mathcal{W}$ by $\mathbb{A}(w^{i_L})$ and a statement A^L by $A(w^{i_L})$. Given a pair of (ordered) distinct perspectives $(w^{i_J}, w^{i_K}) \in \mathcal{W} \times \mathcal{W}$, we write, $\overline{A(w^{i_J}, w^{i_K})} := (A(w))_{w \in \mathcal{W} \setminus \{w^{i_J}, w^{i_K}\}} \in \times_{w \in \mathcal{W} \setminus \{w^{i_J}, w^{i_K}\}} \mathbb{A}(w)$.

Definition 2 Let $\mathcal{F}_G^i = \langle \mathcal{W}, \mathcal{R}^i \rangle$ be a self-projective frame for player i with (w^{i_J}, w^{i_K}) a pair of (ordered) distinct perspective. A $\overline{A(w^{i_J}, w^{i_K})}$−**relational valuation** is a mapping,

$$V_{w^{i_J}, w^{i_K}}(\cdot, \cdot; \overline{A(w^{i_J}, w^{i_K})}) : \mathbb{A}(w^{i_J}) \times \mathbb{A}(w^{i_K}) \to \{0, \mathtt{i}, 1\}.$$

Relational valuations have a clear-cut interpretation. For example, in the case of a canonical frame, $V_{w^{i_i}, w^{i_{-i}}}(A^{\sigma^i}, A^{\sigma^{-i}})$ gives the truth-value of statement A^{σ^i} when i considers his own decision problem at w^{i_i}, given the truth-value of statement $A^{\sigma^{-i}}$ that has been determined (or not) at $w^{i_{-i}}$, when he puts himself into "the shoes of the others".

Definition 3 Let $\mathbb{V}_{\mathfrak{r}}$ denote the class of all relational valuations. Given a self-projective frame $\mathcal{F}_G^i$ for player i in G, we call a **self-projective model** for player i the structure, $\mathcal{M}_G^i = \langle \mathcal{F}_G^i, \mathbb{V}_{\mathfrak{r}} \rangle$. In the special case where $\mathcal{F}_G^i$ is the canonical frame, we say that $\mathcal{M}_G^i$ is the **canonical self-projective model** of player i in G.

Thereafter we say that a statement $A(w^{i_J}) \in \mathbb{A}(w^{i_J})$ is **determined as being** A^{-J}**-relatively true** in $\mathcal{M}_G^i$ at w^{i_J} if $V_{w^{i_J}, w^{i_L}}(A^J, A^L; \overline{A(w^{i_J}, w^{i_L})}) = 1$ holds for every $(w^{i_J}, w^{i_L}) \in \mathcal{W} \times \mathcal{W}$ with $J \neq L$.

4.3 Self-knowledge operators

The above self-projective Kripke model structure clearly calls for the definition of a self-modal logic. Consider a M-self-projective frame $\mathcal{F}_G^i = \langle \mathcal{W}, \mathcal{R}^i \rangle$, or the canonical self-referential frame of player i in a game G. Given the accessibility relation R^{i_J}, the possibility correspondence of player i in perspective w^{i_J} is defined as

$$\mathcal{P}^{i_J}(w^{i_J}) = \left\{ w' \in \mathcal{W} : w^{i_J} R^{i_J} w' \right\}.$$[27]

[27] Conversely, given a possibility correspondence $\mathcal{P}^{i_J}$, the associated accessibility relation R^{i_J} is obtained as follows: $\forall w' \in \mathcal{W}, w^{i_J} R^{i_J} w'$ iff $w' \in \mathcal{P}^{i_J}(w^{i_J})$.

Given a frame $\mathcal{F}_G^i$, we add a function, $f : \mathbb{A} \to 2^{\mathcal{W}}$ that associates with every atomic statement $A \in \mathbb{A}$, the set of perspectives where A is true. For every formula $F \in \mathbb{F}$, the **truth set** of F in $\mathcal{M}_G^i$, denoted by $\|F\|^{\mathcal{M}_G^i}$ is defined recursively as follows:
(1) If $F = (A)$ where A is an atomic statement, then $\|A\|^{\mathcal{M}_G^i} = f(A)$;
(2) If $\neg \|A\|^{\mathcal{M}_G^i} = \|\neg A\|^{\mathcal{M}_G^i}$;
(3) If $\left\|A \vee A'\right\|^{\mathcal{M}_G^i} = \|A\|^{\mathcal{M}_G^i} \bigcup \left\|A'\right\|^{\mathcal{M}_G^i}$;
(4) $\left\|\Box^{i_J} A\right\|^{\mathcal{M}_G^i} = \left\{w \in \mathcal{W} : \mathcal{P}^{i_J}(w) \subseteq \|A\|^{\mathcal{M}_G^i}\right\}$.
The intended interpretation of $\Box^{i_J} A$ is "*player i knows A at perspective w^{i_J}.*" If $w \in \|F\|^{\mathcal{M}_G^i}$ we say that F is true in $\mathcal{M}_G^i$ at perspective w. Thus, according to (4), player i in perspective w^{i_J} knows F if and only if F is true at every other perspective(s) that i considers at w^{i_J}. If E is the truth set of some formula F (that is $E = \|F\|^{\mathcal{M}_G^i}$), and $K : 2^{\mathcal{W}} \to 2^{\mathcal{W}}$ is the knowledge operator, then $K_{i_J} E$ is the truth set of the formula $\Box^{i_J} E$, that is $K_{i_J} E := \left\|\Box^{i_J} E\right\|^{\mathcal{M}_G^i}$. Henceforth, we say that player i in perspective w^{i_J} **knows** event $E \subseteq \mathcal{W}$ (or more precisely, the statement represented by event E) if $w^{i_J} \in K_{i_J} E$. For every event E,

$$K_{\otimes_{J\in M} i_J} E := \bigcap_{J\in M} K_{i_J} E,$$

is the event that player i has **self-knowledge** of event E in his M-frame, $\mathcal{M}_G^i$. Thus an event E is (self)-known by player $i = \otimes_{J\in M} i_J$ if player i knows E in each of his perspectives, $w^{i_J} \in \mathcal{W}$. We are now in a position to state the equivalence result illustrated in the Gedankenexperiment: Given a game G satisfying the axiom of `no-supplementary data`, player i assigns a (relational) sharp-truth value 1 to a rational strategy σ^i if and only if he assigns a (relational) sharp-truth value 1 to a Nash equilibrium $\sigma = (\sigma^1, ..., \sigma^i, ..., \sigma^n)$ of G. Formally:
Lemma 0 *Consider a game G where the axiom of `no-supplementary data` holds. Then, player i can determine σ^i as being a rational choice in G if and only if i determines profile σ^{-i} as being rational for players $j \neq i$ with (σ^i, σ^{-i}) a Nash equilibrium of G in a frame, $\mathcal{F}_G^i$, with at least two perspectives i.e. $|M| \geq 2$.*
Proof. See the Appendix.
In the sequel we will use this result as the starting point of our derivation of the quantum nature of mixed Nash equilibrium.

5 Algebraic representation of a self-projective model

Our aim in this section is to provide the preliminary definitions in order to formally establish the isomorphism between the knowledge structure of a rational player as characterized in Theorem 1, and the primitives of the QM formalism.

5.1 Faithful representation of the (self)-knowledge structure of a player

We shall prove the main Theorem of this paper by characterizing the *graph representation* of the Kripke model of $\mathcal{M}_G^i$ of a player i having determined a rational strategy i.e. a player who has "computed" a Nash equilibrium in his mind (see Theorem 1). More specifically, we shall observe that the frame $\mathcal{F}_G^i = \left\langle \mathcal{W}, \mathcal{R}^{Oi} \right\rangle$ induced in an equilibrium intrinsic state of i can be identified by a finite **weighted directed graph** Γ_G^i. Formally, a weighted directed self-projective graph for $\mathcal{M}_G^i$ is a tuple $\Gamma_G^i = \langle \mathcal{W}, \mathcal{E}, \Omega \rangle$ where $\mathcal{W}$ corresponds now to the (nonempty) set of *vertices* of the graph. $\mathcal{E}$ consists of all ordered pairs of elements $\mathcal{E} = \mathcal{W} \times \mathcal{W}$. For each ordered pair $(w, w^{'}) \in \mathcal{E}$, $e_{w,w'} = (w, w^{'}) \in \mathcal{E}$, is called an **ordered edge** with source w and target $w^{'}$. To *every* such edge we assign a weight, $\omega_{(w,w')}$ and Ω denotes the set of all weights of Γ_G^i. When dealing with perspectives $\mathcal{W}$ of i, $\omega_{e_{JK}}(s^i)$ denotes the value of weight $\omega_{e_{JK}}$ for edge $e_{JK} := (w^{i_J}, w^{i_K}) \in \mathcal{E}$ for pure strategy s^i. $\mathfrak{g} = \{\omega_{e_{JK}} : J \neq K, e_{JK} \in \mathcal{E}\} \subset \Omega$ represents the family of weights $\omega_{e_{JK}}$ in Γ_G^i where the source w^{i_J} is *different* from the target w^{i_K}. The graph Γ_G^i represents the knowledge (or mental) structure of a player deliberating about his own strategies and beliefs. So, it should be clear that the directed weight vectors of this graph will provide a purely algebraic representation of this epistemic (self)-interactive structure. This is the gist of the following definition.

