1-10-2024

What Does '(Non)-absoluteness of Observed Events' Mean?

Emily Adlam
Chapman University, adlam@chapman.edu

Foundations of Physics (2024) 54:13
https://doi.org/10.1007/s10701-023-00747-1

RESEARCH

What Does '(Non)-absoluteness of Observed Events' Mean?

Emily Adlam[1,2]

Received: 14 September 2023 / Accepted: 29 November 2023

Abstract

Recently there have emerged an assortment of theorems relating to the 'absoluteness of emerged events,' and these results have sometimes been used to argue that quantum mechanics may involve some kind of metaphysically radical non-absoluteness, such as relationalism or perspectivalism. However, in our view a close examination of these theorems fails to convincingly support such possibilities. In this paper we argue that the Wigner's friend paradox, the theorem of Bong et al and the theorem of Lawrence et al are all best understood as demonstrating that if quantum mechanics is universal, and if certain auxiliary assumptions hold, then the world inevitably includes various forms of 'disaccord,' but this need not be interpreted in a metaphysically radical way; meanwhile, the theorem of Ormrod and Barrett is best understood either as an argument for an interpretation allowing multiple outcomes per observer, such as the Everett approach, or as a proof that quantum mechanics cannot be universal in the sense relevant for this theorem. We also argue that these theorems taken together suggest interesting possibilities for a different kind of relational approach in which *interaction* states are relativized whilst observed events are absolute, and we show that although something like 'retrocausality' might be needed to make such an approach work, this would be a very special kind of retrocausality which would evade a number of common objections against retrocausality. We conclude that the non-absoluteness theorems may have a significant role to play in helping converge towards an acceptable solution to the measurement problem.

Keywords Wigner's Friend · Perspectivalism · Retrocausality · Relationalism · Measurement problem

✉ Emily Adlam
eadlam90@gmail.com

1 Philosophy Department, Chapman University, Orange, CA 92866, USA

2 The Rotman Institute of Philosophy, 1151 Richmond Street, London, ON N6A5B7, Canada

Published online: 10 January 2024

1 Introduction

Recently there have emerged an assortment of theorems relating to the 'absoluteness of emerged events.' Various interpretations of these results are possible, but one influential school of thought suggests these theorems demonstrate that if certain kinds of experiments were to have the results predicted by unitary quantum mechanics, then we would most likely have to accept that even observed events are not 'absolute'[1][1–3]. For example, it has been suggested that this could involve a picture in which measurements only have definite outcomes relative to individual observers - i.e. there is no absolute, 'third-person' view from which we can say which outcome a measurement actually had[4, 5].

Now, clearly one way in which measurements could be 'non-absolute' in this way would be if a given observer sometimes observes more than one outcome for a given measurement, as for example in the Everett interpretation or other multiple-outcome-per-observer (MOPO) approaches. But the non-absoluteness theorems have sometimes been used to argue for some *other* kind of non-absoluteness - in particular, for a version of relationalism or perspectivalism in which there is only one outcome per observer per measurement, but different observers may disagree about the outcome of a given measurement. However, in our view a close examination of these theorems fails to convincingly support such a possibility. In this paper we will argue that the Wigner's friend paradox[6], the Bong et al theorem[7] and the Lawrence et al theorem[8] do indeed demonstrate that if quantum mechanics is universal, and if certain auxiliary assumptions hold, then the world inevitably includes various forms of 'disaccord,' by which we mean circumstances in which observers may reasonably fail to agree about the outcome of a given measurement - but there is no compelling reason to interpret this 'disaccord' in terms of metaphysically radical forms of relationalism or perspectivalism. Meanwhile, we find that the theorem of Ormrod and Barrett[9] could be regarded as offering a genuine argument for metaphysically radical non-absoluteness, but this is achieved only by making an unusually strong assumption about the circumstances in which unitary quantum mechanics makes correct predictions, and our view is that the theorem should be interpreted either as an argument in favour of MOPO approaches or simply as a reductio ad absurdum against this assumption.

However, this does not mean that the non-absoluteness theorems are not useful; on the contrary, we believe that the emphasis on metaphysically radical interpretations of these theorems may be obscuring some very important lessons that could be drawn from them. In particular, we argue that the Wigner's friend paradox could be regarded as demonstrating that if quantum mechanics is universal then *interaction states* must be relativized, even though the events that actually occur are absolute. We note that the non-absoluteness theorems make it clear that such an approach would either have to violate Locality or would have to exhibit something like

[1] We emphasize that this response is not necessarily endorsed by the original authors of the theorems, who in many cases adopt an attitude of neutrality towards the various possible ways one could respond to their theorem.

superdeterminism or retrocausality; but they also help us see that the kind of 'retrocausality' required is of a very special kind, such that common objections against retrocausality may not apply to it. We also demonstrate that this vision of quantum mechanics with relational interaction states and retrocausality is already realised by seversal existing interpretations. We conclude that the non-absoluteness theorems have significantly narrowed the space of viable interpretations of quantum mechanics by demonstrating that workable approaches must have some quite specific properties, so they may play a significant role in helping us converge towards an acceptable solution to the measurement problem.

2 Background

The non-absoluteness theorems are descendants of the Wigner's friend paradox[6], which describes a scenario in which some observer, Chidi, performs a measurement $\{|S_i\rangle\langle S_i|\}$ on a system S, and then another observer, Alice, performs a measurement on the joint system of Chidi and S. The supposed paradox is that Chidi will presumably have seen a definite outcome to his measurement, so he will ascribe some state $|S_i\rangle$ to the system S, and yet if we believe that quantum mechanics is universal then the correct way to describe the interaction between Chidi and S is to say that they end up in a state $\psi = \sum_i c_i|C_i\rangle|S_i\rangle$, where $|C_i\rangle$ is the state of Chidi corresponding to him having seen the outcome $|S_i\rangle$ to his measurement. Moreover, if this experiment is repeated many times Alice can in principle confirm that Chidi and S do end up in this state, provided that she is able to maintain complete coherent control of the joint Chidi-S system long enough to perform tomographic measurements. Yet the state ψ appears to represent Chidi as not obtaining any definite outcome to his measurement, so how can it be the case that Chidi has observed a definite outcome if Alice subsequently finds him in the state ψ ?

As just stated this 'paradox' is perhaps puzzling but does not lead to an outright contradiction, so the class of theorems that we will refer to as the 'non-absoluteness theorems' have set out to provide something more watertight using 'extended Wigner's friend' scenarios. The first such theorem that we will discuss is due to Bong et al[7]. It should be emphasized that Bong et al originally derived their theorem in a theory-independent way, arriving at some inequalities which must be obeyed by any theory satisfying their assumptions Absoluteness of Observed Events, Locality and No-Superdeterminism. But in this paper we are specifically interested in understanding whether Bong et al's theorem can be used to argue that the universality of quantum mechanics would entail the existence of some kind of non-absoluteness, so we will follow the presentation of ref [10] and work directly in the context of quantum theory. Thus suppose we have two agents, Chidi and Divya, each in a closed laboratory, and each in possession of one particle from an entangled pair. Chidi and Divya each perform a measurement of a certain fixed observable of their particle, obtaining measurement results C, D. Then we have another observer Alice who performs a 'supermeasurement' on the whole system of Chidi's closed lab, obtaining an outcome A; and a fourth observer, Bob, who performs a supermeasurement on Divya's closed lab, obtaining an outcome B.[2] The word 'supermeasurement' here just refers to the fact that one observer is performing a measurement on another observer, and that this is being done in a basis which does not commute with the variables encoding that observer's memories, so for example here Alice is performing on Chidi a measurement which does not commute with the variable encoding the result of Chidi's measurement C, and likewise for Bob and Divya. We use this terminology to emphasize that in practice, performing such a measurement requires Alice to maintain complete coherent control over Chidi and his system, which means she must ensure that no information whatsoever escapes Chidi's closed laboratory, and she must be able to exert fine control over each individual degree of freedom making up Chidi and his system.

Bong et al argue that 'absoluteness of observed events' (AOE) means, first of all, that there are is exactly one outcome of each measurement per observer, so in this case we have exactly four measurement outcomes witnessed by the person who performs the relevant experiment - A, B, C and D. In addition, Bong et al state that AOE also entails that if Alice, instead of performing a supermeasurement, chooses to perform a measurement which tells her the result of Chidi's measurement (e.g. by simply going into the laboratory and asking Chidi what result he got) then the outcome of this measurement must match Chidi's outcome C; and likewise mutatis mutandis for Bob and Divya. Bong et al also invoke assumptions that they call Locality and No-Superdeterminism to argue that the values of C and D cannot depend on whether Alice and Bob choose to perform their supermeasurements or to ask Chidi and Divya what results they obtained, so we have a fixed set of four outcomes A, B, C, D regardless of what measurements Alice and Bob perform. Now, let us suppose that after the experiment Bob shares his result with Alice; thus a single observer, Alice, could ultimately come to know the values of any one of the pairs in the set $\{AB, AD, CB, CD\}$, drawn from the fixed set of outcomes A, B, C, D. Then since we have chosen to work directly in the context of quantum theory, we may proceed on the basis of the assumption that quantum mechanics is universal, in the sense that Alice will never see an outcome which is in contradiction with unitary quantum mechanics; it then follows that in any run of this experiment the four fixed outcomes must match the predictions made by quantum mechanics for each of the pairs of variables in the set $\{AB, AD, CB, CD\}$. However, it is possible to choose the states and measurements used in this experiment such that there is no possible choice for the values of A, B, C and D which reproduces the predictions of unitary quantum mechanics for all of these pairs. Thus if it is really the case that unitary quantum mechanics is universal, and we are not willing to deny Locality or

[2] Ref [7] also includes the more general case in which Alice and Bob can choose between two or more supermeasurements, which is useful to show that the set of correlations obeying the inequalities derived from their assumptions is strictly larger than the set obeying inequalities obtained just from Bell's locality assumptions. However, only one supermeasurement per observer is needed to derive the 'non-absoluteness' result we are interested in here, so we will focus on that simple case.

No-Superdeterminism, it appears we must accept that observed events are not absolute, in the sense in which that term is used by Bong et al.[3]

Another such theorem was proved by Lawrence et al[8]. It uses a scenario similar to the Bong et al experiment, except that it employs GHZ states rather than Bell pairs, and there are only two agents: first Alice measures all three GHZ qubits, each in a fixed basis, and then Bob performs three different supermeasurements on the joint system composed of Alice and her three qubits, each in a fixed basis, with each of Bob's measurements accessing the Hilbert space associated linearly with one of the original qubits. Thus Lawrence et al argue that if there are 'relative facts' about the outcomes of these measurements relative to Alice and Bob respectively, this experiment must produce exactly six measurement outcomes, $A_1, A_2, A_3; B_1, B_2, B_3$, each relativized to the person performing the relevant measurement. Lawrence et al contend that these outcomes should obey certain constraints imposed by quantum mechanics: for example, they show that quantum mechanics predicts that with an appropriate labeling convention, the product of the values $B_1 B_2 B_3$ will always be equal to 1, and similar constraints can be obtained for the trios $B_1 A_2 A_3$, $A_1 B_2 A_3$, and $A_1 A_2 B_3$. But it is straightforward to show that there is no possible assignation of values to $A_1, A_2, A_3; B_1, B_2, B_3$ which obeys all four of these constraints, and thus Lawrence et al argue that 'relative facts' are inconsistent with quantum mechanics.

Although Lawrence et al are concerned with relative facts rather than absoluteness of observed outcomes here, their theorem can be rewritten so as to follow the same argument pattern as that of Bong et al. One could use AOE rather than the existence of relative facts to argue that there are exactly six measurement outcomes witnessed by the person who performs the relevant experiment; and second, one could follow the reasoning of Bong et al to argue that if Bob performs a measurement on Alice seeking to learn her measurement outcomes, he must learn the actual values of A_1, A_2 and/or A_3. Thus one could conclude that Bob is principle able to come to know the values of any one of the trios $\{B_1 B_2 B_3, B_1 A_2 A_3, A_1 B_2 A_3, A_1 A_2 B_3\}$,[4] and then one could invoke the universality of quantum mechanics in the first-person sense to argue that on any run of the experiment the values of the outcomes $A_1, A_2, A_3; B_1, B_2, B_3$ must obey the constraints imposed by quantum mechanics for

[3] Supermeasurements on real observers are not possible using current technology, and most likely will never be possible using any realistic technology, so this experiment is and will likely remain a thought experiment. However, a later version of this theorem[11] considers the possibility of replacing Chidi with something like an artificial intelligence, focusing on the conditions that would need to be met in order for us to agree that such a device has genuinely made an observation which we would naturally expect to be 'absolute.' This version of the experiment might well be performable at some time in the future. However, in this paper we will focus on the original Bong et al theorem, because as long as one is convinced that a real observation can be made by the relevant device in this experiment, its foundational consequences seem roughly the same as the consequences of the original experiment. One exception to this is that the two experiments may possibly have different consequences for approaches suggesting that the interpretation of quantum mechanics is linked in some intrinsic way to (human) consciousness, but we are not considering such possibilities here, and thus we think most of our analysis in this paper will also apply to the newer theorem.

[4] Obviously Bob could also come to know $B_1 B_2 A_3, B_1 A_2 B_3, A_1 B_2 B_3$ or $A_1 A_2 A_3$, but these trios are not needed for the proof.

all four trios, which is known to be impossible. Thus the Lawrence et al argument could be used just like the Bong et al argument to argue that the universality of quantum mechanics implies that observed events are not absolute. Note that Lawrence et al do not explicitly mention assumptions about superdeterminism and retrocausality, but an examination of their derivation makes it clear that they are in fact making an implicit assumption very similar to Bong et al's No-Superdeterminism assumption, since they take it that the quantum constraints must always be obeyed for all four trios, whereas if the values of A_1, A_2, A_3 were allowed to depend on what variables Bob chooses to measure then the quantum constraints would only need to be obeyed for the one trio whose value Bob does actually come to know. The Lawrence et al result also requires an assumption of Locality, because the proof requires us to assume that when Bob measures B_1 the value of B_1 is independent of whether he measures B_2 or A_2, and so on mutatis mutandis; so to make this argument work we probably need to imagine that 'Bob' is really several people spread out in space as in the Bong et al case, in order that we can apply Locality to conclude that the results of these measurements are independent.

