Explaining the *Un*observed—Why Quantum Mechanics Ain't Only About Information

Amit Hagar[*] Meir Hemmo[†]

February 28, 2006

Abstract

A remarkable theorem by Clifton, Bub and Halvorson (2003) (CBH) characterizes quantum theory in terms of information–theoretic principles. According to Bub (2004, 2005) the philosophical significance of the theorem is that quantum theory should be regarded as a "principle" theory about (quantum) information rather than a "constructive" theory about the dynamics of quantum systems. Here we criticize Bub's principle approach arguing that if the mathematical formalism of quantum mechanics remains intact then there is no escape route from solving the measurement problem by constructive theories. We further propose a (Wigner–type) thought experiment that we argue demonstrates that quantum mechanics on the information–theoretic approach is incomplete.

Keywords: Quantum information; Collapse theories; Crucial experiments; Principle vs. Constructive theories.

---

[*] Philosophy Department, University of Delaware, Newark, DE 19716 USA; Email: hagar@udel.edu

[†] Philosophy Department, University of Haifa, Haifa, Israel 31905; Email: meir@research.haifa.ac.il

1 Introduction

Quantum information theory has by now become to a large extent a new orthodoxy in the foundations of quantum mechanics. It is sometimes further claimed that the information–theoretic approach brings out the so–called "futility" of the long-standing debates over the interpretations of quantum mechanics.[1] A major conceptual tool enhancing the information–theoretic approach is the remarkable theorem by Clifton, Bub and Halvorson (2003, CBH henceforth) according to which quantum theory can be characterized by three information–theoretic principles: no signaling, no broadcasting and no (unconditionally secure) bit commitment (NO BIT henceforth). The aim of this paper is to examine the information–theoretic approach to quantum mechanics focusing on Bub's (2004, 2005) recent analysis of it and some of its implications.

On the basis of the above three principles of the CBH theorem, Bub (2004, p. 242; see also 2005) argues for the following three theses:

1. A quantum theory is best understood as a theory about the possibilities and impossibilities of information transfer as opposed to a theory about the mechanics of nonclassical waves or particles.[2]
2. Given the three information–theoretic constraints, any mechanical theory of quantum phenomena that includes an account of the measuring instruments that reveal these phenomena must be empirically equivalent to a quantum theory.
3. Assuming the information–theoretic constraints are in fact satisfied in our world, no mechanical theory of quantum phenomena that includes an account of measurement interactions can be acceptable, and the appropriate aim of physics at the fundamental level then becomes the representation and manipulation of information.

In his recent paper in this journal, Bub (2005) depicts the philosophical significance of the CBH theorem as analogous to Einstein's shift from the constructive view of theories—attributed to Lorentz and FitzGerald—towards

---

[1] For such a claim see Fuchs (2002); For a response—Hagar (2003)

[2] By *information* Bub means information in the physical sense as measured, e. g., in quantum mechanics by the von Neumann entropy.

the principle view of theories in the context of the special theory of relativity.[3] His idea is that the distinction between constructive theories and theories of principle is suitable to characterize the difference between the information–theoretic approach and all other interpretations of quantum theory. As the above theses show, Bub believes that if, indeed, the three information–theoretic principles of the CBH theorem hold in our world, then no constructive theory for quantum phenomena is possible that yields different predictions than those of quantum theory.

Bub's three theses have very important implications regarding the understanding of quantum mechanics and the future direction in which the research at the foundations of the theory will go. We agree with Bub's second thesis (that any constructive theory for quantum phenomena which satisfies the three information–theoretic principles of the CBH theorem is empirically indistinguishable from quantum theory). Nevertheless, here we wish to examine Bub's approach and present an alternative view on the philosophical significance of the CBH theorem and on the role of constructive quantum mechanical theories.[4] As Bub (2004) himself notes (and in accord with his second thesis) a certain constructive theory for quantum phenomena, namely the collapse theory by Ghirardi, Rimini and Weber (GRW) *does* give different predictions from those of quantum theory. We conjecture, following Bub (2005), that the GRW theory violates the (above) NO BIT constraint (see below for some discussion about this point). However, the GRW theory itself implies that this violation is *compatible* with everything we know empirically about the physical world. Roughly, the NO BIT constraint implies the *unrestricted* validity of the superposition principle, and in particular, it entails that macroscopic massive systems can be in nonlocal entangled EPR–type states even with respect to their spatial degrees of freedom. But the existence of such macrostates has never been experimentally confirmed, so we do not

---

[3] For more on the constructive approach to special relativity see Jánossy (1971), Bell (1976), and Brown (2006).