Definition 4 Fix a self-projective M-Krikpean model for player i in G, $\mathcal{M}_G^i$, where σ^i has been determined as being true at a $w^{i_L} \in \mathcal{W}$. Let $(R^i, \wedge)$ be a set closed under a binary operation $\wedge$ generated by $R^i := \cup_{L \in M} R^{i_L}$ in $\mathcal{M}_G^i$, and $(\mathfrak{g}, \wedge)$ be a set closed under a binary operation $\wedge$ generated by a family of weights $\mathfrak{g}$ over a field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We say that a family of weights $\mathfrak{g} \subset \Omega$ of Γ_G^i is a **faithful representation** of the knowledge structure $(R^i, \wedge)$ of player i in Γ_G^i if
1. $\forall w^{i_L} \in \mathcal{W}$, $\omega_{(w^{i_L}, w^{i_L})} = \sigma^i \in \Omega \setminus \mathfrak{g}$;
2. there is an isomorphism, $h : (R^i, \wedge) \to (\mathfrak{g}, \wedge)$ with a pair $(w^{i_J}, w^{i_K}) \in \mathcal{E}, J \neq K,$ such that

$$h((w^{i_J} R^i w^{i_K}) \wedge (w^{i_K} R^i w^{i_J}))(s^i) = \omega_{(w^{i_J}, w^{i_J})}(s^i), \forall s^i.$$

Briefly stated, a faithful representation of the knowledge structure of player i aims at constructing a one-to-one mapping h of the structure of knowledge $(R^i, \wedge)$ of i—as determined in Theorem 1—onto the algebraic structure $(\mathfrak{g}, \wedge)$ of the induced weighted directed graph, Γ_G^i preserving all the algebraic relations (the inverse function h^{-1} behaves likewise) of $(\mathfrak{g}, \wedge)$. Alternatively put, in a faithful representation, the structure of $(\mathfrak{g}, \wedge)$ encodes the structure of knowledge of a player determining a rational strategy from a purely algebraic viewpoint. For obvious reasons, we shall refer to elements, $\left|\omega_{e_{-i,i}}\right\rangle \wedge \left\langle\omega_{e_{-i,i}}\right|$ as the **global equilibrium intrinsic state** of player i and $\omega_{e_{-i,i}}$ as his **local equilibrium intrinsic state**.
As shown in Theorem 1, a faithful representation of the knowledge of player i is necessary if one wants to have a genuine empirical model.

5.2 Observable epistemic states and empirical self-projective model

The Born rule establishes a link between the state of a system prior a measurement, and the state of the system *after* the measurement. So we also need to have an epistemic characterization of the state of mind of a player when he is making a definite choice *during* an experiment. The objective of this section is therefore to extend our definition of a self-projective model in order to analyze the observable implications of the players without any supplementary data in an experiment, via our derivation of the Born rule. To do so, we need to append to the set of mental perspectives, a set $\mathcal{S}_i^O$ of **observable epistemic states** or **observable mental states**. More specifically, $\mathcal{S}_i^O$ will describe the *real* state of mind of player i during his choice of a rational pure strategy, s^i, in the support of his determined rational mixed strategy σ^i, in an experiment. Thereafter, each element $w^{s^i} \in \mathcal{S}_i^O$ is supposed to be in one-to-one correspondence with a strategy $s^i \in S^i$ and it represents the epistemic state of player i fixing his mind on the single action $s^i \in S^i$.

Definition 5 Let $\mathcal{M}_G^i$ be a self-projective model for i in G with a frame $\mathcal{F}_G^i$. The **empirical self-projective model** for i in G is a self-projective model denoted, $\mathcal{M}_G^{iO}$, whose frame has been extended i.e. $\mathcal{F}_G^{iO} = \left\langle \mathcal{S}^i, \mathcal{R}^{iO} \right\rangle$ where $\mathcal{S}^i = \mathcal{W} \bigcup \mathcal{S}_i^O$ such that $R^{i0_J} \in \mathcal{R}^{iO}$ is R^{i_J} extended to $\mathcal{S}^i \times \mathcal{S}^i$.

The following axiom gives a formal content to the fact that a player in a game "fixes his mind" on a particular pure action (in the support of his mixed strategy). It says that, given his M-frame, the choice $s^i \in S^i$ of a player i can be observed at $w^{s^i} \in \mathcal{S}_i^O$ if and only if each perspective of player i knows this choice i.e. in an experiment, the player mentally fixes his mind on exactly the *same* pure action at every of his perspectives.

[**Indivisibility**] Let A^{s^i} be the atomic statement in $\mathbb{A}^i$ "*pure strategy s^i is optimal in G*". The statement A^{s^i} is **observed** at $w^{s^i} \in \mathcal{S}_i^O$ if and only if $w^{s^i} \in K_{\otimes_{J \in M} i_J} E^{s^i}$.

In other words, player i makes a definite choice s^i if and only if he has self-knowledge that "*s^i is a rational action*" in each of his perspectives. Note that this axiom will automatically imply the "wave-particle duality" principle of QM; when player i is asked to make a definite choice, he has to be *perfectly correlated* on the choice of a pure action across all his different perspectives, while the indifference condition requires that he knows that all the pure actions in the support of his mixed strategy are simultaneously rational. The perfect correlation of the player with himself is therefore equivalent to the "wave packet collapse" postulate of QM.

The following definition relates the epistemic or mental state of a player with the probability that a pure action occurs in an experiment.

Definition 6 Let $\mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ be the lattice induced by $\mathcal{S}_i^O$. Fix an empirical self-projective Krikpean model for player i in G, $\mathcal{M}_G^{iO}$, where σ^i has been determined as being true. A probability measure

$$e^i : \mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right) \to [0, 1],$$

satisfying $e^i(w^{s^i}) = \sigma^i(s^i)$, $\forall s^i \in S^i$, is called the **empirical distribution** for the set of strategies S^i of player i.

The intuition behind the definition of an empirical distribution is simple. It says that if a player takes a particular pure action (in the support of his mixed strategy), this implies that his state of mind is fixed on this particular action during the measurement. In short, an empirical distribution translates the observable mental states of a player to its probability of occurrence in an experiment.
In QM, it is standard to study the logic of verifiable propositions of a physical system through the algebra of the so-called "quantum logic" introduced by Birkhoff and von Neumann [31]. In fact, it is well-known that an essential feature of the structure of such elementary events called "*yes no experiments*" is to generate a non-Boolean lattice (also called non-Boolean algebra) i.e. an orthocomplemented lattice which is non-distributive. Lemma 1 below shows that the same phenomenon occurs in the classical game model: The lattice of observable mental states of a player, $\mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$, exhibits the same non-classical properties as the "*yes no experiments*" of the logic of verifiable propositions of a physical system.

Lemma 1 *Fix an empirical self-projective Krikpean model for player i in G, $\mathcal{M}_G^{iO}$, where (σ^i, σ^{-i}) has been determined as being true. Then, $\mathcal{S}_i^O = \left\{K_i(E^{s^i} \wedge E^{\sigma^{-i}}) : s^i \in \mathtt{supp}(\sigma^i)\right\}$ and $\mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ is an orthocomplemented non-distributive lattice.*

Proof. See the Appendix.

5.3 Equivalence of the Axiom of Indivisibility and the projection postulate of QM

As already hinted, the "wave packet collapse" or "projection postulate" of QM arises automatically in the classical game model. Here are the formal details of the proof. Given the unit vector $a_{s^i} \in \Delta^i$, $\Pi_{a_{s^i}} := |a_{s^i}\rangle\langle a_{s^i}|$ denotes the corresponding orthogonal projector expressed in terms of the Dirac notation. As in QM, (see e.g. Birkhoff and von Neumann [31]), game-theoretic statements like "*upon measurement player i chooses a strategy in subset $\widehat{S}^i \subseteq S^i$*", can be represented by particular linear subspace of a Hilbert space $\mathbb{H}$ (or projection operators Π).[28]

Lemma 2 *Let $\mathcal{P}(\mathbb{H})$ be the lattice of all orthogonal projections (or closed subspaces) on $\mathbb{H}$ generated by the set $\left\{\Pi_{a^{s^i}}\right\}_{s^i \in S^i}$. There exists a probability measure $\mu : \mathcal{P}(\mathbb{H}) \to [0,1]$ with $\mu(\Pi_{a^{s^i}}) = f(\mathtt{Im}(\Pi_{a^{s^i}}))$ such that $e^i(w^{s^i}) = \mu(\Pi_{a^{s^i}})$, $\forall w^{s^i} \in \mathcal{S}_i^O$.*

[28] In QM, Birkhoff and von Neumann [31], show that if one introduces the concept of "operational proposition" and its representation in standard QM by an orthogonal projection operator of the Hilbert space, then the set of "experimental propositions" does not form a Boolean algebra, as it the case for the set of propositions of classical logic.

Proof. See the Appendix.

Hence, the projection postulate of QM follows immediately from the Indivisibility of a player (recall that in QM a quantum system is also an indivisible entity in the physical sense), which implies that there is a perfect "coordination" on the pure strategy $s^i \in \mathtt{supp}(\sigma^i)$ amongst all the perspectives of player i. Intuitively, player $i = \otimes_J^m i_J$ makes a choice $s^i \in S^i$ if and only if every perspective, w^{i_K},$K = 1, ..., m$, of player i has the same projection, $\Pi_{a_{s^i}}$. The upshot is thus that this perfect correlation implies that the probability that the unit vector $a_{s^i} \in \Delta^i$ is actually chosen by player i corresponds to the probability of the orthogonal projection $\Pi_{a_{s^i}}$ of any vector (or $\omega^\dagger_{e_{KJ}}$ when $\omega_{e_{KJ}}$ is a co-vector) $\omega_{e_{KJ}} \in \mathfrak{g}, e_{KJ} \in \mathcal{E}$ onto the one-dimensional subspace spanned by a_{s^i} i.e. $\Pi_{a_{s^i}} \omega_{e_{KJ}}, K \neq J$.

6 Main result: Isomorphism between a pure quantum state and a self-projective model

We are now in a position to build the isomorphism between the structure of the binary relations of a self-projective graph and its family of weights and the QM state-space. Our main result below provides a link between the self-knowledge structure of a player deliberating about a rational choice and the algebraic structure of quantum mechanics.

First some notations and definitions. For our purposes, it will be sufficient to focus on finite-dimensional vector spaces $\mathbb{C}^n$ from linear algebra. These spaces are Hilbert spaces in the usual inner product $(w, v) = \sum_{k=1}^n w_k v_k$. A peculiarity of the linear algebra used in QM is the so-called Dirac notation for vectors [32]. In this notation, the symbols $|\, v\rangle$ and $\langle v|$ denote the vectors *in* and the *linear forms* on Hilbert space, respectively. In QM, $\langle v|$ is called a "bra" vector and $|\, v\rangle$ a "ket". Hence, if $v := |\, v\rangle$ is a vector in Hilbert space $\mathbb{H}$ over a real or complex field $\mathbb{K}$, then $|\, v\rangle$ is just another notation for v, and $\langle v|$ means the mapping $u \mapsto \langle v\,|u\rangle$, a linear form $\mathbb{H} \to \mathbb{K}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) defined using the Euclidean or Hermitian inner product $\langle \cdot\,|\,\cdot\rangle$ on $\mathbb{H}$. Given a vector v, it is also convenient to consider $\langle v|$ as a *row vector* $v^\dagger$ in the dual space of $\mathbb{H}$, denoted $\mathbb{H}^*$ and the superscript $\dagger$ means the usual conjugate-linear operation: $\langle v|^\dagger = |\, v\rangle$, and $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$. Finally, a **ray** of $\mathbb{H}$ is a linear subspace consisting of all kets of the form $\alpha\,|\omega\rangle$, where α is any complex number.