Finally we have the theorem of Ormrod and Barrett[9], which uses a similar experimental setup to Bong et al. Ormrod and Barrett consider four different spacelike slices on which we could in principle collapse the wavefunction (one including the measurements of Alice+Bob, one including the measurements of Chidi+Divya, one including the measurements of Alice+Divya, and one including the measurement of Bob+Chidi), and show that if we try to make predictions for measurement outcomes in four different ways, by collapsing the wavefunctions on each of these four spacelike slices, then certain pairs of results on the Chidi-Divya, Bob-Chidi and Alice-Diya spacelike slices are assigned probability zero. Then if it is assumed that observed events are absolute in the sense that for each run of this experiment there exist exactly four non-relativized outcomes A, B, C, D, we can infer that a certain pair of results on the Alice-Bob spacelike slice must also be impossible - but unitary quantum mechanics predicts that this pair of results *is* possible, so if unitary quantum mechanics is universal this pair of results would presumably be seen if the experiment were repeated enough times. Thus Ormrod and Barrett conclude that if unitary quantum mechanics is universal in the sense that we can always choose to collapse the wavefunction on any arbitrary spacelike slice, then observed events cannot be absolute. The theorem of Ormrod and Barrett requires neither a Locality assumption nor a No-Superdeterminism (and/or retrocausality) assumption.

This does not exhaust the space of non-absoluteness theorems in the literature - we will not be able to cover them all in this paper, but ref [10] provides a helpful overview of some other non-absoluteness theorems, including explanations of how they relate to the theorems discussed here. In our view, none of these other theorems convincingly support the possibility of metaphysically radical non-absoluteness either, for reasons similar to the ones we will shortly set out with regard to the Bong et al, Lawrence et al and Ormrod and Barrett theorems.

2.1 Universality

Now, we emphasize that none of the experiments referenced in the non-absoluteness theorems have yet been performed, and there is no prospect that they will be performed in the immediate future. So it is not possible to use these theorems to argue directly for the non-absoluteness of observed events on the basis of real empirical results. Rather it is necessary to begin by making some assumption about what the results of these experiments will be - and as noted above, typically those who wish to use the theorems to argue for non-absoluteness proceed by assuming 'the universality of unitary quantum mechanics,' so it can be assumed the results of these experiments will be as predicted by unitary quantum mechanics.[5] However, there are different ways in which quantum mechanics could be 'universal,' and therefore in this paper we will need to keep in mind two different conceptions of universality.

The first is 'first-person' universality: to say that unitary quantum mechanics is 'universal' in this sense is to say that it is always a correct description of all the observations made by any individual observer, but it does not describe the relation between observations made by different observers. Here we should allow that the observations made by an individual observer, Alice, may include reports made to Alice by other observers about the outcomes of their own experiments, or records that she consults about observations made by other observers, which may be modelled quantum-mechanically as interactions between Alice and other observers/records, or as measurements by Alice on other observers/records. Alice has no direct access to observations made by others, so her only information about them comes from the results of her interactions with other observers/records, and thus first-person universality requires only that Alice's own measurement outcomes are correlated with the results of Alice's measurements on observers/records in the way that quantum mechanics dictates - first-person universality does not require that Alice's outcomes are correlated with the *actual* observations made by other observers in any particular way. Therefore first-person universality places no constraints on the relations between observations made by two distinct observers, except in scenarios where there is some subsequent interaction or measurement in which the observers are able to share information about their outcomes.

First-person universality is the only kind of universality needed in order to use the theorems of Bong et al and Lawrence et al to argue for non-absoluteness, since in these scenarios the sets of variables used to prove the theorems are all in principle available to a single individual. First-person universality is also the kind of universality which is relevant for relational and perspectival interpretations of quantum mechanics[12, 13], in which quantum mechanics is typically understood as a

[5] We note that a more recent version of the Bong et al theorem, in ref [11], assumes only that universal quantum computing is valid; and more generally, it is clear that to make these theorems work one need only assume that the predictions of unitary quantum mechanics are correct for one particular experiment, not that they are always correct. However, presumably 'the universality of unitary quantum mechanics' implies at least that its predictions are correct for this particular experiment, so the reasons we adduce in this section for believing in the universality of unitary quantum mechanics are automatically also reasons for believing any such weaker assumption.

single-user, first-person theory designed to characterise observations from the point of view of the individual observer who is currently applying it. Indeed, some relational and perspectival approaches claim that it is not even meaningful to compare measurement outcomes obtained by different observers, so we can't even pose a question about whether a collection of outcomes obtained by different observers are jointly consistent with the predictions of quantum mechanics unless the observers share their results by means of some kind of physical interaction.

However, one might also think that unitary quantum mechanics should be universal in a stronger sense. For example, the theorem of Ormrod and Barrett uses a stronger notion of universality; motivated by relavistic considerations, Ormrod and Barrett argue that no spacelike slices should be privileged and thus the 'universality of unitary quantum mechanics' should entail that we can always make correct predictions by choosing an arbitrary spacelike slice and then applying unitary evolution up until a collapse of the wavefunction on that spacelike slice, with predictions on that slice obtained from the Born rule. They refer to this prescription as 'Frame-Independent Quantum Theory' (FIQT). Evidently universality in the FIQT sense is a stronger requirement than the first-person approach to universality, since it involves applying quantum mechanics directly across observations made by different observers on the same spacelike slice, even though some of these observers do not and could not possibly share their outcomes with one another - the first-person approach would not recognise as legitimate such applications of the theory to sets of outcomes which could never all be available to a single observer. Although universality in the FIQT sense may seem very reasonable, no one could ever directly verify that this kind of universality really holds in the actual world, because ultimately the only thing anyone can ever *directly* verify is the fact that quantum mechanics makes the right predictions for a single observer. And therefore, although the theorem of Ormrod and Barrett does not assume No-Superdeterminism or Locality and is thus in some sense stronger than the Bong et al and Lawrence et al theorems, in another sense it is weaker because it assumes a more demanding version of 'the universality of unitary quantum mechanics.'

2.2 Responses

There is one obvious route which allows us to accommodate the theorems introduced in section 2 without accepting any kind of non-absoluteness - we need only adopt any interpretation of quantum mechanics which says unitary quantum mechanics applies only in certain regimes and is therefore not universal in either of the two senses discussed above. For example, spontaneous collapse interpretations[14, 15] and interpretations where the wavefunction is collapsed by consciousness[16] do not predict that the experiments referenced in the non-absoluteness theorems would have the results described in section 2 which are used to derive contradictions, and therefore proponents of spontaneous collapse and consciousness-based collapse interpretations clearly do not need to worry about non-absoluteness.

However, there are good scientific reasons for being wary of approaches in which unitary quantum mechanics is not universal in at least one of these senses. For a start, we have not yet found any direct evidence that quantum mechanics ceases to apply in some regimes, so postulating some kind of regime change like von Neumann cut involves adding a considerable amount of additional structure without any clear empirical basis. Additionally, Wallace[17] points out that no extant spontaneous collapse interpretation can reproduce all the predictions of quantum field theory, and he also adduces various structural reasons to think that other approaches involving something like a von Neumann cut are likely to face similar difficulties, so as things stand it seems possible that only approaches which uphold the universality of unitary quantum mechanics will ultimately be empirically adequate. Thus it is certainly tempting to respond to the non-absoluteness theorems by simply accepting that indeed, observed events are not absolute. As a physicist one may feel this is a small cost to pay if the alternative requires us to deny the universality of unitary quantum mechanics and then do all the hard work of constructing what would essentially be a whole new scientific theory incorporating some additional kind of dynamics.

However, accepting the non-absoluteness of observed events is not as harmless as it may seem: in fact this move has very serious consequences for the epistemology of science. For example, one way of denying AOE is to adopt an Everettian approach in which all of the possible outcomes for a given measurement do actually occur in different branches of the wavefunction - but as argued by ref [18], it is difficult in this kind of picture to give an account of probability adequate to make sense of probabilistic confirmation. Another way of denying AOE is to say that different observers may fail to agree about the outcome of a given measurement - yet as argued by ref [19], this may lead to a picture in which observers cannot ever share any information about measurement outcomes, meaning that it is impossible to use observations made by other observers or even past versions of oneself for the purpose of empirical confirmation. So accepting the non-absoluteness of observed events should not be done lightly - this move risks endangering our scientific methodology to the extent that it may no longer be possible to regard quantum mechanics as empirically confirmed, and we certainly cannot rationally believe an interpretation of quantum mechanics which denies that quantum mechanics has been empirically confirmed.

However, the terms ‘absolute’ and ‘relative’ are somewhat vague and abstract, and discussions of the non-absoluteness theorems often seem to be equivocating between different meanings of these words. So it may be that it is possible to come up with an approach which is ‘absolute’ enough to avoid the kinds of epistemic problems that we have just described, while still being ‘relational’ enough to be compatible with the universality of quantum mechanics. We will now seek to understand what such an approach might look like.

3 Non-absoluteness

Non-absoluteness theorems typically define ‘non-absoluteness’ by negation: that is, they provide a definition for ‘absoluteness of observed events’ and then ‘non-absoluteness of observed events’ is understood to refer to any scenario in which that definition fails to hold. For example, both Bong et al and Ormrod and Barrett specify that AOE involves the stipulation that '*an observed event is a real single event, not relative to anything or anyone*' and 'non-absoluteness of observed events' is simply the denial of this.

However, if we are to use these theorems to argue that the universality of unitary quantum mechanics implies the 'non-absoluteness of observed events' (conditional on some auxiliary assumptions), it is crucial to say something about what 'non-absoluteness' might actually look like. For if non-absoluteness is merely defined as the negation of some notion of 'absoluteness,' that leaves open the possibility that there just is not *any coherent way* in which events could fail to be absolute in this particular sense. And if so, the relevant non-absoluteness theorem has not proved the existence of non-absoluteness: rather, it has proved that, if we are not willing to let go of any of the auxiliary assumptions, then we must accept that quantum mechanics is not universal, i.e. the theorem has become a reductio ad absurdum argument against the universality of unitary quantum mechanics.

Now as matter of fact, we think there is at least one coherent way in which observed events could fail to be absolute: it could be possible for measurement events to have multiple outcomes per observer (MOPO), as in the Everett interpretation. Clearly in an Everettian world measurement outcomes are not absolute - there is no fact of the matter about which outcome of a measurement actually occurred, given that all of the possible outcomes do actually occur and are witnessed by different versions of the observer who performed the measurement. But in this paper our aim is to investigate possible alternatives to these multiple-outcome pictures, so we will henceforth rule out all approaches in which the observer who performs a given measurement can have conscious experiences of two or more distinct outcomes for that measurement.

What remains once we rule out the MOPO approaches? Well, it is commonly suggested in the literature that there is some *other* coherent way in which observed events could fail to be absolute. In particular, in 'relational' or 'perspectival' approaches[13, 20–30], typically no measurement has more than one outcome relative to any observer, but nonetheless there is no fact of the matter about which outcome of a measurement actually occurred, because a measurement can have different outcomes relative to different observers. These kinds of views are typically associated with quite strong metaphysical claims - for example, proponents of such views have suggested that there is no objective reality[3]; that all physical facts must be relativized to an observer[25]; that facts are subjective[3]; that it is not even meaningful to compare the perspectives of different observers[31]; and so on. We will henceforth refer to possibilities which deny that there is any absolute fact about the outcome of a measurement as 'metaphysically radical' non-absoluteness. (In this paper we will not include the Everett interpretation and other MOPO approaches in the category of metaphysically radical non-absoluteness).

Now, one may naturally wonder how it can be that a measurement can have different outcomes relative to different observers. After all, in each of these approaches it is accepted that there is a specific observer who actually performs the measurement, and since these are not MOPO approaches, the observer who performs the measurement has a conscious experience of exactly one outcome. So even if there is some reason for other observers to regard this measurement as having some different outcome, one might naturally think that the unique outcome of the measurement *as it is experienced by the observer who actually performed that measurement* should be regarded as a unique, 'absolute' fact about the outcome of the measurement. For if Chidi is the person who actually performed the measurement, isn't he the ultimate authority on his own experience? If Alice fails to agree with him about that outcome, why should we say that there is a different measurement outcome relative to her? Isn't she just *wrong*?

Thus we think it is not particularly helpful to simply gesture at the possibility that observers may fail to agree about measurement outcomes; it is important to be clear about what that actually means. And in fact, an examination of the non-absoluteness theorems discussed in section 2 reveals at least three different ways in which one observer may fail to agree with another about the outcome of a measurement. Note that, since it is arguably the case that standard quantum mechanics does not offer unequivocal answers about what will be witnessed by the various different observers in a Wigner's Friend or Extended Wigner's Friend scenario, we will here make use of the concept of an 'extension' of quantum mechanics, by which we mean simply some probabilistic or deterministic specification of what all of the observers will see in these scenarios. Since we are imagining that all MOPO approaches have been ruled out, we will require that an extension of quantum mechanics specifies no more than one outcome per observer for each measurement. Thus we can specify our three types of 'disaccord'[6]:

1. Type-I Disaccord: Suppose that Divya performs some measurement D_1 and then Alice performs some measurement A_1 and subsequently applies quantum mechanics to make some inference about the result of Divya's measurement D_1. We will say that an extension of quantum mechanics exhibits Type-I disaccord if it tells us that Alice's inference may not agree with what Divya herself experienced the outcome of measurement D_1 to be, even if Alice applied quantum mechanics correctly.
2. Type-II Disaccord: Suppose there is some quantum state ψ which gives correct predictions for the relative frequencies over a large number of trials in all the measurements that Alice could possibly perform on Chidi at some time t; and suppose this quantum state ψ is naturally interpreted as representing something about Chidi's experiences at or shortly before the time t. We will say that an extension of quantum mechanics exhibits Type-II disaccord if it tells us that the 'natural' representation of Chidi's experiences suggested by the state ψ may not agree with what Chidi is actually experiencing at or shortly before the time t.

[6] Note that for the sake of conciseness, in these definitions we use the phrase 'does not agree' to mean either that two things can be compared and they are different, or that they just cannot be compared at all. This is why we have used the term 'disaccord' rather than 'disagreement' - 'does not agree' is not necessarily synonymous with 'disagree' here, since it also includes cases where there is neither agreement nor disagreement.