[4] We set aside the more general issue of the aim of physics as stated by Bub's third thesis. In this context it is interesting that Maxwell, although accepting the distinction between the physics of principles and the construction of models and even admitting that *in principle* indefinitely many dynamical models can explain certain phenomena, nevertheless devoted his career almost solely to the construction of such models (Harman 2001). Einstein himself, after promoting in 1905 the distinction he borrowed from Maxwell and Poincaré between principles and constructions, shifted back to the constructive view and later on abandoned what he called "the new fashion" which he himself helped creating (Balashov and Janssen 2003).

really know empirically whether or not such states are physically feasible. In this paper we focus only on collapse theories as constructive alternatives to Bub's principle approach.[5] We shall argue for the following alternative theses.

I. No collapse quantum mechanical theories (including information–theoretic approaches, if the latter are committed to no collapse dynamics) and alternative collapse theories, such as the GRW theory are *empirically* distinguishable. This is obvious but in the present context deserves attention.[6]

II. Constructive quantum mechanical theories[7] are necessary to solve the measurement problem (given the current mathemtaical formalism of quantum theory).

III. Quantum mechanics under the information–theoretic interpretation is *incomplete*, and moreover, it leads to inconsistent predictions of the (statistical) outcomes of (some) measurements.

As to Thesis II, we shall also argue that interpreting quantum mechanics as a principle theory is *not* the right epistemological stance given the theoretical basis of quantum mechanics, and that information–theoretic interpretations of quantum mechanics as a principle theory (see Bub 2004, 2005) don't warrant abandoning alternative constructive dynamical theories, in particular theories which differ empirically from no collapse quantum mechanics. We shall further argue that the lesson one should take from the CBH theorem lies not only in the quantum phenomena that are captured by its three

---

[5] We do not address here the subtler issue of whether or not Bohmian mechanics might be distinguished empirically from other no collapse theories such as modal and many worlds theories and Bub's principle approach. We agree with Bub (2005) that Bohmian mechanics is empirically equivalent to no collapse quantum mechanics (but compare Valentini 2002). On the other hand, questions of theory choice depend on quite complex factors and not only on empirical content.

[6] There are some definite suggestions of crucial tests between the GRW theory and standard quantum mechanics which bear on the implications of the different rates of GRW collapses and decoherence (see Adler 2005, Adler *et al.* 2005, Bassi *et al.* 2005, Hemmo and Shenker 2005).

[7] By this we include GRW–type collapse theories, hidden variables theories such as Bohm's and modal theories, and also many worlds theories.

information–theoretic conditions, but also in the predictions of quantum mechanics, given the general validity of the CBH conditions, which up to now remain *un*observed. As to Thesis III (above), we shall argue by explicit construction that the informationtheoretic approach must be supplemented by further principles over and above those suggested by the CBH theorem. Also, we shall argue that the notion of quantum information cannot be taken as a primitive but rather requires (as in any other quantum theory) a quantum mechanical analysis of measurement of the kind suggested by constructive theories. Finally, we shall argue that a *complete* and consistent information-theoretic approach is bound to rely on *constructive* models, mainly because the measurement problem in quantum mechanics can only be solved by constructive theories; it cannot be otherwise bypassed by theories that treat measurements as “black boxes”.

The paper is structured as follows. We briefly review in Section 2 the CBH theorem and the purported philosophical significance Bub attaches to it. In Section 3 we explain how the GRW collapse theory bears on the CBH information–theoretic principles, focusing (Section 3.1) in particular on the NO BIT principle that Bub sees as constraining any constructive model for quantum phenomena. In Section 3.2 we argue for thesis II above, and in particular that the issue at stake is different explanations of as yet *un*observed quantum predictions. In Section 4, we focus on the empirical inequivalence between the constructive GRW theory and information–theoretic approaches: we first challenge the latter with a thought experiment that establishes our thesis III above (Section 4.1); and we consider various possible replies to our argument in Section 4.2. Finally, in Section 5 we consider another often stated argument against collapse theories and explain why this argument is unacceptable.