In the next result we exclude "exceptional" games having equilibrium strategy profiles $(\sigma^i)_{i\in N}$ wherein $\mathtt{supp}(\sigma^i)$ is contained in the set of pure equilibrium strategies of player i. We comment on this pathological situation in the next section (see Corollary 1). We also exclude games where Nash equilibria exist in strongly dominant strategies for some players. Hereafter we say that $\mathcal{M}_G^i$ is the **canonical self-projective model** if it is a model with the set of perspectives $\mathcal{W} = \left\{w^{i_i}, w^{i_{-i}}\right\}$ (modulo the equivalence relation on M-frames with $|M| = 2$).

Theorem 1 *Fix a game $G = \langle S^1, ..., S^n; \pi^1, ..., \pi^n \rangle$ and assume the axiom of* `no-supplementary data` *holds with $(\sigma^i)_{i\in N}$ a Nash equilibrium of G. Fix a self-projective model for player i in G, $\mathcal{M}_G^{iO}$, where σ^i has been determined as being a rational choice. If player i has more than two pure strategies i.e. $m_i \geq 3$, the unique self-projective model of player i which admits an empirical distribution is the canonical model (modulo the equivalence relation on M-self-projective models with $|M| = 2$) where $e^i = \omega_{(w',w')} = \sigma^i \in \Omega \setminus \mathfrak{g}$, for $w^{'} = w_{e_{i,i}}, w_{e_{-i,-i}}$, is derived from a faithful representation of player i's knowledge in Γ_G^i such that:*
(1) the set $(\mathfrak{g}, \wedge)$ is generated by the set

$$\mathfrak{g} = \left\{ \left|\omega_{e_{-i,i}}\right\rangle, \omega_{e_{i,-i}} := \left\langle \omega_{e_{-i,i}} \right| \right\}$$

over the complex field $\mathbb{K} = \mathbb{C}$, where $\left|\omega_{e_{-i,i}}\right\rangle$ is called a "wave function";
(2) the set of all possible equilibrium local intrinsic states of player i like, $\omega_{(w^{i_{-i}}, w^{i_i})}$, (of a faithful representation $\mathfrak{g} \subset \Omega$ of Γ_G^i) forms a unit ray of $\mathbb{C}^{m_i}$;
(3) the equilibrium global intrinsic state of player i $\left|\omega_{e_{-i,i}}\right\rangle \wedge \left\langle \omega_{e_{-i,i}} \right| = \left|\omega_{e_{-i,i}}\right\rangle \left\langle \omega_{e_{-i,i}} \right|$ is a projection operator;
(4) the empirical distribution is given by the Born rule i.e. $e^i(s^i) = \left|\omega_{(w,w^\dagger)}(s^i)\right|^2, \forall s^i$.

Proof. See the Appendix

So, Theorem 1 provides a derivation of the complex Hilbert state-space structure of QM from the knowledge or mental structure of a rational player in the classical game model. A sizable literature has sought to apply the quantum formalism to explain various human decision processes. Theorem 1 provides the converse direction. The main message of this result is thus that the basic algebraic structure of QM describes the mental state-space structure of a rational player without any supplementary data. This theorem also relates the vast literature on quantum and modal logic with classical game theory in a clean way. This amounts to the construction of an isomorphism between *the only possible* graph structure of a rational player and the QM formalism. In fact, as proved in the Appendix, the mere existence of an empirical distribution requires a faithful representation of the self-projective graph: *There exists an empirical distribution for a self-projective model of a player if and only if the representation of its graph is isomorphic to the QM formalism, which is only possible for the canonical model with two perspectives.* The underlying quantum-mechanical structure of a game therefore arises from the epistemic states of a player without any supplementary data deliberating about his rational actions. Formally, this derivation follows from two observations:
(i) The determination of a rational strategy, together with the existence of an empirical self-projective model require exactly two vertices (perspectives) and;
(ii) The relational structure of such two-vertices weighted directed graphs is captured via the algebraic duality between the strategy space of i at w^{i_i} and the belief space of $w^{i_{-i}}$.
Property (ii) has already been discussed above. So, let us consider (i). The existence of a well-defined empirical model constraints the number of vertices of the self-projective graph i.e. the player's perspectives. As shown in Theorem

1, the determination of a rational behavior requires at least two perspectives. This yields a lower bound. In addition, as demonstrated in the Appendix, there cannot exist a well-defined empirical probability measure if the player uses more than two perspectives during his introspection. This gives an upper bound. From this we conclude that a player must have exactly *two* perspectives. Hence, the graph will only have two directed edges representing the binary relations R^{i_i} and $R^{i_{-i}}$.

Also, notice that the axiom of indivisibility (the perfect correlation of the player across his perspectives) ensures that the weights on the diagonal of the projector form a well-defined probability measure corresponding to the determined rational mixed strategy σ^i. That is, the operator representing the state of mind of a player must have a trace 1. Together with the axiom of independence, this implies that a faithful algebraic representation of the self-projective graph is given by a pair of dual unit weight vectors.

To recap, the sketch of the proof of Theorem 1 is as follows.

In a first step, we construct an isomorphism between the weighted directed graph Γ_G^i of the *canonical* model of a player i having determined a rational strategy. Theorem 1 implies that the accessibility relations $R^{i_i}, R^{i_{-i}}$, (the knowledge structure of player i) of this model verify $w^{i_i} R^{i_i} w^{i_{-i}}$ and $w^{i_{-i}} R^{i_{-i}} w^{i_i}$. Hence, the graph representation of the canonical model, Γ_G^i, must have two directed opposite edges, $e_{i,-i}$ and $e_{-i,i}$ if one wants to establish a one-to-one relationship between the two structures. As proved in the Appendix, the knowledge structure between $(R^i, \wedge)$ and $(\mathfrak{g}, \wedge)$ is preserved if and only if the properties (2)-(4) of Theorem 1 are fulfilled.

The second step of the proof consists in showing that any non-canonical self-projective model cannot induce a proper empirical distribution in an experiment, which will single out the canonical model as the only possible one. We do so by noting that a faithful representation of such graphs would lead to a violation of Gleason's Theorem [33].

6.1 Interpretation of the "Nash equilibrium mixtures"

The construction of the isomorphism of Theorem 1 already indicates that Nash equilibrium mixtures must have an interpretation in terms of quantum superpositions. However, the notion of a "mixture" as an expression of a classical ignorance also exists in quantum mechanics. We therefore need to complete the analysis in order to pinpoint the exact nature of the probabilities arising in mixed strategies. We start with some preliminary definitions. In the previous section, we have shown that the notion of orthogonal projector allows to describe the equilibrium state of mind of a player before his choice. In general, however, as in QM, the widely used density matrices turn out to the natural tools to describe the intrinsic state(s) of a player. First some general definitions. Let $\mathbb{H}$ be a Hilbert space, $\{a_n\}_{n\in\mathbb{N}}$ an orthonormal basis. We define the trace of an operator ρ by $\mathtt{tr}(\rho) = \sum_{n\in\mathbb{N}} \langle a_n \mid \rho a_n \rangle$. Hereafter, we set $\mathtt{tr}(\rho) = 1$. If further ρ is positive semi-definite, then ρ is said to be a **den-**

sity matrix (operator). The set of density matrices on $\mathbb{H}$ is denoted by $\rho(\mathbb{H})$. The set of density matrices on the complex unit sphere $\mathbb{S}$ in Hilbert space $\mathbb{H}$ corresponds to the convex hull of $\{|\omega\rangle\langle\omega| : |\omega\rangle \in \mathbb{S}\}$.

Definition 7 Let $\{\omega\}$ be a family of unit vectors, $\{p_\omega\}$ a family of weights forming a non degenerate probability distribution over the set of orthogonal projectors, $\{\Pi_\omega\}$. A density operator, ρ_M, is a **statistical mixture** or a **mixed state** if $\rho_M = \sum_\omega p_\omega \Pi_\omega$. A **pure state** is a density matrix where $\{p_\omega\}$ forms a degenerate probability distribution i.e. $\rho_M = \Pi_\omega$.

Theorem 1 shows that the *pure* equilibrium intrinsic states of a players are the density matrices with a single term i.e. $\rho = \Pi_\omega$. By contrast, the mixed intrinsic state of a player is represented by a sum of such projectors upon pure intrinsic states ω with associated statistical weights $0 \leq p_\omega \leq 1$, with $\sum_\omega p_\omega = 1$.[29] In order to drive the wedge between these different situations—and in line with QM—we say that the equilibrium intrinsic state of a player has the **ignorance interpretation** if and only if the equilibrium intrinsic state of player i is really in one of the pure state Π_ω but the observer does not know which one. In this interpretation, the probabilities $\{p_\omega\}$ are subjective and merely reflect the degree of ignorance of the observer. In fact, it follows from Theorem 1 that the interpretation of mixed-strategy Nash equilibria as an expression of the ignorance of an outside observer (or player) is in general impossible.[30]

Corollary 1 *Fix a game $G = \left\langle S^1, ..., S^n; \pi^1, ..., \pi^n \right\rangle$ with a generic mixed Nash equilibrium strategy profile $\sigma = (\sigma^1, ..., \sigma^i, ..., \sigma^n)$. If the empirical distribution is, $e^i = \sigma^i$, then the global equilibrium intrinsic state of player i is pure as it forms a coherent superposition of the pure actions $s^i \in$ `supp`(σ^i).*

Proof. See the Appendix.
This result states that if one takes the formal definition of the classical game model, then the resulting axiom of `no-supplementary data` entails that we can no longer interpret mixed strategies as the reflect of the ignorance of some outside observer (or the other players). There is a simple reason for that: Such an "epistemic interpretation" of probabilities would only be valid if the player had fixed his mind on a particular pure action of the support before a measurement. However, Corollary 1 shows that prior to making a definite choice in an experiment, the mental state of a player is described by a wave function, whose interference terms imply that the player has not yet fixed his mind on a particular pure action. Again, this is just the modal characterization of the usual "indifference condition". The upshot is thus that probabilities cannot be

[29] It is tempting to interpret the weights p_ω as probabilities, so that p_ω is the probability that player i is in the equilibrium pure local intrinsic state ω. However, we should be cautious here. It is indeed well-known that, in general, a density matrix admits uncountably many decompositions into pure states.