3. Type-III Disaccord: Suppose Chidi performs some measurement C_1, and then Alice performs a measurement A_2 on Chidi himself, with the aim of establishing what his measurement result was - for example, this could involve simply asking him what the result was. We will say that an extension of quantum mechanics exhibits Type-III disaccord if it predicts that even under the most favourable conditions possible, the result of Alice's measurement A_2 may not agree with what Chidi himself experienced the result of the measurement C_1 to be.

Note that one might think that Type-III disaccord is simply a special case of Type-I disaccord in which the measurement A_1 is performed on Divya herself. However, we emphasize that Type-I disaccord involves performing a measurement and then using *quantum mechanics* to make an inference about an outcome obtained by another observer, whereas Type-III disaccord involves performing a measurement on an observer and then using *a common-sense understanding of this result* to make an inference about an outcome obtained by another observer. For example, the measurement A_2 could simply involve Alice asking Chidi what outcome he obtained, and the common-sense inference here is that whatever Alice hears Chidi saying matches the outcome that he actually obtained. As we will shortly see, this common-sense inference about the relation between outcomes obtained by different observers doesn't follow from quantum mechanics alone, so Type-III disaccord is distinct from Type-I disaccord, since the inferences involved have different justifications.

We will defer further discussion of Type-I disaccord to section 7. For now, we will proceed as follows: first we will demonstrate that the Wigner's friend scenario and the Bong et al and Lawrence et al theorems can be understood as demonstrating that any extension of quantum mechanics with certain properties (i.e. upholding the universality of unitary quantum mechanics, and No-Superdeterinism, and Locality,) must exhibit Type-II and/or Type-III disaccord. We will then argue that, even if the Everett picture and other MOPO approaches are off the table, the existence of Type-II or Type-III disaccord does not entail the existence of some kind of metaphysically radical non-absoluteness; there is always the option to adopt a deflationary interpretation of disaccord which may have a relational flavour but which maintains that reality is made up out of objective, 'absolute' facts. Thus we will argue that these theorems do not in and of themselves provide any compelling scientific reason to believe in metaphysically radical non-absoluteness; that is a further interpretational choice which is not mandated by the scientific facts as they are currently understood.

3.1 Type-II Disaccord

The original Wigner's friend paradox demonstrates that any extension of quantum mechanics which is consistent with the universality of quantum mechanics in the first-person sense, but which also maintains that every observer always sees a single definite outcome to any measurement they perform, must exhibit Type-II disaccord. This follows from the fact that the universality of quantum mechanics in the first-person sense entails that Alice's measurements on Chidi in the Wigner's Friend scenario must exhibit the statistics we would expect to see if Chidi were in the state ψ.

Of course, in and of itself this statement about Alice's measurement results is nothing more than a description of the dynamics of the Alice-Chidi interaction, which does not entail any kind of disaccord: if everyone in this scenario knows that unitary quantum mechanics is universal in the first-person sense, then everyone can agree that indeed, ψ is a correct description of the dynamics of Chidi and his system with respect to all of the measurements that Alice is able to make. But we will potentially get disaccord if Alice begins making additional interpretative assumptions about the meaning of the state ψ; for the state ψ appears to have the form of a superposition of different conscious states, so if Alice observes statistics consistent with Chidi being in the state ψ, the natural assumption for her to make is that he has not in fact experienced any definite outcome. And if she does assume this then she will end up making an inference about Chidi's experiences which does not agree with what Chidi himself has actually experienced, since ex hypothesi he has in fact seen a single definite outcome - thus we end up with Type-II disaccord.

Now, as we have just described this scenario, there does not seem to be any metaphysically radical non-absoluteness involved - Alice is simply making an incorrect interpretative assumption which leads her to hold some incorrect beliefs about what Chidi has experienced. But some proponents of relational and perspectival interpretations would like to make a stronger claim: they argue that Chidi *really* has not seen any definite outcome relative to Alice, although of course he has seen a definite outcome from his own point of view, so there is just no fact of the matter about whether or not Chidi has seen a definite outcome.

To analyse this claim, it will be helpful to be more specific about the meaning of the word 'state.' For this term is often used in a way which equivocates between two possible meanings: a) an intrinsic description of a system at a given time or over some relatively short time interval, and b) a mathematical object encoding predictions for the outcomes of measurements performed on a system at some time or over a relatively short time interval. For example, in interpretations of quantum mechanics which deny the existence of hidden variables, quantum states are often regarded as being both an ontological representation of a system at a time, and also a tool which encodes information about the results of possible measurements on that system at that time. In this paper we will refer to a) as the intrinsic condition, and b) as the interaction state, because in a theory which incorporates measurements into its physical description rather than treating them as exogenous, a mathematical object encoding information about outcomes of measurements on a system must ultimately be thought of as providing a description of the dynamics for the interaction of the system with some other system which can be understood as 'measuring' it.

Now, this term 'intrinsic condition' is intentionally vague - it is supposed to refer to some objective, observer-independent description of the relevant system, but we are trying to avoid assumptions about the nature of that description. It might perhaps be like a classical state, but might also be something quite different - for example, in the Bell flash approach[32], in which reality is composed of a distribution of pointlike events across spacetime, the 'intrinsic condition' of a system at a time would simply be the distribution of flashes across some spacetime region which is roughly occupied by the system. We will, however, stipulate that for a system which is a conscious observer, if this observer has anything which can be described as an intrinsic condition, then her intrinsic condition is what determines her conscious experiences and observations.[7] This seems like a reasonable stipulation at least for those of a physicalist persuasion, since the nature of the 'intrinsic condition' can be made quite general to accommodate various different theories of the physical basis of consciousness.

We can be more specific about the term 'interaction state' - we will say that the interaction state of a system, S, relative to some observer, O, in a given prediction context, P, is an encoding of the optimal and complete predictions for the results of any measurement that O might perform on S conditional on all of the information specified in E. By 'prediction context' we mean simply a set of information about the physical situation which is to be used in making the predictions encoded in the interaction state - so for example, the prediction context specifies all the physical features that must be reproduced if we wish to perform repeated experiments which will count as instances of the same experiment for the purpose of checking whether observed relative frequencies match the predictions. Note that the interaction state is relativized *both* to an prediction context *and* possibly to different observers within that prediction context. The relativization of the interaction state to a given prediction context is epistemic in nature, because simply including different information in the prediction context can give rise to different predictions - for example, if one specifies an prediction context which includes information about the future, this will of course allow us to predict with certainty some outcomes which we would not be able to predict with certainty if we only had information about the past. However, the relativization of the interaction state to observers is *not* merely epistemic - to say that a system may have two different interaction states relative to two different observers, in a given prediction context, is to say that the outcomes predicted for measurements performed by one observer are different from the outcomes predicted for measurements performed by some other observer, even though both are predicted using exactly the same information from the same prediction context, and the observers in question may be in possession of exactly the same information. In particular, the fact that a state has different interaction states relative to different observers does not necessarily mean that one or the other of them lacks some information about the system; the difference in the interaction state reflects real differences in the results that they will obtain if they measure the system.

Thus our starting point will be that for a conscious observer, if she has an intrinsic condition then her intrinsic condition determines her experiences, whereas her interaction state simply predicts the outcomes of measurements performed on her. But because it is common to refer to both intrinsic conditions and interaction states as simply 'states,' it is also common to assume that there is a straightforward mapping between the two such such that the conscious experiences of an observer can simply be read off her *interaction* state. Thus for example, in section 2 we defined the set of states $\{|C_i\rangle\}$ by saying that '$|C_i\rangle$ *is the state of Chidi corresponding to him having*

[7] We don't know what would determine the conscious experiences and observations of an observer in a radically perspectival picture in which there cannot be any intrinsic conditions; we leave that for the proponents of radically perspectival pictures to specify.

seen the outcome $|S_i\rangle$.' Here the states $|C_i\rangle$ are being used as descriptions of Chidi's intrinsic condition, but we also stipulated that the state $\psi = \sum_i c_i |C_i\rangle|S_i\rangle$ describes the measurement outcomes that will be obtained when Alice measures Chidi, and thus the states $|C_i\rangle$ are also being used to construct an interaction state. This conflation of the intrinsic condition with the interaction state is entirely standard practice within the field of quantum foundations and beyond - and yet it must be kept in mind that this standard practice is based on a number of interpretative assumptions, as we are not logically compelled to assume any particular link between the intrinsic condition and the interaction state. There may be good reasons for assuming the existence of some such link, of course - we will discuss some such reasons in section 4 - but nonetheless it is a substantive assumption which should not be taken for granted.

Returning to the Wigner's friend scenario, then, the fact that 'the quantum state ψ gives correct predictions for the relative frequencies over a large number of trials in all the measurements that Alice could possibly perform on Chidi' can be expressed in the language we have just set out by saying that in the given prediction context, Chidi has interaction state ψ relative to Alice. If we then also assume that Chidi's interaction state relative to Alice must closely reflect his intrinsic condition, or that it must contain all the same information as the intrinsic condition, it is natural to conclude that Chidi's intrinsic condition is some kind of indefinite superposition, so at least relative to Alice he has not made any definite observation. But at the same time, it has been stipulated that from his own point of view Chidi has indeed made a definite observation, and this is what gets us to the supposed 'non-absoluteness of observed events' advocated by proponents of relational and perspectival approaches. For example, this is essentially the argument made by Cavalcanti in ref [33]: Cavalcanti considers the case in which the external observer is able to coherently reverse the observation made by the internal observer, so all the observations made by the external observer are consistent with the internal observer being in the quantum state ψ (which in the language we have used here amounts to assuming that the internal observer has interaction state ψ relative to the external observer in this prediction context), and then argues that we ought to assume that once a quantum state has been specified, then '*there are no objective facts to further specify your degrees of belief even in principle, at least in the cases where you assign a pure state to a situation*' (which amounts to assuming that the intrinsic condition cannot include any information not already included in the interaction state ψ), and hence concludes that '*there's no objective fact of the matter relative to (the external observer) as to which outcome (the internal observer) observed.*'

However, the supposed inevitability of this non-absoluteness rests on equivocation between intrinsic conditions and interaction states. For as noted above, we are not logically compelled to assume any particular relation between intrinsic conditions and interaction states, and therefore it is open to us to say that Chidi has a well-defined, observer-independent intrinsic condition according to which he has seen a definite outcome, but the intrinsic condition of Chidi does not map to the dynamics of the Alice-Chidi interaction in the way we would naturally expect. This would allow us to maintain that despite the fact that Chidi has seen a definite outcome $|S_i\rangle$, his interaction state relative to Alice does not take the form $|C_i\rangle$ in this prediction context. Moreover, this is not merely a matter of Alice's lack of information about

Chidi's outcome - even if we consider an prediction context which includes the information that Chidi has just seen the outcome $|S_i\rangle$, the best possible predictions for Alice's measurements on Chidi will still specify that if she performs tomographic measurements on Chidi and S she will see the results predicted by the state ψ, rather than $|C_i\rangle|S_i\rangle$. So an approach like this allows us to maintain first-person universality in the Wigner's Friend scenario without denying that Chidi has a definite experience of his measurement outcome. But evidently there is nothing non-absolute about any of this, and nothing prevents us from giving an objective, third-person description of such a scenario - the description will simply specify some definite observed measurement outcome for Chidi and then also note that the dynamics of the Alice-Chidi interaction are as described by the interaction state ψ, despite the existence of the definite outcome.

Of course, although the postulation of non-absoluteness is not compulsory here, one may still voluntarily choose to make interpretative assumptions about the existence of non-absoluteness. For example, one may postulate that the *reason* the Alice-Chidi interaction is described by the state ψ is because the intrinsic condition of Chidi, relative to Alice, is indefinite, even though from his own point of view he is in a single definite state. Or alternatively one may postulate that the reason the Alice-Chidi interaction is described by the state ψ is because there is no such thing as an intrinsic condition, so there is nothing to be said about this scenario beyond the specification of the interaction state ψ. But it must be emphasized that these kinds of assumptions are not necessary to resolve the Wigners's friend paradox, as we have already resolved the paradox simply by postulating that Chidi can have interaction state ψ relative to Alice even though Chidi himself has seen a definite outcome. There is no need to say anything at all about Chidi's experiences or outcomes 'relative to Alice' - perhaps the motivation for this is the idea that these somewhat unusual dynamics can't be explained without some kind of metaphysically radical relativization, but we do always have the option of simply regarding all of this as a brute fact about the dynamics of the theory, or perhaps appealing to some axiomatization of quantum mechanics to explain why the dynamics are the way they are.

Moreover, it's not obvious what meaning claims about Chidi being in an indefinite intrinsic condition 'relative to Alice' are supposed to have. Recall that here we are focusing on relational views which are not MOPO approaches, so by stipulation Chidi can have only one conscious experience of each measurement outcome, which would seem to suggest that he has only one intrinsic condition which cannot really be relativized to anyone.[8] If there are 'copies' of Chidi in some other kind of condition in versions of reality defined relative to other observers, and those copies are in some sense physically real, why are those copies not also having conscious

[8] Of course, it is possible that there exists some coherent relational view in which Chidi has one set of conscious experiences relative to himself and different conscious experiences relative to other observers. However, this would be a MOPO approach, putting it outside of the scope of the present discussion. Moreover, it seems likely that a view in which a given observer has different conscious experiences relative to different observers would end up looking quite similar to the Everett interpretation, with 'observers' playing the role of branches, so it's unclear that this would lead to any genuine alternative to the many-worlds picture.

experiences of the different measurement outcomes occurring in those versions of reality? Similarly, it's not straightforward to make sense of the claim that reality is composed entirely of interaction states without any intrinsic conditions at all, for interaction states are nothing more than sets of possibilities for future interactions, and it's certainly controversial to suggest that there could be a reality containing nothing other than possibilities. And it must be emphasized that all the relevant empirical content in this situation is already contained in the assertion that Chidi has interaction state ψ relative to Alice - further statements about what Chidi has or has not seen relative to Alice tell us nothing additional about what either Chidi or Alice will experience. So in a non-MOPO approach this statement about Chidi having a condition relative to Alice which is different from his actual experiences seems to float free from reality, describing nothing but experiences that nobody can possibly have, and thus in addition to being unnecessary it's arguably incoherent.