2 The CBH Theorem and Its Philosophical Significance

The question raised by CBH is whether we can deduce the kinematic aspects of the quantum–theoretic description of physical systems from the assumption that we live in a world in which there are certain constraints on the acquisition, representation, and communication of information. CBH answered this question positively, supplying three information–theoretic principles (so-called three *no-go's*) that are supposed to filter out the algebraic structure of operators and states that characterize (what they take to be) quantum theory from the more basic structure of C^*-algebras.

The first principle, called *no signaling*, prohibits superluminal transfer of information between spacelike separated systems by carrying out measurements on one of them. In other words, no signaling says that measurements (and in fact any physical manipulation) confined to a remote system cannot possibly change the statistics of the outcomes of measurements that might be carried out on the local system. If Alice and Bob are two physically distinct systems,[8] then when both perform local measurements, Alice's measurements can have no influence on the statistics of the outcomes of Bob's measurements, and conversely. This result follows from the *no signaling* theorem in quantum mechanics according to which local measurements on a system α have no effect whatsoever on the reduced state of a remote system β no matter what the quantum state of $\alpha + \beta$ is. (see Ghirardi *et al.* 1980).

The second principle, called *no broadcasting*, prohibits perfectly broadcasting the information contained in an unknown physical state.[9] No broadcasting ensures that the individual algebras $\mathfrak{A}$ and $\mathfrak{B}$ of the two distinct physical systems are noncommutative. As CBH show, broadcasting and cloning are always possible for classical systems, i.e., in a commutative C^*–algebra there is a universal broadcasting map that clones any pair of input pure states and broadcasts any pair of input mixed states. Conversely, they show that if any two states can be (perfectly) broadcast, then any two pure states can be cloned; and if two pure states of a C^*-algebra can be cloned, then they must be orthogonal. So, if any two states can be broadcast, then all pure states are orthogonal, from which it follows that the algebra is commutative.

---

[8] Consider a composite quantum system A+B, consisting of two subsystems, A and B. For simplicity, assume the systems are identical, so their C^*–algebras $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic. The observables of the component systems A and B are represented by the self-adjoint elements of $\mathfrak{A}$ and $\mathfrak{B}$, respectively. Let $\mathfrak{A} \vee \mathfrak{B}$ denote the C^*–algebra generated by $\mathfrak{A}$ and $\mathfrak{B}$. To capture the idea that A and B are *physically distinct* systems, we assume (as a necessary condition) that any state of $\mathfrak{A}$ is compatible with any state of $\mathfrak{B}$, i.e., for any state ρ_A of $\mathfrak{A}$ and ρ_B of $\mathfrak{B}$, there is a state ρ of $\mathfrak{A} \vee \mathfrak{B}$ such that $\rho|_{\mathfrak{A}} = \rho_A$ and $\rho|_{\mathfrak{B}} = \rho_B$.

[9] In fact, for pure states, broadcasting reduces to cloning. In cloning, a ready state σ of a system B and the state to be cloned ρ of system A are transformed into two copies of ρ. In broadcasting, a ready state σ of B and the state to be broadcast ρ of A are transformed to a new state ω of A+B, where the marginal states of ω with respect to both A and B are ρ.

In elementary quantum mechanics, on the other hand, neither cloning nor broadcasting is possible in general.

These two principles capture two well known features of quantum theory: for a composite system A+B, the no signaling constraint entails that the C^*–algebras $\mathfrak{A}$ and $\mathfrak{B}$, whose self-adjoint elements represent the observables of A and B, commute with each other (this feature is sometimes called 'micro-causality'); and the no broadcasting constraint entails that each of the algebras $\mathfrak{A}$ and $\mathfrak{B}$ is noncommutative. The quantum mechanical phenomenon of interference is the physical manifestation of the noncommutativity of quantum observables or, equivalently, the superposition of quantum states.

The third NO BIT principle prohibits communicating information in a way that implements a given 'bit commitment' with *unconditional* security.[10] In quantum mechanics Alice may send Bob, as a warrant of her bit commitment, one of *two* mixtures associated with the same density operator (where the mixtures correspond to alternative commitments). However, Alice may prepare in advance a suitable entangled state, where the reduced density operator for Bob is the same as that of the mixture she sent him. In this case Alice would be able to steer Bob's system nonlocally into either one of the two mixtures (where Bob cannot be aware of this). So if there are no restrictions on the entangled states that Alice may prepare, Alice can always cheat Bob by pretending to have a secure bit commitment.