[30] Note that it is still possible to interpret σ^i as an expression of the ignorance of an outside observer (or player) in the particular case where the *whole* support of a mixed strategy of player i, σ^i, is included in the set of pure Nash strategy equilibria of player i.

construed as a lack of knowledge that an outside observer (or another player) would hold about which pure action of the support will be played in an experiment. As in QM, the probabilities arising in a Nash equilibrium must therefore be seen as an expression of an "irreducible" randomness and probabilities arising in equilibrium must be construed as "tendencies" or "dispositions" of the players' states of mind. Indeed, as in QM, these probabilities have an "ontological status"—a player has not singled out a particular pure action in his mind prior he is asked to do so—rather than reflecting the ignorance of an outside observer—a player has already fixed his mind on a particular action, but the other players cannot read in his mind.[31] Also, note that these probabilities reflect the fact that a pure choice is a real "future contingent" statement (an unsettled event), as long as a player has not settled his mind on a particular pure action: prior an experiment, the definite pure choice of the player is not an actual state of affair, in the same way as the mixed strategy was initially nonexistent in the mind of the player.
What about the interpretation of a mixed strategy as an explicit randomization? This interpretation can also be ruled out, since in this view, the *actual* choice is made by a "random device", which is not part of the description of the game model. [32] As an aside, note that the addition of such an artificial device would not account for the behavior of the player, but of the device. Alternatively put, we can also say that the description of the mental state of a player as given by the QM formalism is maximally informative, in the sense that the state of knowledge of the player in on a par with the state of knowledge of an ideal observer. We complete this section by stating a result which relates the general case of density matrices with the issue of multiple Nash equilibria in games.

Corollary 2 *Fix a game* $G = \left\langle S^1, ..., S^n; \pi^1, ..., \pi^n \right\rangle$ *with a generic set of mixed Nash equilibria strategy profiles,* $\{\sigma_\omega\}$ *. Suppose player i has determined one of these Nash equilibria in the canonical self-projective model. The empirical distribution is,* $e^i = \sum_\omega p_\omega \sigma^i_\omega$*, if and only if the global equilibrium intrinsic state of player i,* $\rho_M = \sum_\omega p_\omega \Pi_\omega$*, has the ignorance interpretation.*

Proof. See the Appendix
This result states that we can partially recover a usual interpretation of probabilities in games by embedding the existence of multiple Nash equilibria into the analysis. By Theorem 1, this amounts to a player determining several rational strategies i.e. Nash equilibria—with a certain probability distribution $\{p_\omega\}$. Hence, in this case, the outside observer is uncertain in the usual sense about which pure quantum state $|w\rangle$ i.e. Nash equilibrium, is really occurring in the mind of the player. The intuition for the proof of this result is then straightforward: Unlike in the case of a mixed Nash equilibrium, a player is no

[31] This "propensity interpretation" of probabilities is defended by the so-called ontic approaches.
[32] Note also that the "randomization interpretation", would bring us back to the "epistemic" interpretation, since in this case, the device would just be a surrogate of the player choosing a pure action *prior* an experiment.

longer indifferent between the several multiple Nash equilibria, which implies that a player's state of mind cannot be described as a quantum superposition over the various equilibria of the game.

6.2 Mixed strategies as ontic-epistemic states of mind

Corollary 1 indicates that the nature of the probabilities appearing in games take a much more delicate interpretation for they appear to take an ontic *and* epistemic interpretation. To see this point more clearly, there is an instructive visual story inspired by the previous game-theoretic Gedankenexperiment.
Think of the perspectives of Ann, A_A and $A_{B\otimes C}$, as being two distinct persons. Imagine that the atomic perspective of Ann A_A has one head and the meta-perspective $A_{B\otimes C}$ is endowed of two heads (this is a meta-perspective made up of two persons). Suppose that the two perspectives of Ann are sitting side by side. We assume that each perspective of Ann is blindfold. There is an urn containing three hats of different colors.[33] Each perspective of Ann wears the head(s) of the other by drawing a hat out of the urn. Perspective A_A wears the heads of $A_{B\otimes C}$ by drawing two hats out of the urn, while $A_{B\otimes C}$ draws (simultaneously) one hat to wear the head of A_A (still blindfold). In this story, all statements like "*The perspective of Ann A_A wears a white hat*" are indeterminate *before* the drawing of the hats. This is the ontological part of the story. The epistemic part of the story only begins once the hats have been drawn out of the urn. Since each perspective of Ann draws a hat blindfold, she cannot (trivially) infer the color of her hat(s) by watching the head of the other. Similarly, in the process of determination of an equilibrium, Ann herself does not know what will be her definite choice of a pure strategy in an experiment.[34] Thus, an outside observer reading in the mind of Ann, must be in the same epistemic state on her choice, than Ann herself. Thus, we cannot interpret the probabilities as measures of ignorance of an outside observer of the actual unknown values of the player's pure strategies.
The bottom line of this story is thus that in a mixed Nash equilibrium, the state of mind of a player is a state of knowledge. This state is thus epistemic in nature since it is only expressible in terms of the player's knowledge. However, this epistemic state is also ontic i.e. it is a state of reality, because it provides the actual complete specification of all the properties of the player. Hence, the critical difference between the usual definition of an epistemic state is that an equilibrium mixture does *not* represent the relative likelihood that an outside observer (or an another player) would assign to the choice of a player's pure strategy. Instead, the epistemic state of the observer coincides exactly with the actual intrinsic state of the player.

[33] The story is inspired from the well-known "three-hat puzzle".
[34] Formally, recall that this comes from the existence of off-diagonal terms.

7 Concluding remarks

The difficulty of interpretation of mixed strategies is a well-known issue in classical game theory (see e.g. Osborne and Rubinstein, [26]). To the best of our knowledge, Pietarinen [20, 21] is the first to have studied the similarities between the logic of games of imperfect information and quantum theory. In this paper we have proven that the existence of such a connection is fundamentally related to the notion of mixed Nash equilibrium.
Our derivation of the state-space of quantum mechanics proceeds from the following reasoning: Without any supplementary data, the best response structure of a game does not generally contain enough information to make a rational choice. This leads each player to adopt the decision problems of the other players, in order to determine his *own* rational choice. It then becomes a mere internal consistency requirement that the *same* player will have to determine his own strategy by determining an entire profile of beliefs (the choices and therefore beliefs for the others) that are mutually rational to each others, thereby forming a Nash equilibrium. Thus, the above introspective reasoning induces the famous "indifference condition" that must hold in any mixed Nash equilibrium: A player deems equally rational *all* the alternatives contained in the support of his own mixed strategy if and only if the (common) belief he holds when determining the others' player strategies matches his own mixed strategy. This is precisely the very existence of such a *self-referential belief* that is at the origin of the description of a player in terms of a quantum state of mind. By construction, a player can hold such self-referential beliefs *if and only if* he has really not made-up his mind on a pure choice. It turns out that the algebraic characterization of this indifference property is indeed given by a pure quantum state.

7.1 Historical note

In his dissertation, Nash [3] motivates his equilibrium concept as a methodology to rationally predict the behavior of players:

> By using the principles that a rational prediction should be unique, that the players should be able to deduce and make use of it, and that such knowledge on the part of each player of what to expect the others to do should not lead him to act out of conformity with the prediction, one is led to the concept of a solution defined before.
>
> [Nash, [3]].

It is well-known that von Neumann was not convinced by this methodology (see Shubik [34, p.155]). Nash's proposal seemed to be at odds with the "*enormous variety of observed stable social structures*" [Wolfe, [35], p.25] that von Neumann viewed as part of an "irreducible indeterminism":

> The discussion opened with a statement by von Neumann in justification of the enormous variety of solutions which may obtain for n-person games. He pointed out that this was not surprising in view of the correspondingly enormous variety of observed stable social structures; many differing conventions can endure, existing today for no better reason than that they were here yesterday.

[Wolfe, [35], p.25]

On the other hand, it seems also clear that von Neumann had perceived an implicit connection between this inherent indeterministic nature of human behavior and the seemingly "irreducible probabilistic" nature of quantum physics:

Although...chance was eliminated from the games of strategy under consideration (by introducing expected values and eliminating "draws"), it has now made a spontaneous reappearance. Even if the rules of the game do not contain any elements of "hazard"(i.e. no draws from urns) . . . in specifying the rules of behavior for the players it becomes imperative to reconsider the element of "hazard". The dependence on chance (the "statistical" element) is such an intrinsic part of the game itself (*if not of the world*) that there is no need to introduce it artificially by way of the rules of the game: Even if the formal rules contain no trace of it, it still will assert itself.

(Von Neumann [36] p. 26, emphasis added)

As summarized by Leonard [37]:

(...) the prevailing probabilistic view of the world in physics was being reflected in von Neumann's theory of human interaction.

[37, p.734]

In the light of the results of this paper, the quantum-mechanical nature of mixed Nash equilibria can thus be viewed—somewhat ironically—, as accounting for the von Neumann's "philosophy of indeterminism" [the term is taken from Brandenburger (2010)] in game theory. However, there does not seem to be any evidence that von Neumann noticed a formal connection between these two kinds of "indeterminism".

7.2 Relation to the notions of indeterminism and free choice in physics

We note that the present results hint that quantum mechanics is not compatible with a deterministic world, as suggested by the branching space-time framework of Belnap (see Belnap, [38]) and its connection to the existence of indeterminate quantum phenomena (see e.g. Belnap et al. [39,40]). However, the possibility of such a connection remains open.

We also remark that the present results echo the recent literature in physics analyzing the consequence of the free choice postulate on the nature of the quantum state (see e.g. Colbeck and Renner [7,8]).[35] In this paper, the free choice assumption can be related to the under-determination of the game matrix and captured by our axiom of "no-supplementary data".