Possibly the claim about Chidi's measurement outcome 'relative to Alice' should really be understood as asserting that the inference Alice makes about Chidi's outcome is not *just* a mistake in the ordinary sense. To make this point it is helpful to distinguish between two different kinds of mistake Alice could hypothetically make. First, she could think that the interaction state of Chidi and S relative to her in this prediction context is ψ when really it is a state of the form $|C_i\rangle \otimes |S_i\rangle$ - and if she makes this mistake, some of her predictions for the results of measurements on Chidi and S in this prediction context will turn out to be wrong, so such mistakes have real empirical consequences. Clearly this is quite different from the kind of mistake where Alice knows the right interaction state for Chidi relative to her is ψ and then takes this to mean that Chidi has not had a definite experience when in fact he has, for these incorrect beliefs about what Chidi has experienced will never cause her to make predictions which could be empirically falsified. So the statement that Chidi has no definite measurement outcome 'relative to Alice' could be understood as a kind of shorthand intended to express that the fact that the interaction state of Chidi relative to her in this prediction context is ψ, and thus in a sense she is not making a mistake when she says to herself that he has no definite measurement outcome, even though he has in fact experienced an outcome. If this is what is meant by the terminology, we have no particular objection - we would simply caution that this way of using the phrase 'relative to Alice' means that the relativization of states does not entail the existence of any kind of metaphysically radical non-absoluteness, since the whole story is entirely compatible with the idea that there is an objective, mind-independent fact about Chidi's outcome which simply is not reflected in his interaction state relative to Alice for this prediction context in the way one might most naturally expect.

3.2 Type-III

The Bong et al experiment can be used to demonstrate that any extension of quantum mechanics which is consistent with the universality of unitary quantum mechanics in the first-person sense, which maintains that every observer always sees a single definite outcome to any measurement they perform, and which obeys the assumptions of Locality and No-Superdeterminism, must exhibit Type-III disaccord. The Lawrence et al theorem demonstrates roughly the same thing. We will focus on the Bong et al theorem in this section, but we think most of our conclusions would carry over to the Lawrence et al theorem.

For our purposes, the important thing to notice about the Bong et al theorem is that the AOE assumption is really made up of two quite separate assumptions.[9] The first assumption, which we will henceforth refer to as AOE1, is that there is '*a well-defined value for the outcome observed by each observer*' [7]. There is some ambiguity around what exactly is meant by this, but in this paper we will understand it to mean that the observer who performs a given measurement has a conscious experience of exactly one outcome for that measurement, i.e. measurements have only one outcome relative to the observer who performs the measurement. Understood in this way, it is not possible to deny AOE1 without adopting a MOPO approach. The second assumption, which we will henceforth refer to as AOE2, is that if one observer performs a measurement on another aiming to establish the outcome of their measurement, the value obtained by the first observer will be the same as the value that the second observer actually witnessed.[10]

To see that these two assumptions together can be used to play the role of 'absoluteness of observed events,' in the Bong et al theorem, note that from AOE1 and AOE2 we can derive the conclusion that in the reelvant experiment there must exist a well-defined probability distribution over the four outcomes A, B, C, D which reduces to the probabilities predicted by quantum mechanics for the pairs $\{AB, AD, CB, CD\}$, which is exactly the result needed by Bong et al to arrive at their contradiction. To see where this conclusion comes from, note first that AOE1 allows us to conclude that on any run of the experiment there are exactly four measurement outcomes witnessed by the people actually performing the measurements. Now, AOE1 on its own does not imply that these four outcomes coexist in a single reality or can be compared in any way; however, this does follow once we add AOE2, because AOE2 implies that Alice can access the single measurement outcome observed by Chidi, and Bob can access the single measurement outcome observed by Divya, and Alice and Bob can also access one another's measurement outcomes and also any information that the two of them might have about Chidi and Divya's outcomes; so either Alice or Bob can in principle access any pair of values in the set $\{AB, AD, CB, CD\}$. Thus AOE1 and AOE2 together imply that there are exactly four

[9] In the Lawrence et al theorem the argument is less direct: what we have called AOE1 is understood to follow from the 'existence of relative facts,' and the assumption that we have called AOE2 is never referenced explicitly, but seems to be assumed in the authors' discussion of the version of relational quantum mechanics with cross-perspective links. In any case, as we saw in section 2, if the Lawrence et al theorem is to be repurposed as a 'non-absoluteness' theorem in the tradition of the Bong et al result, it does require both assumptions AOE1 and AOE2, since otherwise several of the trios of outcomes to which the quantum constraints are applied could never be accessible to any individual observer, and therefore the universality of quantum mechanics in the first-person sense would not imply anything about them.

[10] Ref [10] also splits the AOE assumption used by Bong et al into two parts, the first being an assumption that 'an observed event is an absolute single event, and not relative to anything or anyone,' and the second, which they call 'Tracking' equivalent to what we have referred to as AOE2.

measurement outcomes in this situation *and* that the four outcomes must coexist in a single reality, in the sense that any pair of values in the set $\{AB, AD, CB, CD\}$ can potentially belong to Alice's or Bob's reality. We can then apply the first-person universality of quantum mechanics to place constraints on these values. We emphasize the importance of AOE2 here - if the observers were not able to share their original measurement outcomes with each other, then nobody in this scenario would ever have access to more than one of the original measurement outcomes A, B, C, D, and therefore nobody would ever see anything incompatible with quantum mechanics even if the probability distribution over the values for the pairs $\{AB, AD, CB, CD\}$ were not as predicted by quantum mechanics. Subsequently other authors have considered relaxing the AOE2 assumption to arrive at inequalities applying to scenarios in which Alice can gain only incomplete information about Chidi's measurement outcome[34] and similarly for the other observers, but of course in these cases it is still assumed that *some* information can be obtained; no theorem of this kind can be proved without Alice having some sort of access to Chidi's observations and Bob's observations and to Divya's observations via Bob.

Since the pair of assumptions AOE1 and AOE2 are together sufficient to derive the same conclusion that Bong et al drive from what they call 'absoluteness of observed events' it follows that anyone who wishes to respond to the Bong et al theorem by rejecting AOE and maintaining the other assumptions must ultimately reject either AOE1 or AOE2. And since we are interested here in whether the non-absoluteness theorems can be used to argue for some kind of metaphysically radical non-absoluteness rather than simply a MOPO approach, we are interested in the case where AOE1 is retained and AOE2 is rejected; and evidently the rejection of AOE2 means that we have Type-III disaccord. Thus one way of interpreting the Bong et al theorem is to see it as an argument that if unitary quantum mechanics is universal in the first-person sense, and measurements always have a single outcome for each observer, and the Locality and No-Superdeterminism assumptions are both correct, then Type-III disaccord must exist.

We note that some proponents of non-absoluteness in the non-MOPO sense might object to this conclusion on the grounds that that they reject neither AOE1 nor AOE2; they simply reject the assumption that there are absolute facts about measurement outcomes, which I will refer to as AOE3. And indeed, it's true that AOE3 can substitute for AOE1 in the derivation of the conclusion needed for the Bong et al theorem: from the assumption that there are absolute facts about measurement outcomes, we can conclude that there's a well-defined probability distribution over the four outcomes A, B, C, D, and then we add AOE2 to conclude that any of the pairs $\{AB, AD, CB, CD\}$ could in principle be accessed by a single person, so we can apply the first-person universality of quantum mechanics to place constraints on these values. However, as described above, in order to derive AOE we do not need to first assume that there is a well-defined joint probability distribution and then place constraints on it using AOE2; we can also start from the assumption that each observer has a conscious experience of a unique outcome and then use AOE2 to insist that observers can sometimes access one another's unique outcomes, meaning that the pairs $\{AB, AD, CB, CD\}$ could in principle be accessed by a single person. And thus, even though we did not assume initially that the unique outcomes associated with the observers all co-exist in the same reality, nonetheless AOE2 ultimately entails that some of them ultimately *do* coexist in the same reality and therefore they must have joint probabilities which reduce to the probabilities predicted by quantum mechanics. Thus, since AOE can be derived either from AOE1+AOE2 or AOE3+AOE2, it is not possible to reject AOE by rejecting AOE3 while maintaining AOE1 and AOE2 - anyone who proposes rejecting 'absoluteness' as a response to the Bong et al theorem is obliged to deny either AOE1 or AOE2. Thus, whatever the proponents of non-MOPO relationalism mean when they say they are rejecting absolute facts about measurement outcomes, this must include rejecting either AOE1 or AOE2; and since rejecting AOE1 appears to lead directly to a MOPO approach, it would seem they must ultimately be committed to the rejection of AOE2 and hence committed to the existence of Type-III disaccord. Or at least, if this is not the intention, then the proponents of non-absoluteness need to offer more clarity on how they can reject AOE1 without ending up with a MOPO approach, or how they can reject the existence of a well-defined joint probability distribution over all four outcomes without rejecting either AOE1 or AOE2.

It should also be noted that the assumption AOE2 is also needed for a variety of other no-go theorems, most notably Bell's theorem - for if we can't assume that the two observers in a Bell experiment are able to compare their results, we certainly can't conclude anything about nonlocality. Thus one might wonder why we are placing so much emphasis on AOE2 in the context of the Bong et al and Lawrence et al theorems, given that it is also needed by many other theorems. And in fact, the reason we are focusing on AOE2 here is simply because denying AOE2 appears to be the route out of the Bong et al conundrum that must ultimately be taken by those who contend that the non-absoluteness theorems are compelling arguments for some kind of non-MOPO metaphysically radical relationalism or perspectivalism; thus, since our aim here is to understand whether the non-absoluteness theorems really offer any compelling argument for these metaphysically radical possibilities, it is important for us to consider what the failure of AOE2 would mean. It is not our intention to argue that denying AOE2 is a good or even plausible idea - indeed, we will shortly argue that it is neither - as our point in this section is simply to show that even if one *does* choose to respond to the Bong et al theorem by denying AOE2 and thus accepting Type-III disaccord, that doesn't inevitably lead to metaphysically radical relationalism or perspectivalism.

Now, Bong et al seem to be assuming, or perhaps defining, that an observed event is 'absolute' only if information about it can be accessed by observers other than the person who initially observed it; and indeed, this idea was subsequently formalised in ref [34], which defines a 'non-absoluteness coefficient' quantifying the extent to which the outcome of Alice's measurement on Chidi is correlated with the result that Chidi himself obtained. But the failure of 'absoluteness' in this sense does not entail that there is anything radically relational or perspectival going on - after all, most realists would presumably accept that there can be a well-defined, objective fact of the matter about what has been observed by some observer even if no other observer ever finds out about it, or could possibly find out about it. Moreover, the definition used in Bong et al arguably doesn't reflect what most people intuitively imagine 'absoluteness of observed events' to mean. Indeed, Bong et al themselves initially define AOE as the requirement '*an observed event is a real single event, not relative to anything or anyone*' and yet the denial of AOE2 and the existence of Type-III disaccord seems entirely compatible with this kind of absoluteness. For example, Ormrod and Barrett use the same preliminary statement of the meaning of AOE, but then cash it out as requiring that there exists a fact of the matter about the measurement outcome which could in principle be written down as part of a non-relativized list of outcomes; and Type-III disaccord is clearly not incompatible with this, as the fact that Alice cannot find out Chidi's outcome does not entail that the outcome could not in principle be written down in a non-relativized list of outcomes, although obviously *Alice* could not write it down.

Indeed, note that Type-III disaccord can, like Type-II disaccord, occur as a result of mismatches between intrinsic conditions and interaction states. For suppose that ψ is the interaction state of Chidi relative to Alice for the prediction context specified in the Bong et al theorem, which as discussed in section 3.1 means there is a mismatch between Chidi's interaction state and instrinsic condition. Now if we *additionally* assume that ψ would still be the interaction state of Chidi relative to Alice even if we included a specification of the outcome of Chidi's measurement in the prediction context, then we immediately get Type-III disaccord. For the state ψ does not contain any information about the definite outcome witnessed by Chidi, so if ψ would still be the interaction state of Chidi relative to Alice even when the prediction context includes Chidi's outcome, this means that Chidi's outcome is not dynamically relevant to Alice - and therefore Alice cannot reliably get information about Chidi's outcome via any dynamical interaction, even if that simply involves asking him about his outcome. So at a conceptual level, Type-III disaccord is not all that different to Type-II disaccord; the only real novelty here relative to the Wigner's Friend case is that we are now allowing the set of possible Alice-Chidi interactions to include an interaction which is naturally described as 'Alice asking Chidi the outcome of his measurement,' and we are insisting that the best possible predictions for the outcome of that interaction is always given by ψ, for any prediction context. From this point of view it is clear that the problem is not that Chidi's outcome is not 'absolute' - the difficulty is simply that the dynamics are defined in such a way that Alice is not able to access that absolute outcome through any physical interaction, even one as straightforward as having a conversation with Chidi. So distinguishing between the interaction state and the intrinsic condition gives us everything we need to make sense of the Bong et al result, without introducing any kind of metaphysically radical non-absoluteness.

Of course, perhaps those who advocate metaphysically radical non-absoluteness have in mind some kind of verificationism which has the consequence that if observers can't possibly compare their measurement outcomes, then it isn't meaningful to talk about the relation between their measurement outcomes at all, which would seem to justify the claim that there can't be any absolute description of the outcomes. But clearly this approach is not *mandated* by the existence of Type-III disaccord as manifested in the Bong et al experiment, so the theorem by itself provides little reason to take that route unless one already has verificationist sympathies. Thus, although we do think that Type-III disaccord involves a more significant revision of our intuitive picture of the world than just Type-II or Type-I, nonetheless we are still not convinced that the existence of Type-III disaccord provides any compelling reason to postulate the existence of metaphysically radical non-absoluteness.

4 Intersubjectivity

The preceding discussion suggests that neither the Wigner's Friend scenario nor the Extended Wigner's Friend scenario point directly to metaphysically radical non-absoluteness - they simply point to some kind of disaccord, which we may or may not choose to interpret in a metaphysically radical way. So can we simply respond to these scenarios by just accepting the existence of Type-II and/or Type-III disaccord, interpreted in our preferred fashion?