The NO BIT constraint prohibits unconditionally secure bit commitment by stipulating that there are no restrictions on the preparation and stability of entangled nonlocal states. Note that the structure that the first two principles filter out from the general C^*–algebra still includes noncommutative theories which are compatible with unconditionally secure bit commitment. In such theories, it might be, for example, that although some nonlocal entangled states (i. e., which permit remote steering) are physically possible, they turn out to be highly unstable (over time) by the dynamical equations of motion (say, given a non-unitary dynamics) and therefore not feasible.[11] So one has

---

[10] Bit commitment is a cryptographic protocol in which, say, Alice sends an encoded bit to Bob as a record of her commitment to either 0 or 1, which allows Bob to ascertain Alice's bit commitment later (only with further information supplied by Alice) so as to make sure that her initial commitment hasn't changed. In classical information theory unconditionally secure bit commitment is always possible in principle.

[11] As noted by Bub, such a possibility in which an entangled state of a composite system quickly decays to a mixture as soon as the component systems spatially separate was raised

to stipulate the feasibility or dynamical *stability* of such states, and this is what the NO BIT constraint does over and above the other two conditions of no signaling and no broadcasting.[12]

Taking stock, the content of the CBH theorem, according to Bub (2005), is this:

> . . . [Q]uantum theories—theories where (i) the observables of the theory are represented by the self-adjoint operators in a noncommutative C^*–algebra (but the algebras of observables of distinct systems commute), (ii) the states of the theory are represented by C^*–algebraic states (positive normalized linear functionals on the C^*–algebra), and spacelike separated systems can be prepared in entangled states that allow remote steering, and (iii) dynamical changes are represented by completely positive linear maps—are characterized by the three information–theoretic 'no-go's': no superluminal communication of information via measurement, no (perfect) broadcasting, and no (unconditionally secure) bit commitment.

In order to flesh out the philosophical significance of the CBH theorem, Bub (2005) makes use of the famous distinction between theories of principle and constructive theories. According to this distinction (which is attributed usually to Einstein although it already appears in the writings of Maxwell and Poincaré),

> [constructive theories] attempt to build up a picture of the more complex phenomena out of the materials of the relatively simple

---

by Schrödinger in 1936.

[12] Timpson (2004, Ch. 9) argues that the NO BIT condition follows logically from no signaling and no broadcasting, and is therefore redundant. However, in theories with non–unitary dynamics (e. g., Schrödinger–type theories) some pure states in the Hilbert space of a system (e. g., superpositions of position states corresponding to spatially separated systems), although permissible, turn out to be dynamically unstable in the sense that they decay extremely quickly by the equations of motion into mixed states that correspond to classical mixtures. In this sense the NO BIT condition in a Schrödinger–type theory (that satisfies no signaling and no broadcasting) might turn out to be unstably sustained for some states. An example of a Schrödinger–type theory of this kind is the GRW theory. We shall argue for our conjecture that the GRW theory does violate the NO BIT condition in the above dynamical sense and that this is the sense relevant to Bub's analysis in Section 3.2, and footnote 19.

> formal scheme from which they start out. Thus the kinetic theory of gases seeks to reduce mechanical, thermal and diffusional processes to the movement of molecules—i.e., to build them up of the hypothesis of molecular motion. [Principle theories, on the other hand,] ...employ the analytic, not the synthetic method. The elements which form their basis and starting point are not hypothetically construed but empirically discovered ones, general characteristics of natural processes, principles that give rise to mathematically formulated criteria which the separate processes or the theoretical representations of them have to satisfy. Thus the science of thermodynamics seeks by analytical means to deduce necessary conditions which separate events have to satisfy, from the universally experienced fact that perpetual motion is impossible (Einstein 1919).