Finally, this paper is also fundamentally related to the "free will theorem" of Conway and Kochen [14,15] who argued that "*if indeed we humans have free will, then elementary particles already have their own small share of this valuable commodity.*" (Conway and Kochen, [15], p. 226). This suggests that the decisions of humans cannot be made a function of the past, in the same manner as a "*particle"s response is not determined by the entire previous history of the universe*" [Conway and Kochen p.1, [14]].

[35] Colbeck and Renner considers a choice "A free if it is uncorrelated with any other variables, except those that lie in the future of A in the chronological structure".

APPENDIX

Proof of <u>Lemma 0</u>. We shall prove the following statement: *Consider a game G where the axiom of* `no-supplementary data` *holds. Then, player i can determine σ^i as being a rational choice in G if and only if i determines profile σ^{-i} as being rational for players $j \neq i$ with (σ^i, σ^{-i}) a Nash equilibrium of G in a frame, $\mathcal{F}_G^i$, with at least two perspectives i.e. $|M| \geq 2$.*
Proof. Let $(A^J)_{J\in M}$ represent the Nash equilibrium profile $\sigma = (\sigma^{i_L})_{L\in M}$.
Sufficiency. Consider a model with a M-self-projective frame $\mathcal{F}_G^i = \left\langle w^{i_J}, R^{i_J} \right\rangle_{J\in M}$ with $|M| \geq 2$. Suppose that E^σ is common knowledge. Thus, we have the set of binary relations, $R^{i_J} \in \mathcal{R}^i$ such that $w^{i_J} R^{i_J} w^{i_K}$, for all $J \neq K$. We must also have a profile of statements, $(A^L)_{L\in M}$, which has been determined as being relatively true. That is, we have a system of relational valuation mappings,

$$A^J \in \arg \max_{A'^J \in \mathbb{A}^J} V_{w^{i_J}, w^{i_K}}(A'^J, A^K; A^{-J,-K}), \forall (w^{i_J}, w^{i_K}) \in \mathcal{W} \times \mathcal{W}, J \neq K,$$

whose solution, $(A^L)_{L\in M}$, corresponds to the Nash equilibrium profile of G, $\sigma = (\sigma^{i_L})_{L\in M}$. The solution of the above system of equations is guaranteed by the standard existence theorem of a Nash equilibrium in finite games (Nash, [3]). Hence, if E^σ is common knowledge in the canonical model, then we necessarily have a Nash equilibrium with $\mathcal{M}_G^i, w^{i_K} \vDash A^{\sigma^J}, \forall w^{i_K} \in \mathcal{W}$. When this is true for each $i \in N$, this implies that σ is a mixed Nash equilibrium (NE) of the game being played.
Necessity. We first show that A^J is true in a M-self-projective model of i only if i determines a Nash equilibrium. By definition, the self-projective model with the single perspective w^{i_N} cannot yield the determination of a truth $A^N \in \mathbb{A}^N$ when there is no strongly dominant actions. [36] Thus, we must pick a M-self-projective model with at least two perspectives. Suppose that there exists an arbitrary $L \in M$ with $R^{i_L} \in \mathcal{R}^i$ such that $w^{i_L} \neg R^{i_L} w^{i_K}$, for some $L \neq K$. Then, by definition, $\emptyset \in \arg\max_{A'^L \in \mathbb{A}^L} V_{w^{i_L}, w^{i_K}}(A'^L, A^K; A^{-L,-K})$, for $L \neq K$, since $V_{w^{i_L}, w^{i_K}}(A'^L, A^K; A^{-L,-K}) = \mathtt{i}$ whenever $\mathcal{M}_G^i, w^{i_L} \vDash \neg\Box^{i_L} A^{\sigma^K}$. Hence, we have that a statement is determined only if $w^{i_J} R^{i_J} w^{i_K}, \forall J \neq K$. Notice that E^σ is a self-evident event to every perspective w^{i_J} of player i i.e. $E^\sigma = K_{i_J} E^\sigma$. This follows since, $w^{i_J} \in \left\| \Box^{i_J} A^{\sigma^{-J}} \rightarrow_{\mathtt{L}} A^{\sigma^{-J}} \right\|^{\mathcal{M}_G^i}$. [37] Hence $w^{i_J} \in \left\| \Box^{i_J} A^{(\sigma^i)_{i\in N}} \right\|^{\mathcal{M}_G^i}$, for all $w^{i_J} \in \mathcal{W}$ whenever the profile of propositions, $(A^J)_{J=1}^m$ is the solution of the above set of equations and corresponds, by construction, to a Nash equilibrium of the game G. **End of proof**.
Proof of Lemma 1. We first characterize the elements of $\mathcal{S}_i^O$. Fix an empirical self-projective Krikpean model for player i in G, $\mathcal{M}_G^{iO}$, where σ^i has

[36] Indeed, recall that in this case $w^{i_N}(A^N) = \mathtt{i}$, $\forall A^N \in \mathbb{A}^N$ and there is no relational valuations.
[37] Recall that the binary relation R^{i_J} is reflexive at every perspective J. Thus, we have that $\forall E \subseteq \mathcal{W}$, $K_{\otimes_{i_J \in M}} E \subseteq E$. which means that the *truth axiom* is satisfied for each player i in $\mathcal{F}_G^i$.

been determined as being true. By Theorem 1, the indifference condition holds i.e. any $s^i \in \mathtt{supp}(\sigma^i)$ is necessarily a best reply to σ^{-i}. Hence, when i is making a rational choice, we must have $w^{s^i} \in \mathcal{S}_i^O$, such that $\mathcal{M}_G^{iO}, w^{s^i} \models \wedge_{L \in M} \Box^{i_L}(A^{s^i} \wedge A^{\sigma^{-i}})$. The indivisibility axiom requires that event $E^{s^i} \wedge E^{\sigma^{-i}}$ is self-knowledge, which proves that $K_i(E^{s^i} \wedge E^{\sigma^{-i}})$ is an element of $\mathcal{S}_i^O$. Next, we verify that $\mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ is an orthocomplemented non-distributive lattice. By definition of a pure choice, we must have that $w^{s^i} \wedge w^{s'^i} = \emptyset$, $\forall w^{s^i} \neq w^{s'^i}$. By definition of a lattice, $\vee$ and $\wedge$ must be idempotent, commutative, and associative operations such that $a \vee (b \wedge a) = a$ and $a \wedge (b \vee a) = a$, $\forall a \in \mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$. Idempotence and commutativity are obvious. We also note that $w^{s^i} \vee (w^{s'^i} \wedge w^{s^i}) = w^{s^i}$ since $w^{s'^i} \wedge w^{s^i} = \emptyset$ whenever $w^{s'^i} \neq w^{s^i}$ by definition of a pure choice and $w^{s^i} \wedge (w^{s^i} \vee w^{s'^i}) = w^{s^i}$. That $\wedge$ is associative follows since $w^{s^i} \wedge (w^{s'^i} \wedge w^{s''^i}) = (w^{s^i} \wedge w^{s'^i}) \wedge w^{s''^i} = \emptyset$ whenever $w^{s^i} \neq w^{s'^i}$ and $w^{s'^i} \neq w^{s''^i}$. Moreover, in equilibrium, the indifference condition implies that any $s^i \in \mathtt{supp}(\sigma^i)$ is a best reply. Thus, $\bigvee_{s^i \in \mathtt{supp}(\sigma^i)} w^{s^i} = \mathcal{S}_i^O$ so that $o := \emptyset$ and $1 := \mathcal{S}_i^O \in \mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ are the two identity elements with $a \wedge 1 = a$ and $a \vee o = a$. Moreover, the lattice is orthocomplemented under the set-theoretic complement, $a^{\sim} = \mathcal{S}_i^O \setminus a, \forall a \in \mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ such that $a \subseteq b \Leftrightarrow a^{\sim} \subseteq (b)^{\sim}$ and the partial ordering $\subseteq$ corresponds to the material implication "$\Rightarrow$". That this lattice is non distributive[38] follows by noting that $w^{s^i} \wedge (w^{s'^i} \vee w^{s''^i}) \subseteq w^{s^i} \neq \emptyset$ whenever $s^i \in \mathtt{supp}(\sigma^i)$. This implies that, $w^{s^i} \wedge (w^{s'^i} \vee w^{s''^i}) \neq \emptyset$ while $(w^{s^i} \wedge w^{s'^i}) \vee (w^{s'^i} \wedge w^{s''^i}) = \emptyset$ whenever $s^i \neq s'^i$ and $s'^i \neq s''^i$. **End of proof**.

Proof of Theorem 1. We first construct the isomorphism between the complex Hilbert space structure of QM together with its postulates and the self-projective graph induced by a canonical model i.e. a model whose M-frame has exactly *two* perspectives. Then, in a second step, we will use this result as a lemma to show that *any* other model with more than two perspectives (i.e. the M-frame is such that $|M| > 2$) does not induce a well-defined empirical distribution e^i in an experiment. From this we will conclude that the only possible self-projective model is the canonical model. The isomorphism between the canonical model and the QM structure then terminates to demonstrate that the mental state space of a rational player in the classical game model is always isomorphic to the state-space structure of QM.

The derivation of the QM formalism of Theorem 1 builds on the following result.

Proposition 1 *Fix the canonical self-projective Krikpean model for player i in G, $\mathcal{M}_G^{iO}$, where σ^i has been determined as being true at $w^\dagger$. There exists a faithful representation $\mathfrak{g}$ of the knowledge structure of player i in Γ_G^i if the set $(\mathfrak{g}, \wedge)$ is generated by the space*

$$\mathfrak{g} = \left\{ \left| \omega_{e_{-i,i}} \right\rangle, \omega_{e_{i,-i}} := \left\langle \omega_{e_{-i,i}} \right| \right\}$$

[38] The distributive law states that $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$ and $a \vee (b \wedge c) = (a \vee b) \wedge (a \wedge c)$.

over the complex field $\mathbb{K} = \mathbb{C}$ such that:
(1) $\omega_{e_{J,-J}}(s^i)\omega_{e_{-J,J}}(s^i) = \omega_{e_{J,J}}(s^i)$, $\forall s^i \in \boldsymbol{supp}(\sigma^i)$;
(2) $\langle \omega_{e_{-i,i}} | \omega_{e_{-i,i}} \rangle = 1$ *is the identity element i.e.* $a \wedge 1 = a, \forall a \in (\mathfrak{g}, \wedge)$;
(3) The multiplicative operation $\wedge$ *is such that* $\langle \omega_{e_{-i,i}} | \wedge | \omega_{e_{-i,i}} \rangle = \langle \omega_{e_{-i,i}} | \omega_{e_{-i,i}} \rangle$ *and* $|\omega_{e_{-i,i}}\rangle \wedge \langle \omega_{e_{-i,i}}| = |\omega_{e_{-i,i}}\rangle \langle \omega_{e_{-i,i}}|$ *is an orthogonal operator;*
(4) $\sigma^i(s^i) = \left|\omega_{(w,w^\dagger)}(s^i)\right|^2, \forall s^i$ *(Born rule) with* $\omega_{(w',w')} = \sigma^i \in \Omega \setminus \mathfrak{g}$, *for* $w' = w, w^\dagger$.