Unfortunately it is not quite so easy, because there is a potentially serious problem for approaches allowing disaccord - they may lead to a kind of extreme failure of intersubjectivity. Both Type-II and Type-III disaccord involve some kind of mismatch between intrinsic conditions and interaction states; but we can only find out about other people's intrinsic states via interaction with them via their interaction states, and so if there were *no* relation at all between intrinsic states and interaction states, this would lead to a picture of reality in which each one of us is trapped inside of our own separate reality, with no ability whatsoever to learn anything about what is going on inside someone else's reality. This is exactly the problem envisaged in ref [19], where it is argued that such an extreme failure of intersubjectivity would undermine the entire practice of science. Thus if we are going to respond to Wigner's Friend scenarios by allowing some kind of disaccord, we probably need to maintain that in general there exists some systematic relation between the intrinsic condition of a system and its interaction states relative to other observers in various prediction contexts.

Still, that leaves a lot of freedom to fix the exact nature of this systematic relation. In particular, for the purposes of upholding the rationality of science it doesn't really matter whether or not 'supermeasurements' yield outcomes which map in an intuitively natural way onto intrisic conditions, given that no supermeasurement has ever yet been performed on or by any human being. What matters for scientific rationality is simply that ordinary kinds of physical interactions - the ones that real people have historically used and continue to use today to establish a shared intersubjective reality - should in general deliver meaningful information about the intrinsic condition of other observers. So it is no problem if Alice's supermeasurements on Chidi give results consistent with him being in state ψ, as long as her outcome matches his in the case where she simply asks him about his result. That is to say, we can straightforwardly accommodate Type-I and Type-II disaccord in a scientific worldview, as long as we don't also have *Type-III* disaccord - or at least, we don't have Type-III disaccord in ordinary physical interactions like conversations between observers, for it is only that kind of disaccord that would seriously undermine the reliability of the means of communication that we use to establish our shared intersubjective reality.

What might an approach allowing Type-II but not Type-III disaccord look like? We will see some examples in section 6, but for now let us just recall that as noted in section that in cases like the Bong et al experiment, we automatically get both Type-II disaccord and Type-III disaccord, *if* we specify that the interaction state ψ would still be the interaction state of Chidi relative to Alice even if we included the information about Chidi's definite outcome in the measurement context. However, we do not necessarily need to make this specification. We do, of course, have to say that when we add more information to the prediction context we get a new interaction state which is compatible with the statistics predicted by ψ for the original prediction context, but this can easily be done - after all, most of the predictions encoded in the interaction state ψ are probabilistic, so it's entirely possible that including additional information will lead to more specific predictions for some measurements. Indeed, to get Type-II disaccord without Type-III disaccord in the Bong et al case, all we need to do is specify that if we add to the prediction context information about the definite outcome S_i witnessed by Chidi, then the interaction state of Chidi relative to Alice is still ψ *except* that we can make a more specific prediction for the case where Alice performs a measurement in a basis corresponding to the variable of Chidi that records his measurement outcome - in that case Alice will definitely get an outcome $|C_i\rangle\langle C_i|$ matching Chidi's definite outcome S_i, but for any other basis the most specific prediction possible is still the probabilistic prediction encoded in the state ψ. This means that Chidi's interaction state relative to Alice in the prediction context where we have the extra information about S_i is not exactly the same as ψ, but it is also not the eigenstate $|C_i\rangle$ associated with the definite outcome S_i - for although we can predict that Alice will definitely get the outcome $|C_i\rangle$ if she measures in a basis for which $|C_i\rangle$ is a possible outcome, for *any other basis* she will instead see outcomes as predicted by ψ, not $|C_i\rangle$. This means that adding an exception for measurements in this special basis does not entail that the state ψ is just a reflection of Alice's lack of knowledge of the true state $|C_i\rangle$, for even in the prediction context including information about Chidi's outcome it is still true that the interaction state is ψ-with-an-exception rather than $|C_i\rangle$. So when we add this exception we still have genuine Type-II disaccord, but we have eliminated Type-III disaccord, thus demonstrating that it's possible to have one without the other.

That said, getting rid of Type-III disaccord comes at a cost. For we saw in section 3.2 that the Bong et al theorem can be used to argue that if quantum mechanics is universal in the first-person sense, and we are not willing to accept a MOPO approach, and we are determined to maintain Locality and No-Superdeterminism, then we *are* in fact obliged to have Type-III disaccord. So although it's indeed possible to have Type-II disaccord without Type-III disaccord, the Bong et al theorem tells us that this can't be achieved in a way which is compatible with the first-person universality of quantum mechanics unless we reject either Locality or No-Superdeterminism. Thus it looks like accepting some type of disaccord is not going to be enough to solve the problem posed by the Bong et al theorem - since we can't reasonably accept Type-III disaccord, regardless of whether it is interpreted in a metaphysically radical way or not, we're going to have to reject one of the other assumptions going into the theorem.

In summary, if we rule out MOPO approaches and we want to maintain first-person universality, the original Wigner's Friend scenario gives us good reason to allow certain kinds of mismatch between intrinsic conditions and interaction states - but we must stop short of allowing Type-III disaccord, and therefore it seems that under these circumstances we are ultimately going to have to deny either the Locality assumption or the No-Superdeterminism assumption.

4.1 Retrocausality

So if these are the options available, is it better to reject Locality or No-Superdeterminism? We will not make a judgement on that point here, but we do want to make the point that the failure of No-Superdeterminism in this particular context may not be as problematic as it first appears. To see this, first recall that the No-Superdeterminism assumption is written in such a way as to rule out both superdeterminism and retrocausality. We will now argue that the type of 'retrocausality' that would be needed to avoid Type-III disaccord in the Bong et al experiment is of a very special kind, and some of the usual reasons one might have for wishing to rule out retrocausality do not apply in this context. Indeed, we think the fact that the 'retrocausality' needed to avoid Type-III disaccord in the Bong et al experiment happens to be of this very special kind may not be a coincidence - rather it should be taken as an indication that 'retrocausality' could be precisely the right solution to the problem posed by these theorems!

4.1.1 Grandfather Paradoxes

One of the main reasons for being wary of retrocausality is the fact that it can in principle be used to create logical contradictions. For example, it is easy to construct a retrocausal version of the grandfather paradox: suppose that instead of time traveling, our intrepid experimenter makes use of a retrocausal mechanism to cause her grandfather to die before he can father any children. But then the time traveler will not exist, so she won't be able to cause her grandfather's death. Thus if the retrocausal mechanism is deterministic and perfectly reliable, there is no logically consistent way to resolve this set of events; so evidently this kind of retrocausal mechanism cannot exist if the world cannot contain logical contradictions.

Of course, many of the processes we are concerned with in quantum mechanics appear to be indeterministic, so this kind of argument may not directly apply. But we can easily imagine an indeterministic analogue, in which a set of probabilistic processes are composed in a way which makes it impossible for them all to exhibit relative frequencies exactly equal to or close to the values of their theoretical probabilities; we will refer to this as a 'probabilistic contradiction.' For example, suppose we gather a collection of time travelers and have them make use of a retrocausal mechanism which is supposed to have fifty percent chance of causing the death of the person on whom it is directed at a chosen time. If all of the time travelers direct this mechanism on their grandfathers at a time before these grandfathers have fathered any children, we would naively expect around fifty percent of the grandfathers to die. But in fact, it is not logically possible for any of the grandfathers to die, because any grandfather who dies before he fathers any children cannot have a grandchild who uses this retrocausal mechanism on him. So no matter how many times we try this experiment, the relative frequency of death will always be zero, which is very far from the expected frequency of fifty percent. Now, probabilistic contradictions are evidently not impossible in the way that logical contradictions are, but there is still something undesirable about them: for surely any reasonable account of probability would tell us that if a probabilistic process reliably and robustly deviates from its theoretical probabilities in a certain context, then those theoretical probabilities are simply not the right ones, and we are really just dealing with a different process altogether. Thus it seems reasonable to think that retrocausal mechanisms also cannot exist if they could be used to create probabilistic contradictions, since they will necessarily just become different mechanisms which don't produce contradictions.

Now, observe that one way of using retrocausality or superdeterminism to avoid Type-III disaccord in the Bong et al scenario is to simply postulate that Chidi's specific measurement outcome may be different depending on whether or not Alice performs a supermeasurement. Note that in order to maintain empirical adequacy this dependence relation should leave the overall relative frequencies unchanged, such that Charlie sees relative frequencies compatible with quantum mechanics and Alice also sees relative frequencies compatible with quantum mechanics in the set of experiments in which she asks Charlie about his outcome. However, the constraint that the overall relative frequencies should remain the same is perfectly compatible with the hypothesis that the *specific* outcome that Chidi obtains, on some run of the experiment in which Alice chooses not to perform a supermeasurement, is not the same as the outcome that he would have obtained had Alice chosen to perform a supermeasurement: and this hypothesis is enough to avoid the contradiction derived in the Bong et al theorem, because it means that the values for the outcome pairs CB, CD only need to obey the predictions of quantum mechanics in the case where Alice actually measures C, and not in the case where she instead performs a supermeasurement; thus we need not find a probability distribution over four fixed outcomes A, B, C, D such that all of the pairs $\{AB, AD, CB, CD\}$ obey the predictions of quantum mechanics. That is, the fact that quantum mechanics only predicts relative frequencies and not specific values in this instance makes it possible to use retrocausality to block the argument to the existence of a set of four fixed values without violating the empirical predictions of the theory.

This dependence of Charlie's outcome on Alice's choice need not necessarily be strictly 'causal,' but it does appear to have the character of a dependence relation going backwards in time, and in that sense it might be described as retrocausal. But note that this particular instantiation of backwards-in-time dependence cannot possibly be used to create logical contradictions or probabilistic contradictions! This is because logical or probabilistic contradictions, as in the grandfather paradox, require cyclic dependencies: some past event (e.g. the death of the grandfather) depends on some future event (e.g. the birth of the grandchild) *and vice versa*. But we cannot create such cyclic dependencies here. For example, to create a logical or probabilistic contradiction using the dependence of Chidi's measurement outcome on Alice's choice about whether or not to perform a supermeasurement, we would need to set up an experiment such that Alice's choice about whether to perform a supermeasurement or not also depends on Chidi's measurement outcome, and then arrange the experiment in such a way that these cyclic dependencies produce a logical or probabilistic contradiction. But quantum mechanics tells us that Alice can perform a 'supermeasurement' on Chidi only if she preserves him in a coherent state inside his laboratory, and preserving him in a coherent state means ensuring that no information from inside of his laboratory escapes to the outside. Therefore Alice can only have the option of performing a supermeasurement as long as she has no information about his outcome, and therefore her choice cannot possibly depend on his measurement outcome, because as soon as she has information about his outcome the option to perform a supermeasurement is no longer available to her. Thus it is impossible to set up any cyclic dependencies here, since Chidi's outcome depends only on whether or not Alice performs a supermeasurement, but whether or not Alice performs a supermeasurement cannot possibly depend on Chidi's outcome. Note also that nothing here depends on Alice being a consious observer - we could replace her with a quantum computer or an automaton and the same argument would apply, since the outcome of the automaton's decision process can't depend on Chidi's outcome unless some information about that outcome is available to the automaton, and if that information is available to the automaton then the automaton is no longer maintaining coherent control over Chidi and hence it cannot perform a supermeasurement on him.[11]

So why should we be worried about this kind of retrocausality? There is a certain kind of temporal prejudice which inclines people to think that 'retrocausality' is just impossible - it is not the kind of thing the universe could possibly contain. But here is an alternative proposal: retrocausality is impossible *precisely when it could be used to create logical and/or probabilistic contradictions*. This proposal implies that retrocausality *should* be possible in special circumstances where it cannot be used to create logical or probabilistic contradictions - and that is exactly what we see if we try to introduce retrocausality in order to avoid Type-III disaccord in the Bong et al scenario, so from a certain point of view it would not be so shocking to discover that some kind of retrocausal effect occurs in this setting.

4.1.2 Differing Conceptions of Retrocausality

Another reason for being wary of retrocausality is that the very idea of it seems suspect in a metaphysical or ontological sense. Are we really supposed to posit two different, opposing arrows of causality which somehow 'collide' in the middle to determine intermediate measurement outcomes? To many people this sounds ludicrous. Furthermore, many philosophers take the view that 'causation' is really a macroscopic phenomenon associated with the temporal direction derived from the thermodynamic gradient[36], and to someone who understands causation in this way, it makes no sense to imagine that causation could proceed backwards in time from

[11] A similar point was made by Price in ref [35] about the kind of retrocausality needed to preserve locality in Bell scenarios, but we are not limiting ourselves to local retrocausality here.

Alice's choice to Chidi's outcome: causation is tied to the thermodynamic gradient, and the experiments involved in the non-absoluteness theorems do not involve any abnormal fluctuations of entropy which could conceivably be seen as reversing the thermodynamic gradient.

However, the kind of 'retrocausality' needed to resolve the non-absoluteness theorems need not be understood in terms of a literal backwards arrow of causality. All we need is to say that Chidi's result may depend on Alice's choice - it is not necessary to interpret this dependence as causal in any strong sense. For example, ref [37] argues that the right way of understanding this kind of dependence is to think of the entire set of events as being determined 'all-at-once,' so there is neither a forwards nor a backwards arrow of causality at the fundamental level; rather the past and the future mutually depend on one another, in which case we would naturally expect to find instances where there is something like a dependence relation going backwards in time. In such a picture it is no trouble at all to have Chidi's outcome depending on Alice's choice, since the events are mutually adjusted to fit the constraints imposed by the relevant quantum-mechanical laws - which will always be possible, since we have just seen that the dependence of Chidi's outcome on Alice's choice cannot possibly produce any contradictions. And note that this fundamentally symmetric picture is very friendly to views of causation which see it as a macroscopic phenomenon associated with the thermodynamic gradient, as discussed in greater detail in refs [38]. So metaphysical or ontological objections to the notion of retrocausality are not necessarily good reasons to rule out a dependence of Chidi's outcome on Alice's measurement in the Bong et al scenario.

5 Relational Interaction States

Based on the preceding discussion, we do not believe that the original Wigner's Friend scenario or the more recent non-absoluteness theorems offer any compelling reason to postulate some kind of metaphysically radical non-absoluteness, even if one is determined to maintain the universality of unitary quantum mechanics and one has already ruled out all MOPO approaches. However, our discussion has nonetheless shown that these theorems may suggest the existence of a less radical kind of non-absoluteness, according to which *interaction states* are relativized to observers whilst intrinsic conditions and hence observed events are still objective and observer-independent. After all, even in a completely mundane world composed entirely of observer-independent, 'absolute' events, we can imagine a scenario in which the statistics for Bob's measurements on Alice are different for the statistics for Chidi's measurements on Alice in a given prediction context, meaning that Alice may have different interaction states relative to Bob and Chidi in that prediction context, even though she still has just one intrinsic condition.

This possible avenue for making sense of the non-absoluteness theorems has probably been obscured by the long-standing tradition of conflating intrinsic conditions with interaction states. And of course, we should acknowledge that historically there has been a good reason for treating the two as interchangeable - locality. For locality tells us that the interaction between a system and a measuring device cannot depend on anything other than the intrinsic conditions of the system and the measuring device and their relative arrangement, and therefore given the intrinsic condition of a system we can always write down a unique 'interaction state' which specifies the outcome or probability distribution over outcomes that would be obtained for every local interaction this system could possibly have with a measuring device. Thus in a local theory, it will always be possible to construct a simple map from the intrinsic condition of a system to its interaction state in a way which is valid for any possible measuring instrument, so there is little harm in using just one mathematical object to represent both of them.

But in a non-local theory, such as quantum mechanics arguably appears to be, it need not be the case that the interaction between a system and a measuring device depends only on the intrinsic conditions of the system and the measuring device and their relative arrangement, so this simple prescription for mapping an intrinsic condition to a interaction state may no longer work. And indeed, the Wigner's friend scenario and non-absoluteness theorems appear to be telling us that if unitary quantum mechanics is universal in a first-person sense this must be the case. For example, in the prediction context of the Wigner's Friend scenario after Chidi has measured system S, if unitary quantum mechanics is correct in the first-person sense then the predictions for the result that Chidi would get if he were to perform another measurement on S are different from the predictions for the result that Alice would get if she performed the same measurement on S, even if the two of them were to use identically configured measuring instruments.[12] This suggests that in a quantum context, even if we fix the prediction context we will not in general be able to write down a unique 'interaction state' for a system specifying just one probability distribution over outcomes for each possible configuration of the measuring device, because we need to take into account facts about the person who is operating or who will later read the measuring device, and/or facts about recent history, and/or facts about the broader context of the measurement, all of which may influence the measurement interaction in subtle spatially and temporally non-local ways. So we will end up with a complex, non-local, and probably somewhat 'retrocausal' dynamics for the theory - but this complex dynamics can still be understood as producing well-defined, non-relative events and states of affairs, so although the dynamics are unusual, the metaphysics need not be particularly radical.

The special features of interaction states in the quantum context also help show why it is a mistake to elide the interaction state with the intrinsic condition. For intrinsic conditions (if they exist) are facts about individual systems in and of themselves, whereas interaction states are facts about the relation between a system and a measuring device, an observation which becomes less trivial in the non-local quantum context. This explains straightforwardly why interaction states can be

[12] Probably they could not actually use completely identical measuring instruments, since Alice must perform her experiment in a way which maintains full coherent control over the Chidi-S system, but perhaps she could arrange that at least the part of her device which interacts locally with S is configured in the same way as Chidi's instrument would have been.

relativized even if intrinsic conditions are non-relative: a measurement is a dynamical interaction between the system being measured and the system doing the measuring, so it is entirely reasonable that the outcome of Alice's measurements on Chidi may not be entirely determined by facts about Chidi, since features of Alice and her history and her broader context may also be relevant. In addition, interaction states must be relativized to an prediction context, since adding additional information may change the predictions, whereas intrinsic conditions are not relativized to an prediction context. And finally, intrinsic conditions (if they exist) are features of a system at a time or at least over a relatively short temporal interval, whereas we know that in quantum mechanics the quantum state can only be established by performing tomographic measurements on a large number of identically prepared systems, so if interaction states are to be identified with quantum states (or quantum states with exceptions for measurements in certain bases, as suggested in section 4), then possibly they should perhaps be thought of in terms of some kind of spatially and temporally non-local coordination, rather than as features of individual systems. Thus although classical physics allowed us to get away with identifying interaction states with intrinsic conditions, in a quantum context it appears that intrinsic conditions and interaction states are not at all the same kind of object - indeed, arguably a lot of misunderstanding may have arisen from the decision to call the quantum wavefunction a 'state' even though it is arguably better thought of as an *interaction* state, and therefore it is disanalogous in many ways from the traditional classical concept of state, which is closer to our notion of intrinsic condition.

So we can certainly have a coherent and interesting 'relational' view which is compatible with the universality of quantum mechanics in the first-person sense but which does not require any metaphysically radical non-absoluteness. For the hypothesis that the outcomes of measurement interactions are determined by the features, histories and/or broader context of both systems involved is perfectly compatible with, and indeed *requires*, the hypothesis that these systems have well-defined non-relative features, histories and/or contexts which determine the outcome of the measurement interaction. So metaphysically radical hypotheses are not necessary to maintain the universality of quantum mechanics in the face of the Wigner's Friend paradox: the suggestion that *interaction* states are relativized to observers is already a powerful and far-reaching idea which can potentially resolve a number of puzzles about quantum mechanics without any need to deny the absoluteness of observed events.

5.1 Universality Versus Completeness

At this juncture one might object that postulating intrinsic conditions for quantum systems which are distinct from interaction states looks like it might be incompatible with the assumption of the universality of unitary quantum mechanics, if we take it that the interaction state for a given prediction context is the same as the quantum state we would naturally assign for that context. That would be problematic, since the desire to maintain the universality of unitary quantum mechanics was a key motivation for making interaction states relational in the first place. However, it is important to distinguish between the proposition that unitary quantum mechanics is ‘universal’ and the proposition that it is ‘complete.’ To say that unitary quantum mechanics is ‘universal’ in the first-person sense is to say that it is always a correct description of all the observations made by any individual observer, which is a claim about the dynamics of the theory, whereas to say that quantum mechanics is ‘complete’ is to say that once we have specified the quantum state of a given system (at a given time), there is nothing further to say about it (at that time) - and this is evidently a claim about kinematics or perhaps ontology, rather than about dynamics. Moreover, we have noted already that since quantum mechanics makes only probabilistic predictions, it is possible that quantum mechanics is ‘universal’ in the sense that it always predicts the correct relative frequencies for a certain kind of observation, whilst failing to be ‘complete’ because there are certain measurements for which one could in principle make more specific predictions if one had additional information - for example, we noted that this occurs in RQM+CPL. Furthermore, we have seen throughout this paper that upholding the ‘universality’ of quantum mechanics tends to lead towards views in which the quantum state is relational, whereas once we start making quantum states relational the claim that quantum mechanics is ‘complete’ doesn’t even seem well-formulated, because in that case one cannot simply specify the quantum state of a given system; one must always specify the state relative to some observer, and then any time we have specified a quantum state of a system there will always be something further to be said about it, since it may have different states relative to other observers. So it seems clear that quantum mechanics can be ‘universal’ in the sense that it always predicts the correct measurement outcomes for any individual observer, even if it is not ‘complete.’

Moreover, the scientific reasons for thinking that quantum mechanics must be universal, as discussed in section 2.2, are principally to do with the problems that come from trying to stitch together two separate kinds of dynamics (e.g. the fact that we have no empirical evidence that the unitary dynamics of quantum mechanics ever gives way to any other dynamics, and the fact that it is difficult to combine two different dynamical processes together in a way that works for quantum field theory). So these reasons for thinking that quantum mechanics may be universal do not necessarily give us any reason to think that it must be *complete*, for the claim that quantum mechanics is complete is an assertion about kinematics or ontology, and therefore denying it does not entail that there must exist some second kind of dynamical process that we must integrate with the quantum dynamics. Thus there is arguably less justification for insisting on completeness than there is for insisting on universality in this context, so we should at least be open to views where interaction states diverge from intrinsic conditions in such a way that quantum mechanics is universal but not complete.

6 Examples

We have offered a sketch of a solution to the measurement problem which, in virtue of making use of relativized interaction states and a subtle kind of retrocausality, is able to maintain the universality of unitary quantum mechanics and the existence of only one outcome per observer without postulating any metaphysically radical non-absoluteness. Lest anyone suspect that this is a mere fantasy, we will now give some examples of existing interpretations of quantum mechanics which implement this vision.

6.1 Relational Quantum Mechanics with Cross-Perspective Links

Relational quantum mechanics (RQM), in its original form set out in ref [39], exhibits Type-III disaccord - for one of the postulates of RQM given in ref [39] specifically tells us that it is not even meaningful to compare the perspectives of different observers, except by invoking a third system with respect to which the comparison is made. And clearly if we cannot even compare the perspectives of different observers from a non-relativized perspective, we cannot maintain that there is some physical interaction which reliably brings these perspectives into agreement from a non-relativized perspective, so Type-III disaccord must be generic in this picture.

Recognising the epistemic problems posed by this failure of intersubjectivity, ref [40] proposes an alteration to RQM. The prohibition on comparing perspectives is removed, and instead a new postulate known as Cross-Perspective Links (CPL) is added: '*In a scenario where some observer Chidi measures a variable V of a system S, then provided that Chidi does not undergo any interactions which destroy the information about V stored in Chidi's physical variables, if Alice subsequently measures the physical variable representing Chidi's information about the variable V, then Alice's measurement result will match Chidi's measurement result.*' Ref [40] specifies that the information about V is destroyed (relative to Alice) precisely when Alice performs a measurement on Chidi which does not commute with the variable associated with that information - which is to say, once Alice has performed a super-measurement, the information about Chidi's measurement is no longer accessible to her in subsequent measurements.

RQM+CPL exhibits Type-II disaccord: it tells us that in a Wigner's friend scenario, Chidi's intrinsic condition is that he has seen a single definite measurement outcome, but nonetheless over many repetitions of the experiment Alice will see measurement outcomes consistent with the state ψ, so the interaction state for this prediction context is ψ, However, this is a case as described in section 4, where adding information about Chidi's outcome to the prediction context changes the interaction state - in the new context the interaction state is ψ with the exception that if Alice measures in the basis corresponding to the variable of Chidi that records his measurement outcome, then she is guaranteed to get the same outcome as Chidi. One might worry that this will lead Alice to see outcomes which are incompatible with quantum mechanics, but in fact it follows from the linearity of unitary quantum mechanics that the expected statistics for Chidi's measurement are the same as the expected statistics for Alice's measurement in the corresponding basis on the state ψ, so over many repetitions of this scenario Alice will indeed see the statistics predicted by the state ψ.[13] Thus we have Type-II disaccord but not Type-III disaccord: although there are some mismatches between intrinsic conditions and interaction states CPL is specifically designed to ensure that, when Alice correctly measures the variable encoding Chidi's measurement result, Alice's measurement result will match what Chidi himself saw.

Moreover, although in RQM+CPL observed outcomes are dynamically relevant only to measurements performed in one special basis, macroscopic observers almost exclusively end up measuring one another in that special basis, because to use any other basis they would have to maintain complete coherent control in order to perform a supermeasurement, and decoherence makes this more or less impossible. Thus although the complete interaction state of a macroscopic observer relative to another observer, in an prediction context specified to a high level of microscopic detail, would not usually be the eigenstate that we would naturally associate with their intrinsic condition, nonetheless the full details of this interaction state will generally be inaccessible to other observers, because the tomographic measurements that would reveal it are almost impossible to actually perform. This explains how it is that in this picture we can expect to end up with a stable quasi-classical intersubjective reality in which most interactions simply deliver information about intrinsic conditions, depiste the fact that interaction states do not generically map onto intrinsic conditions in the way one would naturally expect.

Note that since RQM+CPL does not exhibit Type-III disaccord and does not allow more than one outcome per observer, but nonetheless also maintains the universality of unitary quantum mechanics in the first-person sense, it must violate one of the other assumptions of the theorem of Bong et al. And in fact we think a sensible formulation of RQM+CPL would certainly violate No-Superdeterminism.[14] Specifically, the kind of subtle retrocausality discussed in section 4.1.2 actually seems quite natural in the context of RQM+CPL, because in that picture we are required to

[13] We have not been able to find any cases in which the cross-perspective-links postulate definitely leads to a prediction which is inconsistent with the predictions of unitary quantum mechanics, but no one has given a general proof of consistency, so it is possible that some inconsistency could still be found - it would certainly be a worthwhile project to investigate this further! However, we caution that attempts to show inconsistency should keep in mind that RQM+CPL can be formulated in a way that allows a subtle kind of retrocausality, which may help it avoid possible inconsistencies - indeed, we will shortly see an example of this with reference to the Lawrence et al theorem.

[14] It is also possible that it would violate Locality. Certainly, RQM+CPL is in a sense non-local, if it is understood in terms of an 'all-at-once' model, since such a model allows generic correlations between spacelike, timlike and lightlike separated events without any need for local mediation. The question then is whether RQM+CPL violates Parameter Independence (PI), in which case it violates the Locality condition of Bong et al, or just Outcome Independence (OI), in which case it does not violate the Locality condition. As a matter of fact, we suspect that the distinction between PI and OI only makes sense in a causal, time-evolution model, so in the all-at-once context it may not be possible to separate the generic nonlocality out into PI and/or OI; thus it is likely that RQM+CPL would violate PI in a formal way, though in the all-at-once context that would not have the same implications it is claimed to have in the time-evolution context, and in particular it would not necessarily entail superluminal action or a preferred reference frame.

postulate the existence of a large network of *absolute, observer-independent* events distributed over spacetime in such a way that they exhibit non-local quantum correlations across both spacelike and timelike separations, and it is has been recognised in the context of the Bell flash ontology that in order to have relativistic covariance in a picture based on an ontology of correlated pointlike events, we typially have to '*renounce any account of the coming-into-being of the (events).*' That is, we must allow that the network of events is determined in an 'all-at-once' manner, rather than being generated in some temporal order. And it was already noted in section 4.1.2 that backwards influences which look superficially like retrocausality will likely be generic in an 'all-at-once' picture, so we probably shouldn't be too surprised to learn that in RQM+CPL we will need to violate the No-Superdeterminism assumption.