Proof. We first state the following lemma.

Lemma 3 *Fix the empirical canonical self-projective Krikpean model for player* i *in* G, $\mathcal{M}_G^{iO}$, *where* σ^i *has been determined as being true. Then the set* $(R^i, \wedge)$ *generated by the transitive closure of* $R^i := R^{i_i} \bigcup R^{i_{-i}}$ *must have the multiplicative operation* $\wedge :=$*"and" such that:*
1. $(wR^iw) \wedge (wR^iw) = (wR^iw)$ *(idempotence);*
2. $(w'R^iw) \wedge (wR^iw') \Rightarrow (w'R^iw')$ *(transitivity).*

Proof. We first check that any $(w, w') \in \mathcal{W} \times \mathcal{W}$ lies in R^i. By Theorem 1, in equilibrium, there is common knowledge, $K_*^i E^\sigma = \{w^{i_i}, w^{i_{-i}}\}$, which amounts to saying that the transitive closure $R^i := R^{i_i} \cup R^{i_{-i}} = \mathcal{W} \times \mathcal{W}$. Hence, by definition of the transitive closure, this implies that the set R^i is equipped with the non-commutative operation $\wedge : \mathcal{W} \times \mathcal{W} \to \mathcal{W}$ such that (1)-(2) hold. **End of proof**.
The above lemma shows that any $(w, w') \in \mathcal{W} \times \mathcal{W}$ lies in R^i. Next, we prove the existence of a faithful representation $\mathfrak{g}$ together with its characterization. We construct an isomorphism h for the canonical model with $h : (R^i, \wedge) \to (\mathfrak{g}, \wedge)$ as follows. Let $\omega_{(w,w^\dagger)} : S^i \to \mathbb{C}^{m_i}$ with $h((wR^iw^\dagger))(s^i) = \omega_{(w,w^\dagger)}(s^i)$, $\forall s^i$ so that $h((wR^iw^\dagger)) = |w\rangle, h((w^\dagger R^iw)) = \langle w|$, and the multiplicative map $\wedge$ over $\mathfrak{g}$ is defined such that $h((w^\dagger R^iw) \wedge (wR^iw^\dagger)) = h(w^\dagger R^iw)) \wedge h((wR^iw^\dagger)) = \langle w| |w\rangle$ and $h((wR^iw^\dagger) \wedge (w^\dagger R^iw)) = h(wR^iw^\dagger)) \wedge h((w^\dagger R^iw)) = |w\rangle \langle w|$ and we set $h((w^\dagger R^iw^\dagger)) = 1$, which implies that $h((wR^iw)) = \Pi_{\omega_{(w,w^\dagger)}}$ is an orthogonal projector.
We check idempotence (1) as follows:

$$h((wR^iw) \wedge (wR^iw)) = |w\rangle \underbrace{\langle w| \wedge |w\rangle}_{1} \langle w| = |w\rangle \langle w|$$

and $h((w^\dagger R^iw^\dagger) \wedge (w^\dagger R^iw^\dagger)) = \langle w| |w\rangle \wedge \langle w| |w\rangle = 1$ since 1 is the identity element. Thus, we have constructed an isomorphism that induces a set $\mathfrak{g}$ with: $\omega_{((w',w)\wedge(w,w'))}(s^i) = \omega_{((w',w))}(s^i)\omega_{((w,w'))}(s^i), \forall s^i$, whenever $w' \neq w$, where $\omega_{((w',w))}(s^i)$ is the complex conjugate of $\omega_{((w,w'))}(s^i)$. The condition that $h((w^\dagger R^iw^\dagger)) = 1$ implies that $\omega_{((w,w^\dagger))}$ must lie in the complex unit sphere and that the induced vectors, $(\omega_{((w',w)\wedge(w,w'))}(s^i))_{s^i \in S^i} = (\omega_{((w',w'))}(s^i))_{s^i \in S^i}$, are in the unit simplex Δ^i, $\forall w' = w^\dagger, w$. This means that the property $\omega_{((w',w'))} := (\omega_{((w',w'))}(s^i))_{s^i \in S^i}$ with $\omega_{((w',w'))} = \sigma^i$, $\forall w'$ is met, which immediately implies the Born rule (4). **End of proof**.

Necessity of a faithful representation in the canonical model. Next, we show that the existence of an empirical distribution e^i in the canonical model implies necessarily the faithful representation given in Proposition 2. We first start with a Lemma which proves that the Indivisibility Axiom is equivalent to the "Projection postulate" of QM.
Proof of Lemma 2. Given $\widehat{S}^i \subseteq S^i$, let $\Pi_{\widehat{S}^i} := \sum_{s^i \in \widehat{S}^i} \Pi_{a^{s^i}}$ so that

$$H_{\Pi_{\widehat{S}^i}} = \left\{\Psi \in \mathbb{H} : \Pi_{\widehat{S}^i}\Psi = \Psi\right\}$$

is a closed subspace of $\mathbb{H}$ which corresponds to statement $\widehat{S}^i$:="$\widehat{S}^i$ *is a rational subset of strategies for player* i". It is well-known that a set of subspaces like $H_{\Pi_{\widehat{S}^i}}$ ordered by inclusion $\subseteq$, with the complement $\sim$ defined by orthocomplementation $\perp$ and meet and join operations defined by intersection $\cap$ and direct sum $+$ of subspaces forms an orthocomplemented and non-distributive lattice. Hence, we can also identify the logical connectives of the lattice $\mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right)$ in terms of the lattice of projectors, $\mathcal{P}(\mathbb{H})$, with $H_{\Pi_{\widehat{S}^i}} \bigcap H_{\Pi_{\widehat{S}'^i}} \longleftrightarrow \Pi_{\widehat{S}^i} \wedge \Pi_{\widehat{S}'^i} = \Pi_{\widehat{S}^i}\Pi_{\widehat{S}'^i}$, $H_{\Pi_{\widehat{S}^i}} \subseteq H_{\Pi_{\widehat{S}'^i}} \longleftrightarrow \Pi_{\widehat{S}^i}\Pi_{\widehat{S}'^i} = \Pi_{\widehat{S}'^i}\Pi_{\widehat{S}^i} = \Pi_{\widehat{S}^i}$, $H_{\Pi_{\widehat{S}^i}} + H_{\Pi_{\widehat{S}'^i}} \longleftrightarrow \Pi_{\widehat{S}^i} + \Pi_{\widehat{S}'^i} - \Pi_{\widehat{S}'^i}\Pi_{\widehat{S}'^i}$, $(H_{\Pi_{\widehat{S}^i}})^{\perp} \longleftrightarrow \Pi_{(\widehat{S}^i)^{\sim}} = I - \Pi_{\widehat{S}^i}$, Hence, there exists a one to one mapping,

$$g : \mathcal{P}(\mathbb{H}) \to \mathcal{L}\left(\mathcal{S}_i^O, \vee, \wedge, \sim\right), \Pi_{\widehat{S}^i} \overset{g}{\mapsto} \bigvee_{s^i \in \widehat{S}^i} K_i(E^{s^i} \cap E^{\sigma^{-i}}),$$

with $\Pi_{a^{s^i}}\Pi_{a^{s'^i}} = o_2$, $\forall s^i \neq s'^i$, $g(o_2) = \emptyset$, $g(I) = \mathcal{S}_i^O$ and $\sum_{s^i \in \mathtt{supp}(\sigma^i)} \Pi_{a^{s^i}} = I$. This implies the existence of a measure μ such that $e^i \circ g(\Pi_{a^{s^i}}) = \mu(\Pi_{a^{s^i}})$.
End of proof.
Proof of Theorem 1 (1-2) We prove that the set of all equilibrium local intrinsic states of player i, $\omega_{(w^{i-i}, w^{i_i})}$, in a faithful representation $\mathfrak{g}$ of Γ_G^i are in the complex unit sphere $\mathbb{S}$ and forms a ray. To see that every $\omega_{(w^{i-i}, w^{i_i})} \in \mathfrak{g}$ lies in the unit sphere $\mathbb{S}$ of $\mathbb{H} = \mathbb{C}^{m_i}$, it suffices to note that if i is in an intrinsic equilibrium $\sigma = (\sigma^i, \sigma^{-i})$, then a faithful representation of Γ_G^i implies that the (co)-vectors of weights $\omega_{(w^\dagger, w^\dagger)}$ and $\omega_{(w,w)}$ must fulfill the equality,

$$\omega_{(w^\dagger, w)}(s^i)\omega_{(w, w^\dagger)}(s^i) = \omega_{(w^\dagger, w^\dagger)}(s^i),$$

with $\omega_{(w^\dagger, w^\dagger)}(s^i) = \sigma^i(s^i)$. Together with the condition that σ^i is in the unit simplex, conditions (i) and (ii) imply that $\sum_{s^i \in \mathtt{supp}(\sigma^i)} \omega_{(w^\dagger, w)}(s^i)\omega_{(w, w^\dagger)}(s^i) = 1$ where $\omega_{(w, w^\dagger)}(s^i) = \overline{\omega}_{(w^\dagger, w)}(s^i)$ is the complex conjugate of $\omega_{(w^\dagger, w)}(s^i)$. From this last condition, we infer that $\omega_{(w, w^\dagger)}$ is in $\mathbb{S}$ and $\omega_{(w^\dagger, w)}$ is its co-vector. This shows that in equilibrium, the intrinsic state space of i (at w) is in $\mathbb{S}$. Finally, the fact that the set of vectors $\omega_{(w, w^\dagger)}$ lying in $\mathfrak{g}$ forms a one-dimensional subspace of $\mathbb{S}$ follows easily from the Born rule. **End of proof**.
Proof of Theorem 1 (3) (orthogonal projector)
The analysis of the choice of player i in terms of a probability measure over a set of projectors requires the following definitions.
First, we have to define a measure μ on the complex Hilbert space $\mathbb{H} = \mathbb{C}^{m_i}$.

Let $\mathcal{P}(\mathbb{H})$ be the set of orthogonal projectors of $\mathbb{H}$. Consider a unit vector a_{s^i} and the associated one dimensional projector $\Pi_{a_{s^i}}$. A measure is a mapping which assigns a non-negative real number $\mu(\Pi_{a_{s^i}})$ to each projector $\Pi_{a_{s^i}}$. Hence, we can identify an empirical model, $e^i(\cdot\,|\mathfrak{g})$, as a mapping $\mu : \mathcal{P}(\mathbb{H}) \to \mathbb{R}$ by setting $\mu(\Pi) = f(\mathtt{Im}(\Pi))$. This mapping must be defined such that, if $\{a_{s^i}\}$ are mutually orthogonal, then the measure of $\sum_{s^i \in S^i} \Pi_{a_{s^i}}$ has to satisfy the (sub-)additivity property $\mu(\sum_{s^i \in S^i} \Pi_{a_{s^i}}) = \sum_{s^i \in S^i} \mu(\Pi_{a_{s^i}})$. Any such measure is determined by its values on the one dimensional projections such that:
(1) If $\left\{\Pi_{a_{s^i}}\right\}$ is a family of pairwise orthogonal projectors i.e. $\Pi_{a_{s^i}}\Pi_{a_{s^{i'}}} = 0$ for $s^i \neq s^{i'}$ with $\sum_{s^i \in S^i} \Pi_{a_{s^i}} = I$, then $\left\{\mu(\Pi_{a_{s^i}})\right\}$ is summable with sum $\mu(\mathtt{id})$.
(2) If $\left\{\Pi_{a_{s^i}}\right\}$ is a family of pairwise orthogonal projectors i.e. $\Pi_{a_{s^i}}\Pi_{a_{s^{i'}}} = 0$ for $s^i \neq s^{i'}$, then $\left\{\mu(\Pi_{a_{s^i}})\right\}$ is summable with $\sum \mu(\Pi_{a_{s^i}}) = \mu(\sum \Pi_{a_{s^i}})$.
In this case μ is a *measure* on $\mathcal{P}(\mathbb{H})$.
Next, we will use Gleason's Theorem in order to establish a contradiction, namely that if we use a non canonical self-projective model, with a set of perspectives greater than two, then we cannot derive a well-defined probability measure, e^i.
Gleason's Theorem [33]: *In a Hilbert space of finite dimension $d \geq 3$, any bounded measure on the orthogonal projectors of $\mathbb{H}$ is such that $\Pi \to \mathtt{tr}(A\Pi)$ where A is a linear Hermitian operator of $\mathbb{H}$.*
In the particular case where the measure is assumed to be a probability measure, Gleason's Theorem implies that,

$$\mu(\Pi) = \mathtt{tr}(\rho\Pi),$$

where ρ is a density operator. We have the following lemma.

Lemma 4 *Fix the empirical canonical Krikpean model for player i in G, $\mathcal{M}_G^{iO}$, where σ^i has been determined as being true at perspective w^{i_i}. There exists an empirical distribution $e^i(w^{s^i}) = \mathtt{tr}(\rho\Pi_{a_{s^i}})$ with $\rho = \Pi_{\omega_{e_{-i,i}}}$ if and only if we have the faithful representation with $\rho \in (\mathfrak{g}, \wedge)$.*

Proof. "*Only if*". We have the following lemmata:
lemmata *Suppose an equilibrium σ of G has been determined as true in a self-projective model, $\mathcal{M}_G^i$. If player i has more than two pure strategies i.e. $m_i \geq 3$, then, any empirical distribution of $\mathcal{M}_G^{Oi}$ must be of the form, $e^i(w^{s^i}) = \mathtt{tr}(\rho\Pi_{a_{s^i}})$ with ρ a self-adjoint endomorphism of trace 1 (a density matrix indeed).*
Proof. By Lemma 2, $e^i(w^{s^i}) = \mu(\Pi_{a_{s^i}}), \forall s^i$ with μ a bounded measure. By Gleason's Theorem [33] we have that $\mu(\Pi_{a_{s^i}}) = \mathtt{tr}(\rho\Pi_{a_{s^i}})$ with ρ a self-adjoint endomorphism. The necessity to have a trace 1 follows since we want a probability measure. End of proof.
The above results allow to conclude that the global intrinsic state of i is necessarily $\rho = h((w, w^\dagger)) \wedge h((w^\dagger, w)) \in (\mathfrak{g}, \wedge)$. The fact that $h((w, w^\dagger)) \wedge h((w^\dagger, w))$ is an orthogonal projector, $\Pi_{\omega_{(w,w^\dagger)}} = \left|\omega_{(w,w^\dagger)}\right\rangle \left\langle\omega_{(w,w^\dagger)}\right|$, follows immediately

from the fact that $\omega_{(w,w^{\dagger})}$ is a unit vector by property (1) of Theorem 1 above.
Proof of Theorem 1 (4) (Born rule)
By definition of an empirical distribution, we must have $e^i(w^{s^i}) = \sigma^i$. If we have a faithful representation $\mathfrak{g}$, proposition 2 implies that $e^i(w^{s^i}) = \overline{h}((w^{'}, w))(s^i)h((w^{'}, w))(s^i)$, $\forall s^i$, which can be rewritten as

$$e^i(w^{s^i}) = \left|\omega_{(w,w^{\dagger})}(s^i)\right|^2.$$

End of proof.
Lemma 5 establishes the existence of an empirical distribution in the canonical model and states that we must have a faithful representation of the self-projective graph to have an empirical distribution. We now give the proof that an empirical distribution does not exist for self-projective models $\mathcal{M}_G^i$ with more than two perspectives, whenever $m_i \geq 3$.
Before we embark in the proof of this result, we need some preliminary definitions.

Definition 8 Let $\mathcal{M}_G^{'i}$, and $\mathcal{M}_G^i$, be two self-projective models for G with $\mathcal{W}^{'} \subset \mathcal{W}$. If A^{σ^i} can be determined as being relatively true in both models, then we say that $\mathcal{M}_G^i$, is a **reducible** Krikpean model for G. When $\mathcal{M}_G^i$, is not reducible it is said to be **irreducible**.

We will refer to the irreducible self-projective model with $\mathcal{W} = \left\{w_i^i, w_{-i}^i\right\}$ as the **canonical** self-projective model for G.

Lemma 5 *In a reducible self-projective model for a game G, $\mathcal{M}_G^i$, with $m_i \geq 3$ if player i determines a rational strategy σ^i as being relatively true i.e. a mixed NE profile $(\sigma^i)_{i\in N}$, then $\mathcal{M}_G^i$ has no empirical distribution $e^i = \sigma^i$.*

Proof. *Step 1.* Every self-projective model of G, $\mathcal{M}_G^i$, with a set of $m > 2$ perspectives, is reducible.
Consider a self-projective model of G, $\mathcal{M}_G^i$, with its associated weighted directed self-projective graph, $\Gamma_G^i = \langle \mathcal{W}, \mathcal{E}, \Omega\rangle$ where $\mathcal{W} = \left\{w_J^i : J \in M\right\}$ is an arbitrary class of subsets of W with a cardinality $|\mathcal{W}| > 2$. This induces a graph Γ_G^i with $m > 2$ vertices. Theorem 1 asserts that relative truth σ^i can indeed be determined in any such a model. Hence, this immediately shows that $\mathcal{M}_G^i$ is indeed reducible.
Step 2. We prove that any $\mathcal{M}_G^i$ that is reducible has no empirical distribution e^i. By step 1, any $\mathcal{M}_G^i$ that is reducible has a set of perspectives of cardinality greater than 2. Next, our aim is to apply Gleason's Theorem [33] in order to show that e^i cannot form a probability measure—hence an empirical distribution—when the set of vertices of Γ_G^i is of cardinality greater than 2. Let Π be projection operator Π on a Hilbert space $\mathbb{H}$.[39] Expand the unit vector of weights, $\left|\omega_{(w^{i_K}, w^{i_L})}\right\rangle \in \mathbb{C}^{m_i}$ in the standard basis, $\{a_{s^i}\}_{s^i\in S^i}$, as $\left|\omega_{(w^{i_K}, w^{i_L})}\right\rangle = \sum_{s^i\in S^i} \left\langle a_{s^i} \mid \omega_{(w^{i_K}, w^{i_L})}\right\rangle |a_{s^i}\rangle$. Lemma 5 tells us that $e^i(w^{s^i}) = \mu(\Pi_{a_{s^i}})$. Hence, the empirical distribution must be a function

[39] In other words, Π is an idempotent and self-adjoint operator.

of $\Pi_{a_{s^i}}\omega_{(w^{i_K},w^{i_L})}$, for $\omega_{(w^{i_K},w^{i_L})} \in \mathbb{H}$, with $(w^{i_K}, w^{i_L}) \in \mathcal{E}$. Thus, there must exist a function f such that: $e^i(w^{s^i}) = f(\Pi_{a_{s^i}}\omega_{(w^{i_K},w^{i_L})}), \forall s^i$. A direct application of Gleason's Theorem allows to conclude that the unique empirical probability measure e^i is of the form

$$e^i(w^{s^i}) = \left\|\Pi_{a_{s^i}}\omega_{(w^{i_K},w^{i_L})}\right\|^2 \forall s^i.$$

Next, we note that the Indivisibility axiom implies necessarily the Born rule in the canonical model, without using Gleason's Theorem. From this, we will be able to show that the structure of the graph of any non canonical model does not induce the Born rule, which allows to conclude that such model does not induce a well-defined empirical distribution e^i as a consequence of Gleason's Theorem.