This response also applies to the argument of Lawrence et al, which was originally intended specifically as a criticism of RQM. As pointed out by ref [41], the argument of Lawrence et al clearly fails to land if it is directed at the original version of RQM, because this version of RQM tells us that the measurement outcomes obtained separately by Alice and Bob can never be compared, so the universality of unitary quantum mechanics in the first-person sense does not entail that these outcomes must obey the quantum constraints obtained by Lawrence et al. However, one might initially think that the argument would work if it is directed at RQM+CPL, since the whole point of adding the CPL postulate is to allow observers to gain access to the observations of other observers. And indeed, Lawrence et al argue that in RQM+CPL all of the outcomes obtained by Alice and Bob are real for both of them: '*in fact all outcomes exist for all observers, but some of them are just unknown*' - and thus, they appear to think, in this case the six outcomes must definitely all of obey the constraints obtained from quantum mechanics.

As a matter of fact, we do think it is correct to say that in RQM+CPL all outcomes exist for all observers, with some of them simply being unknown; and we also agree that in RQM+CPL, if Bob measures Alice in an appropriate basis he can learn the values of A_1, A_2 and A_3, so he could in principle access any one of the four trios in the set $\{B_1B_2B_3;B_1A_2A_3;A_1B_2A_3, A_1A_2B_3\}$. Thus the constraints derived by Lawrence et al will have to be satisfied for the trio that he does in fact access. However, Lawrence et al are wrong to assume that all four constraints must always be obeyed. Bob cannot access all four trios *on any given run of the experiment*, since measuring B_1 is incompatible with measuring A_1 and so on. And in order to uphold the universality of unitary quantum mechanics in the first-person sense, RQM+CPL does not need to respect the quantum constraints for all outcomes which could in principle have been accessed by a given observer; it only has to respect the quantum constraints for all outcomes which are *actually* accessed by a given observer on a given run of the experiment. Thus even with CPL, universality in the first-person sense does not entail that all four trios must obtain the quantum constraints obtained by Lawrence et al, so the proof does not go through.

Lawrence et al might try to argue, like Bong et al, that since Bob could measure any combination of these variables, then if Bob's results always match the predictions of quantum mechanics we must be able to pre-specify values for all six variables which obey the four constraints obtained from quantum mechanics, which their proof shows to be impossible. However, as in Bong et al, we only need to be able to pre-specify values matching the quantum constraints for all four trios if we assume that the outcomes obtained by Alice are independent of Bob's subsequent choices and that Bob's individual outcomes B_1, B_2, B_3 are all independent of what other measurements Bob chooses to make, and thus this argument, like the Bong et al one, assumes Locality and No-Superdeterminism. Therefore since RQM+CPL is perfectly compatible with a subtle retrocausal effect of Bob's choices on Alice's measurements, it does not run into a contradiction here; so, contrary to their claims to have ruled out relative facts, the argument of Lawrence et al does not rule out either the original version of RQM or RQM+CPL, properly understood.

6.2 Kent's Lorentzian Solution to the Quantum Reality Problem

Another way of implementing Type-II disaccord without Type-III disaccord involves postulating a model in which there are fewer observed events then we would naturally imagine. That is, we could say that there is not always an actual observed event associated with an instance in which the unitarily evolving wavefunction appears to represent an observer as making an observation, thus allowing us to evade the non-absoluteness theorems by saying that all observed events are absolute but some of the events featuring in these theorems are not actually observed.

For example, this occurs in Kent's Lorentzian solution to the quantum reality problem[42–44]. Kent's proposed interpretation of quantum mechanics is based on a simple idea: there is no collapse of the wavefunction, so we just allow the wavefunction to undergo its standard unitary evolution over the whole course of history, and then, at the end of time, we imagine a single measurement being performed on the final state to determine the actual content of reality. In Kent's words, '*an event occurs if and only if it leaves effective records in the final time... measurement*' [42]. So for example in ref [44] Kent imagines a final measurement on the positions of photons which have been reflected off matter at various points, and then defines the beables at x by the expectation value of some operators (e.g. the stress-energy tensor components) at x conditional on the detections of photos outside the future lightcone of x.

In Kent's picture, all observed events are 'absolute' but there are a few cases where one might expect an event to be observed but as a matter of fact no event gets observed. In general, Kent's approach entails that macroscopic events will reliably occur as expected, since decoherence encodes these events robustly in the state of their environment and thus they will be actualised by the final measurement. And of course 'observed events' are always macroscopic events - for even if the event itself is not macroscopic, the observer and the records made in their brain are macroscopic - so 'observed events' are typically actualised, which is to say they actually occur and are actually observed. But in a special case where all records of a macroscopic events are subsequently destroyed, that macroscopic event will *not* be actualised by the final measurement; so if the macroscopic event is an observed event, that event will not occur and will not be observed. And this is exactly what happens in the non-absoluteness theorems - for by definition, a supermeasurement can be performed on an observer only if decoherence is controlled and all records of their previous observations are completely erased by the supermeasurement. So the non-absoluteness theorems pertain to exactly the kinds of special cases in which Kent's model says certain observations do not in fact occur. For example, in the Bong et al experiment, whenever Alice performs a supermeasurement on Chidi that supermeasurement destroys all the records of Chidi's measurement, so according to Kent's approach Chidi's measurement outcome is simply not an 'observed event' in this instance, and therefore absoluteness of observed events does not entail that it must have a unique, well-defined outcome.

As with RQM+CPL, we can think of Kent's approach as postulating a divergence between interaction states and intrinsic conditions. For example, in the Bong et al experiment, Kent's approach predicts that Alice's measurement outcomes will always be as if Chidi is in the state ψ, so Chidi has interaction state ψ relative to Alice in this prediction context. But we also know that if Alice in fact chooses to perform a supermeasurement, Chidi's measurement is not actualised by the final measurement, so Chidi may not have an intrinsic condition at all during this time, or at least, his intrinsic condition will not be at similar to the intrinsic condition normally associated with a classical observer having a conscious experience - in a sense he doesn't really exist at all during this time. Yet Alice is able to dynamically access his degrees of freedom and perform measurements on him even though he 'does not exist,' thus exhibiting a dramatic divergence between his intrinsic condition and his interaction state relative to Alice. Thus Kent's approach exhibits a version of Type-II disaccord in the context of the Bong et al experiment: if Alice assumes that her observations indicate that Chidi is having some kind of indefinite 'superposed' experiences relative to her, or at least that he is having *some* experiences, then her inference will fail to match Chidi's experiences, since actually he is not having any experiences at this time.

However, Kent's approach does not exhibit Type-III disaccord: if Alice asks Chidi about his measurement outcome, then the conversation itself forms a part of the records of his measurement outcome, and by construction the records arriving at the final state must match the event which is actualised by the final measurement, so what Alice hears Chidi saying will indeed match what he originally perceived. And note that quantum mechanics is universal in the first-person sense in Kent's view, because it entails that no observer will ever observe anything that is incompatible with unitary quantum mechanics, although there are some cases where this is achieved only in virtue of certain observers failing to observe anything at all. Thus we can conclude that Kent's approach must violate one of the other assumptions of the Bong et al theorem; and indeed, it evidently exhibits the subtle kind of retrocausality discussed in section 4.1, since we have seen that the result of Chidi's measurement depends on what Alice chooses to measure. Note however that while in RQM+CPL Chidi may get two different results depending on what Alice chooses to measure, in Kent's picture it is instead the case that if Alice chooses to ask Chidi about his outcome Chidi gets some result, whereas if Alice performs a supermeasurement Chidi does not get any result at all. So although both RQM+CPL and Kent's approach make use of relational interaction states and retrocausality, they implement this in quite different ways[15].

7 Ormrod and Barrett

We have deferred discussion of the Ormrod and Barrett result to this final section because it is importantly different to the other cases. For the Wigner's Friend scenario and the Bong et al and Lawrence et al theorems, we were able to show that even if MOPO approaches are ruled out, and even if we are determined to maintain Locality and No-Superdeterminism, what is demonstrated by these theorems is only the existence of disaccord, not necessarily any kind of metaphysically radical non-absoluteness; therefore we were able to suggest a way of responding to these theorems in a way which maintains Locality, No-Superdeterminism, the existence of only one outcome per observer and the universality of unitary quantum mechanics in the first-person sense without any metaphysically radical non-absoluteness, by simply relativizing interaction states rather than the observed events.

But no such option exists for the Ormrod and Barrett case: in this experiment, FIQT predicts that there is a non-zero probability that Alice and Bob get outcomes such that it is mathematically impossible that all observers have exactly one non-relativized outcome and all of these unique outcomes obey FIQT for any choice of spacelike slice. Therefore on runs of the experiment when Alice and Bob do get this outcome, it follows that if all four observers have a unique non-relative outcome, on at least one spacelike slice there must be a pair of outcomes which is predicted by FIQT to be impossible; so FIQT cannot be universally correct if all measurement outcomes are unique and non-relativized. Relativizing interaction states doesn't help here, since the theorem doesn't require any observers to communicate their results to one another and thus it doesn't make any assumptions about the relation between the outcomes of dynamical interactions with an observer and the actual experiences of that observer. Adopting a picture like Kent's does help in a certain sense - it would resolve the apparent contradiction by simply saying that the measurements of Chidi and Divya don't actually take place, since all records of them are subsequently erased, so we don't

[15] One might worry that Kent's approach has avoided the epistemic problems associated with Type-III disaccord only by introducing another kind of serious epistemic problem, since it entails that some of our beliefs about the past are very seriously mistaken - sometimes we believe an event has occurred when in fact it did not occur at all. This is a significant revision of our usual view of the past, and one might think it could undermine our scientific practice of using records of the past to empirically confirm theories. However, note that in Kent's picture, the only kinds of events which fail to occur are events such that there are no ongoing records of the relevant event; so in Kent's picture, we may rest assured that whenever we *do* have access to a record of an event, that event most likely did occur in the way suggested by the record, and if it did not the reason is just an ordinary one like human error. So even though in Kent's picture it is occasionally true that events don't occur when we expect them to, nonetheless we can still expect records of past events to be generally reliable and thus we can make use of them for scientific confirmation.

get outcomes predicted to be impossible by FIQT on the Chidi-Divya spacelike slice because we don't get outcomes at all on that slice. But presumably Ormrod and Barrett would take the view that such an approach does not fully uphold their assumption of the universality of FIQT, since it entails that some outcomes predicted by FIQT simply don't occur at all.

Nor can we straightforwardly argue that what really follows from the assumptions of the theorem is disaccord rather than non-absoluteness. We *can* obtain Type-I disaccord from this scenario, but only by taking a deflationary appraoch to FIQT and regarding it as a description, not of the actual measurement outcomes that occur, but of some inferences that could reasonably be made by observers in a scenario of this kind. So for example, suppose Alice and Bob compare their results and then make an inference about Divya's measurement result by imagining a collapse on the Alice-Divya spacelike slice and using what they know about Alice's outcome, and an inference about Chidi's measurement result by imagining a collapse on the Bob-Chidi spacelike slice and using what they know about Bob's outcome, and likewise, Chidi makes an inference about Divya's measurement result by imagining a collapse on the Chidi-Divya spacelike slice and using his knowledge of his own outcome, and Divya makes an inference about Chidi's result in a similar way. Then the import of the Ormrod and Barrett theorem is that there exists a possible set of measurement outcomes for Alice and Bob such that it is mathematically guaranteed that not all of these inferences made in accordance with FIQT can be correct. Thus we arrive at Type-I disaccord, because at least one person in this scenario must end up making a wrong inference about someone else's outcome, even though they applied quantum mechanics correctly. But evidently this approach does not really maintain the universality of unitary quantum mechanics in the FIQT sense, since it involves accepting that on at least one spacelike slice the predictions of FIQT are not right.

So the theorem of Ormrod and Barrett seems like it really could provide an argument for metaphysically radical non-absoluteness: if quantum mechanics is truly universal in the FIQT sense, and if MOPO approaches have been ruled out, then the only remaining option is that there must be some kind of radical relativization going on. However, it is important to note that this follows only because of the stronger universality assumption made by Ormrod and Barrett - no such result can be derived using purely first-person universality. And therefore the kind of 'non-absoluteness' involved here cannot be the kind of non-absoluteness that features in ordinary relational and perspectival approaches to quantum mechanics, for approaches of this kind typically postulate only first-person universality, and indeed they often forbid comparisons between measurement results obtained by different observers. Thus in many relational and perspectival approaches it is simply not meaningful to apply quantum mechanics to jointly predict the results obtained by two different observers on the same spacelike slice in the way that Ormrod and Barrett do, unless of course those observers later interact in some way which makes the outcome of one physically relevant to the outcome of the other, which is not the case in the Ormrod and Barrett scenario. Therefore it should be kept in mind that the assumptions going into the Ormrod and Barrett non-absoluteness theorem are incompatible with a number of the metaphysically radical interpretations that are sometimes argued to be supported by the non-absoluteness theorems. Of course, it is true that any relational or perspectival approach will avoid the contradiction that Ormrod and Barrett describe, but most such approaches do this in virtue of failing to be universal in the FIQT sense, and thus they are not really solving the problem posed by Ormrod and Barrett, which is about how to *maintain* universality in the FIQT sense.

And in fact, it seems quite hard to understand what kind of non-MOPO relational or perspectival approach *could* maintain universality in the FIQT sense. First note that in a non-MOPO relational approach, it seems hard to avoid relativizing measurements to observers. For if we instead have outcomes relativized to frames of reference or some such object, we end up with a description of observers seeing different outcomes in different frames of reference; and then if we can find no natural way to decide which frame of reference determines the conscious experience of the observer, we would presumably have to accept that the observer has different conscious experiences in each frame, giving rise to a MOPO approach[16] Moreover, in a relational approach which maintains universality in the FIQT sense, we have to be able to specify facts about Alice's outcome and Divya's outcome relative to the same referent, since FIQT requires us to say that Alice and Divya's measurements together obey the predictions of quantum mechanics. So it seems that in a non-MOPO relational approach which maintains universality in the FIQT sense, we will ultimately have to say that in the Ormrod and Barrett experiment there is some fact about Divya's measurement outcome relative to Alice, or about Alice's measurement outcome relative to Divya.