The axiom of Indivisibility implies that the intrinsic state of player i is $w_0 \in E^{s^i} = K_{i_i \otimes i_{-i}} E^{s^i}$ iff we have $\Pi_{a_{s^i}} \left|\omega_{(w,w^\dagger)}\right\rangle$ for i_{-i} and $\Pi_{a_{s^i}} \left(\left\langle\omega_{(w,w^\dagger)}\right|\right)^\dagger$ for i_i. In the canonical self-projective model, this immediately implies that, $e^i(w^{s^i}) = \left\|\Pi_{a_{s^i}} \left|\omega_{(w,w^\dagger)}\right\rangle\right\|^2$, where $\|\cdot\|$ is the norm derived from the inner product on $\mathbb{C}^{m_i}$. This by definition is the Born rule. This of course also requires to have the faithful representation of the canonical graph given in Theorem 1. Hence, Gleason's Theorem and the axiom of Indivisibility allows to conclude that the Born rule cannot be constructed from a self-projective graph with more than two perspectives. This can be seen by noting that if Γ^i_G has more than two vertices, then the empirical distribution is

$$e^i(w^{s^i}) := \mu(\Pi_{a_{s^i}}), \text{ such that } e^i(w^{s^i}) \neq \left\|\Pi_{a_{s^i}}\omega_{(w^{i_K},w^{i_L})}\right\|^2.$$

This immediately implies that the graph must have a *unique* pair of weights and hence $\mathcal{M}^i_G$ must be canonical i.e. irreducible. If not, either the indivisibility axiom is not met, or Γ^i_G is not well-defined i.e. it does not assign some weights at every edge, which entails that h does not form an isomorphism. From this we conclude that e^i never forms a well-defined probability measure for any reducible self-projective models.

That a canonical model with a faithful representation induces necessarily an empirical distribution is shown in the proof of Theorem 1 (3) below. **End of Proof.**

Since any reducible model has no empirical distribution, the rest of the appendix is devoted to the study of canonical self-projective models. Henceforth it is convenient to set $w^{i_i} := w^\dagger$ and $w^{i_{-i}} := w$.

Proof of Corollary 1 By Theorem 1, any global equilibrium intrinsic state of player i is given by the density matrix $\left|\omega_{(w^{i_{-i}},w^{i_i})}\right\rangle \left\langle\omega_{(w^{i_{-i}},w^{i_i})}\right|$. This operator can be rewritten as,

$$\sum_{s'^i \in \mathtt{supp}(\sigma^i)} \left|\omega_{(w^{i_{-i}},w^{i_i})}(s'^i)\right|^2 \mid a_{s'^i}\rangle \langle a_{s'^i} \mid + \sum_{s^i \in \mathtt{supp}(\sigma^i): s^i \neq s'^i} \omega_{(w^{i_{-i}},w^{i_i})}(s^i)\omega_{(w^{i_{-i}},w^{i_i})}(s'^i) \mid a_{s^i}\rangle \langle a_{s'^i} \mid.$$

By the classical indifference condition, σ^i is an equilibrium strategy for player i if and only if very $s^i \in \mathtt{supp}(\sigma^i)$ is a best reply. On the other hand, a faithful representation implies that in the empirical model, the event,

$$\left\| \Box^{i_i} A^{s^i} \wedge \Box^{i_{-i}} A^{s'^i} \right\|^{\mathcal{M}_G^i},$$

has a non-zero (complex) off-diagonal weight given by,

$$\omega_{(w^{i_i}, w^{i_{-i}})}(s^i)\omega_{(w^{i_{-i}}, w^{i_i})}(s'^i) \neq 0,$$

for all pair of pure actions, $s^i, s'^i \in \mathtt{supp}(\sigma^i)$. This proves the existence of off-diagonal terms. **End of proof**

Proof of Corollary 2

Sufficiency. By Theorem 1, any mixed strategy σ_ω^i gets identified by an orthogonal projector, Π_ω, with complex off-diagonal terms. Hence, if player i is in a *pure* intrinsic state Π_ω with a probability p_ω, we have an empirical distribution of the form, $e^i = \sum_\omega p_\omega \sigma_\omega^i$.

Necessity. By contradiction. Suppose that the global intrinsic state of player i cannot be expressed by $\sum_\omega p_\omega \Pi_\omega$. This means that player i does not determine σ_ω^i as a rational strategy with a probability p_ω. By definition, this implies that $e^i \neq \sum_\omega p_\omega \sigma_\omega^i$. **End of proof**.

Acknowledgements I am grateful to two referees and one editor for helpful input, to Sonia Smets for her organization of the workshop 'Quantum Physics Meets TARK' in Groningen (July 2011), to Samson Abramsky, Alexandru Baltag, Bob Galesloot and Emmanuel Haven for valuable conversations, as well as the participants of seminars at the Department of Economics of Maastricht and the University of Leicester for valuable comments. Finally, I d like to express my thanks to Adam Brandenburger, Pierfrancesco La Mura and Andrés Perea for their words of encouragement.

References

1. Von Neumann, J., & Morgenstern, 0. (1944). *Theory of games and economic behaviour*. Princeton University Press.
2. Aumann, R.J. (1987). Correlated equilibrium as an expression of Bayesian rationality, *Econometrica 55*, 1-18.
3. Nash, J. (1950). Equilibrium points in n-person games. *Proc. Natl. Acad. Sci, 36* 48-49 (1950).
4. Sakurai, J. (1994). *Modern quantum mechanics.* Addison-Wesley, revised edition.
5. Mandel, L. (1999). Quantum effects in one-photon and two-photon interference. *Rev. Mod. Phys. 71*, 274-282.
6. Kochen, S., & Specker, E.P. (1967). The problem of hidden variables in quantum mechanics, *Journal of Mathematics and Mechanics, 17.* 59-87.
7. Colbeck, R. & R. Renner, R. (2011). No extension of quantum theory can have improved predictive power, *Nature Communication, 2.* 4-11.
8. Colbeck, R. & R. Renner, R. (2012). Is a system's wave function in one-to-one correspondence with its elements of reality?. *Physical Review Letters, 108*, 150-402.
9. Brandenburger. A. (2007). The power of paradox: Some recent developments in interactive epistemology. *International Journal of Game Theory*, 35, 465-492.
10. Brandenburger A. (2008). *Epistemic Game Theory: An Overview*. in The New Palgrave Dictionary of Economics, 2nd edition, Palgrave Macmillan.
11. Feynman, R.P. Leighton, R.B., & Sands, M. (1965). The Feynman lectures on physics. (Addison-Wesley, Reading, MA).

12. Khrennikov, A. (2004). Probabilistic pathway representation of cognitive information. *Journal of Theoretical Biology, 231*, 21.
13. Lukasiewicz J.(1930). *Philosophical remarks on many-valued systems of propositional logic*. Reprinted in Selected Works (Borkowski, ed.), Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, 1970, 153-179.
14. Conway, J., & Kochen, S. (2006). The free will theorem. *Foundations of Physics, 36,* 1441-1473.
15. Conway, J., & Kochen, S. (2009). The strong free will theorem. *Notices of the American Mathematical Society, 56*, 226-232.
16. Wheeler, J. A. (1983). 182-213 *in Wheeler and Zurek.*
17. Zurek, W. H. (2009). Quantum Darwinism, arXiv:0903.5082; *Nature Physics 5,* 181-188 (2009).
18. Gisin, N. (2012). Impossibility of covariant deterministic nonlocal hidden-variable extensions of quantum theory. *Physical Review A* 83:020102.
19. Eisert, J, Wilkens. M., & Lewenstein, M. (1999). Quantum games and quantum strategies. *Physical Review Letters, 83*, 3077-3080.
20. Pietarinen, A. (2001). Propositional logic of imperfect information: Foundations and Applications. *Notre Dame Journal of Formal Logic* 42(4), 193-210. ISSN 0029-4527.
21. Pietarinen, A. (2002). Quantum logic and quantum theory in a game-theoretic perspective. *Open Systems & Information Dynamics, vol. 9*, 273-90. Zbl 01812492. MR 1923231. 208.
22. Brandenburger A. (2010). The relationship between quantum and classical correlation in games. *Games and Economic behavior, 69*, 175-183.
23. Zeilinger, A. (1999). A foundational principle for quantum mechanics. *Foundation of Physics 29*, 631-643.
24. Hardy, L., & Spekkens R. (2010). Why Physics Needs Quantum Foundations. *Physics in Canada*, 66 (2), 73-76.
25. Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings of the Royal Society of London A*, *455,* 3129-3137.
26. Osborne M.J., Rubinstein. A. (1994). *A Course in Game Theory.* MIT Press.
27. Kolhberg, E., & Mertens J.-F. (1986). On the Strategic Stability of Equilibria, *Econometrica,54*, 1003-1037.
28. Luce, R. D., & Raiffa, H. (1957). *Games and Decisions.* New York: John Wiley.
29. Kripke, S. (1963). Semantic analysis of modal logic. *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik 9,* 67-96.
30. Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy, 72,* 690-716.
31. Birkhoff, G., von Neumann, J. (1936). The logic of quantum mechanics. *Annals of Mathematics, 37*, 823.
32. Dirac, P. (1958). *Quantum Mechanics* (Clarendon Press,Oxford).
33. Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics, 6*, 885-894.
34. Shubik, M. (1992). *Game Theory at Princeton*, 1949-1955: A Personal Reminiscence, in Weintraub, E.R. (ed.), Toward a History of Game Theory. Duke University Press, 151-164.
35. Wolfe, P. (1955). *Report of an informal conference on Recent Developments in the Theory of Games.* Department of Mathematics (Logistics Research Project), Princeton University, January 31-February 1.
36. Von Neumann, John. Sur la Theorie des Jeu. *Comptes Rendus, Academie des Sciences*, 1928, 1689-91. "Zur Theorie der Gesellschaftsspiele," Mathematische Annalen 1928, 100, 295-320; translated by S. Bargmann as "On the Theory of Games of Strategy," in Contributions to the theory of games. Vol. 4. Eds.: Albert Tucker and R. Duncan Luce. Princeton: Princeton University. Press, 1959, 13-42.
37. Leonard R. J. (1995). From Parlor Games to Social Science: Von Neumann, Morgenstern, and the Creation of Game Theory 1928-1944. *Journal of Economic Literature, Vol.* 33, No. 2, 730-761.
38. Belnap, N. (1992). *Synthese* 92:385-434.
39. Belnap, N., & Szabo L. (1996). Branching space-time analysis of the GHZ theorem, *Foundations of Physics, vol. 26*, no. 8, 989-1002.
40. Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future. Agents and choices in our indeterminist world.* Oxford: Oxford University Press.