However, in most relational and perspectival approaches we are called upon to make reference to 'Divya's measurement outcome relative to Alice' only in cases where Alice physically measures some variable encoding Divya's outcome, or where Divya's outcome is correlated with or has consequences for something which Alice can physically interact with after the experiment has concluded; the existence of real interactions of this kind establishing real physical relations between the observers is what endows the notion of 'Divya's measurement outcome relative to Alice' with physical content. Whereas in the Ormrod and Barrett experiment Divya's measurement outcome is erased before it can have any consequences for Alice, so it is not clear what physical significance claims about 'Divya's measurement outcome relative to Alice' could possibly have in this context - if this relativized outcome is not relevant to any of Alice's future predictions or interactions, and it is not a fact about Divya's experiences or Alice's experiences, what *is* it a fact about? Perhaps one might think that 'the outcome of Divya relative to Alice' in this context has *epistemic* rather than physical significance: Divya's measurement outcome 'has a value' relative to Alice in the sense that Alice believes something about it, or it would be rational for her to believe something about it based on her understanding of quantum mechanics. But as we have already noted, this epistemic approach would surely not satisfy anyone who is determined that quantum mechanics must be universal in the FIQT sense, since it just amounts to saying that inferences made using FIQT may be reasonable but they are nonetheless sometimes wrong.

[16] Ormrod and Barrett themselves have been developing a novel MOPO approach of this kind[45].

Thus we are not convinced that there is any coherent version of metaphysically radical non-absoluteness which can be supported by the Ormrod and Barrett theorem. Rather, the theorem is probably best understood as an argument in support of a MOPO approach, such as the Everett interpretation, or an approach with outcomes relativized to reference frames[45]. And if MOPO approaches were to be ruled out, then perhaps the appropriate conclusion to draw from the theorem would simply be that quantum mechanics cannot possibly be universal in the FIQT sense - if there is no coherent version of 'non-absoluteness' which would allow us to reconcile the universality of quantum mechanics in the FIQT sense with the stipulation that all observers always see a single outcome to a given measurement, then for those who believe that all observers should always see a single outcome, the theorem becomes a reductio ad absurdum against FIQT.

In this connection, it is important to note that although FIQT is motivated by the desire to have compatibility with relativity, it is certainly possible to have relativistically covariant approaches which exhibit universality only in the weaker first-person sense. For example, Kent's approach is explicitly designed to be Lorentz-covariant, and yet it has the consequence that quantum mechanics is universal only in the first-person and not in the FIQT sense, since it tells us that some of the events predicted by FIQT do not actually occur. Similarly, RQM+CPL appears to be relativistically covariant when formulated as an 'all-at-once' theory, but it tells us that quantum mechanics is universal only in the first-person and not in the FIQT sense, since it does not require that quantum mechanics makes correct predictions when applied across the perspectives of two observers who never have an opportunity to share their results. So the claim made by Ormrod and Barrett, '*Given a version of quantum theory that models measurements unitarily and which fits naturally with special relativity, it simply cannot be true that there are absolute observed outcomes*' is not completely right: their vision of what it might look like for a version of quantum theory to 'fit naturally with special relativity' is too narrow, as it fails to take into account the possibility of models in which there are fewer observed events than we might naturally imagine, or all-at-once models with relational interaction states, or other novel possibilities that have not yet even been thought of. So in our view it is reasonable to respond to the Ormrod and Barrett theorem by simply rejecting FIQT, although of course the MOPO route still remains open.

Additionally, it is very interesting that the universality of quantum mechanics in the first-person sense appears to be perfectly compatible with the existence of only one outcome per observer *and* the absence of any metaphysically radical non-absoluteness - only a much stronger universality assumption could compel us to accept either a MOPO approach or some kind of metaphysically radical non-absoluteness. Since universality in the first-person sense is naturally associated with relational and perspectival views, this suggests that the strong metaphysical claims often tied to such views may be a misunderstanding of what quantum mechanics is really telling us about relationality - such approaches should all along have been understood as being *dynamically* rather than metaphysically relational. Additionally, since we can only really have direct empirical evidence for universality in the first-person sense and not in the FIQT sense, this conclusion may perhaps be regarded as an argument in favour of the absoluteness of observed events: our immediate empirical evidence seems to be telling us that the world has relational features but that it stays strictly within the limits of relationality compatible with having one outcome per observer without any metaphysically radical non-absoluteness, and one natural explanation for that would be that observed events are, in fact, absolute!

8 Conclusion

In this article we have reached a somewhat negative conclusion about the idea of 'non-absoluteness of observed events.' That is, we agree that the Everett interpretation and other MOPO approaches would genuinely involve something that could be referred to as non-absoluteness of observed events, but we are not convinced that the Wigner's Friend scenario or the non-absoluteness theorems offer compelling evidence for any alternative approach involving metaphysically radical non-absoluteness. For we have seen that the Wigner's Friend scenario and the Bong et al and Lawrence et al theorems can be understood simply as demonstrating that if unitary quantum mecahnics is universal and certain auxiliary assumptions hold, then there must exist instances of Type-II and/or Type-III disaccord, and we have argued that these kinds of disaccord do not need to be understood in terms of metaphysically radical non-absoluteness, since we always have the option of relativizing the *interaction* state whilst maintaining the absoluteness of observed events. Of course Type-III disaccord is certainly unappealing in many ways and we agree that it should be avoided, but postulating metaphysically radical non-absoluteness does not help at all with this problem and indeed makes things significantly worse. Meanwhile, the theorem of Ormrod and Barrett does seem like it could offer a real argument for metaphysically radical non-absoluteness, but it only achieves this by making an unusually strong universality assumption, and it is unclear that accepting metaphysically radical non-absoluteness is a better option than simply rejecting this strong universality assumption (or alternatively, adopting a MOPO approach).

Now, we suspect that some of our negative comments about metaphysically radical non-absoluteness may invite a response of the form: 'You are prejudiced against this kind of picture because you are too attached to your naive classical worldview and are thus incapable of properly comprehending a radically relational/perspectival approach.' And this may be so. But our central point in this article is simply that these kinds of radical approaches are not *necessary*; and given that they appear to lead to severe problems for the epistemology of science, we probably should not adopt them unless there is genuinely no other choice. Moreover, it seems that relativizing interaction states while maintaining 'absolute' intrinsic conditions and observed events gives us nearly everything we might hope to gain from a relational or perspectival approach but without the associated dangers for scientific rationality, so we believe that the possibility of developing views which are relational in a less radical way deserves more attention than it has so far received.

Finally, although we don't think the non-absoluteness theorems should be interpreted as demonstrating the non-absoluteness of observed events in a metaphysically radical sense, we do think they provide important new insights into quantum mechanics. For example, the scenarios and theorems discussed in this paper together both make a clear case for approaches with relational interaction states and also give a sense of the shape which an approach with dynamical interaction states would have to take - in particular, the results of Bong et al and Lawrence et al suggest that if we want to avoid Type-III disaccord whilst also continuing to maintain the universality of unitary quantum mechanics in the first person sense and the existence of only one outcome per observer, we may be forced to compromise by allowing the violation of Locality, and/or allowing some superdeterminism or retrocausality. So in our view, these recent results demonstrate that the common prejudice against the measurement problem as unscientific and unsolvable is unfounded: the combination of these non-absoluteness theorems, the need to avoid epistemic irrationality, and the mandate to successfully reproduce the predictions of QFT is now forcing us down an increasingly narrow bottleneck, so progress is certainly being made and there is cause for optimism that we may eventually converge on a fully acceptable solution.

Acknowledgements Thanks to Nick Ormrod for reading and discussing a draft of this paper, and thanks to Howard Wiseman and Yilé Ying for very helpful conversations on these topics. This publication was made possible in part through the support of the ID #62312 grant from the John Templeton Foundation, as part of the project 'The Quantum Information Structure of Spacetime' (QISS). The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

Funding Open access funding provided by SCELC, Statewide California Electronic Library Consortium.

References

1. Brukner, Č: Wigner's friend and relational objectivity. Nat. Rev. Phys. **4**(10), 628–630 (2022)
2. Bitbol, M.: Is the life-world reduction sufficient in quantum physics? Cont. Philos. Rev. **54**, 12 (2021)
3. Fedrizzi, A., Proietti, M.: Quantum Physics: Our Study Suggests Objective Reality Doesn't Exist. Accessed (August 2023)
4. Oldofredi, A.: The relational dissolution of the quantum measurement problems. Found. Phys. **53**(1), 1–24 (2022)
5. Pienaar, J.: QBism and relational quantum mechanics compared. Found. Phys. **51**(5), 96 (2021)
6. Wigner, Eugene, P.: Remarks on the mind-body question. In: Philosophical Reflections and Syntheses, pp. 247–260. Springer, New York (1995)
7. Bong, K.-W., Utreras-Alarcón, A., Ghafari, F., Liang, Y.-C., Tischler, N., Cavalcanti, E.G., Pryde, G.J., Wiseman, H.M.: A strong no-go theorem on the wigner's friend paradox. Nat. Phys. **16**(12), 1199–1205 (2020)
8. Lawrence, J., Markiewicz, M., Żukowski, M.: Relative facts of relational quantum mechanics are incompatible with quantum mechanics. Quantum **7**, 1015 (2023)
9. Ormrod, N., Barrett, J.: A No-Go Theorem for Absolute Observed Events Without Inequalities or Modal Logic (2022)
10. Schmid, D., Yīng, Y., Leifer, M.: A Review and Analysis of Six Extended Wigner's Friend Arguments (2023)

11. Wiseman, H.M., Cavalcanti, E.G., Rieffel, E.G.: A “thoughtful" local friendliness no-go theorem: a prospective experiment with new assumptions to suit. Quantum **7**, 1112 (2023)
12. Fuchs, C.A., Mermin, N.D., Schack, R.: An introduction to QBism with an application to the locality of quantum mechanics. Am. J. Phys. **82**, 749–754 (2014)
13. Rovelli, C.: Relational quantum mechanics. Int. J. Theor. Phys. **35**(8), 1637–1678 (1996)
14. Frigg, R.: GRW Theory (Ghirardi, Rimini, Weber Model of Quantum Mechanics), pp. 266–270. Springer, Berlin (2009)
15. Tumulka, R.: A relativistic version of the Ghirardi Rimini Weber model. J. Stat. Phys. **125**, 821–840 (2006)
16. London, F., Bauer, E., Langevin, P.: La théorie de l’observation en mécanique quantique. No Title (1939)
17. Wallace, D.: The sky is blue, and other reasons quantum mechanics is not underdetermined by evidence. Eur. J. Philos. Sci. **13**(4), 54 (2023)
18. Adlam, E.: The problem of confirmation in the Everett interpretation. Stud. Hist. Philos. Sci. Part B **47**, 21–32 (2014)
19. Adlam, E.: Does science need intersubjectivity? The problem of confirmation in orthodox interpretations of quantum mechanics. Synthese **200**(6), 522 (2022)
20. Bohr, N.: The Philosophical Writings of Niels Bohr. Ox Bow Press, Woodbridge (1987)
21. Heisenberg, W.: Physics and Philosophy: The Revolution in Modern Science. Harper, New York (1958)
22. Pauli, W., Enz, C.P., von Meyenn, K.: Writings on Physics and Philosophy. Springer, Berlin (1994)
23. Zeilinger, A.: A foundational principle for quantum mechanics. Found. Phys. **29**(4), 631–643 (1999)
24. Zeilinger, A.: Bell’s Theorem, Information and Quantum Physics, pp. 241–254. Springer, Berlin (2002)
25. Brukner, C.: On the quantum measurement problem. Quantum [Un] Speakables II: Half a Century of Bell‘s Theorem (2015)
26. Bub, J.: Bananaworld: Quantum Mechanics for Primates. Oxford University Press, Oxford (2012)
27. Demopoulos, W.: Generalized probability measures and the framework of effects. In: Ben-Menahem, Y., Hemmo, M. (eds.) Probability in Physics, pp. 201–217. Springer, New York (2012)
28. Janas, M., Cuffaro, M.E., Janssen, M.: Understanding Quantum Raffles: Quantum Mechanics on an Informational Approach-Structure and Interpretation (Foreword by Jeffrey Bub). Springer, New York (2021)
29. de Muynck, W.M.: Towards a Neo-Copenhagen interpretation of quantum mechanics. Found. Phys. **34**(5), 717–770 (2004)
30. Fuchs, C.A., Mermin, N.D., Schack, R.: An introduction to QBism with an application to the locality of quantum mechanics. Am. J. Phys. **82**, 749–754 (2014)
31. Di Biagio, A., Rovelli, C.: Relational quantum mechanics is about facts, not states: a reply to Pienaar and Brukner. Found. Phys. **52**(3), 62 (2022)
32. Esfeld, M., Gisin, N.: The GRW flash theory: a relativistic quantum ontology of matter in spacetime? *ArXiv e-prints*, (October 2013)
33. Cavalcanti, E.G.: The view from a Wigner bubble. Found. Phys. **51**(2), 39 (2021)
34. Moreno, G., Nery, R., Duarte, C., Chaves, R.: Events in quantum mechanics are maximally nonabsolute. Quantum **6**, 785 (2022)
35. Price, H.: A neglected route to realism about quantum mechanics. Mind **103**(411), 303–336 (1994)
36. Price, H.: Causal perspectivalism. In: Price, H., Corry, R. (eds.) Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited. Oxford University Press, Oxford (2005)
37. Adlam, E.: Two roads to retrocausality. Synthese **200**, 422 (2022)
38. Adlam, E.: Is there causation in fundamental physics? New insights from process matrices and quantum causal modelling. Synthese **201**(5), 1–40 (2023)
39. Di Biagio, A., Rovelli, C.: Relational quantum mechanics is about facts, not states: a reply to Pienaar and Brukner. Found. Phys. **52**(3), 62 (2022)
40. Adlam, E., Rovelli, C.: Information is Physical: Cross-perspective Links in Relational Quantum Mechanics (2022)
41. Cavalcanti, E.G., Di Biagio, A., Rovelli, C.: On the consistency of relative facts (2023)
42. Kent, A.: Solution to the Lorentzian quantum reality problem. Phys. Rev. A **90**(1), 012107 (2014)
43. Kent, A.: Lorentzian quantum reality: postulates and toy models. Philos. Trans. R. Soc. A **373**(2047), 20140241 (2015)
44. Kent, A.: Quantum reality via late-time photodetection. Phys. Rev. A **96**(6), 062121 (2017)

45. Ormrod, N.: Quantum influences and event relativity, (September 2023). PIRSA:23090026