In his analysis of quantum mechanics as a principle theory, Bub appeals to two different historical analogies where scientific progress has been clearly achieved. In his (2004) he considers the transition from the constructive ether-theory of Lorentz–FitzGerald to the abstract geometric formalism of Minkowski's spacetime and argues that the transition was only made possible by Einstein's principle approach to special relativity. And in (2005) Bub focuses on the transition (in the 'opposite' direction) from thermodynamics (as a sort of a principle theory) to the constructive theory of statistical mechanics (in the special case of the kinetic-molecular theory). In both cases Bub's historical analysis seems to be plausible (but these issues are under dispute; compare Brown 2003, 2006), but we are doubtful as to the conclusions he draws about quantum mechanics. Bub argues that the CBH theorem plays the same role in a principle approach to quantum mechanics as the one played by Einstein's principle approach to relativity theory. Focusing on Bohmian mechanics as a constructive mechanical model of quantum mechanics, Bub's argument consists of essentially three elements: First, in special relativity the structure of spacetime is understood in terms of a new primitive—i. e., a *field*—which is not reducible to mechanical motion (e. g., of particles relative to the ether as in the Lorentz theory). Similarly, in quantum mechanics, the algebraic structure of observables is to be understood in terms of a new primitive, i. e., *quantum information*, not reducible to the behavior of mechanical systems (e. g., particle trajectories in configuration space). Second, in both cases the principle approaches are simpler and more fruitful. In the case of quantum theory the CBH theorem is taken to explain *away* (using the above constrains on information flow) some problematic notions in Bohmian mechanics such as sourceless fields that guide the trajectories of particles in configuration space (that sometimes even result in e. g., surreal trajectories and contextual probabilities).[13] At the same time the information–theoretic approach brings out new implications of quantum mechanics such as the use of entanglement as a possible new physical resource for quantum computation. Third, and most crucially in the present context, the constructive mechanical alternatives to quantum theory are *empirically* indistinguishable from the principle theory.

The point of Bub's second analogy (i. e., the transition from thermodynamics to the kinetic-molecular theory) is precisely to bring out the immense importance of empirical distinguishability in theory choice. Here the argument is that the kinetic-molecular theory would not be regarded seriously as an alternative constructive model for thermodynamics if (contrary to fact) it had no new empirical predictions that differ from those of thermodynamics, and if those predictions were not experimentally confirmed (e. g., Einstein's prediction of fluctuations in Brownian motion). By contrast, in the case of Bohmian mechanics it is provable that (i) once the distribution of the particles' positions is given by the square of the amplitude of the wavefunction at each point (i. e., by Born-like probabilities) this distribution is preserved at all later times by the dynamical equations of motion (see Dürr, Goldstein and Zanghi 1992); and (ii) given the Born distribution, Bohm's theory is empirically equivalent to quantum mechanics. This means, in the information–theoretic approach, that Bohm's theory must be equivalent to quantum theory if the CBH constrains on the information flow were satisfied even once in the past (since, roughly, in a no collapse theory these constrains hold if and only if the Born probability distribution holds). Consequently, Bohm's theory, quite unlike the case of Brownian motion, can yield no predictions of 'fluctuations' that deviate from the predictions of quantum mechanics. And therefore, given Bub's arguments above, the rational epistemological stance, is to reject it, and prefer the principle information–theoretic approach.

In what follows we question Bub's reading of the present state of affairs in quantum mechanics. Although we largely agree with Bub's analysis (sketched

---

[13] Here we try to flesh out Bub's preference for a principle approach to quantum mechanics over a constructive theory such as Bohm's. Obviously, supporters of Bohm's theory judge surreal trajectories and contextual probabilities as *un*problematic.

above) of the features of alternative hidden variables theories, we think that the analysis doesn't capture all the relevant aspects related to theory choice in the case of the present state of quantum mechanics. In particular, there are two crucial points that seem to us not appropriately addressed in Bub's analysis. First, there are *other* constructive quantum mechanical theories (in particular the GRW *collapse* theory) which generally violate the NO BIT constraint (as a dynamically stable constraint; see Section 3.2). The GRW theory differs in its empirical predictions from quantum theory, while it is perfectly compatible with our experience so far. By Bub's own standards (see Bub 2005, Sec. 4), therefore, the GRW theory is acceptable as an alternative constructive theory. But adhering to Bub's principle approach would result in loosing sight of theories like the GRW theory on what seems like an *a–priori* rather than an empirical basis. Second, information–theoretic approaches (i. e., both Bub's principle approach and the Bayesian approach) are *incomplete*, and as we said (in our thesis (III) above) need be supplemented by further principles that are quite hard to justify (see Section 4).

On the basis of these two points we now proceed to argue for our thesis (II), namely that constructive theories are necessary to solve the measurement problem in standard quantum mechanics, and that at least partially the issue at stake lies not in the information–theoretic description of the *observed* quantum phenomena, but rather in the explanation of the predictions of quantum theory which up to now remain *un*observed.

3 Explaining the *Un*observed

Standard no collapse quantum theory predicts the unrestrictive existence of superpositions in spatially separated entangled states. This is tantamount to saying, using the CBH theorem, that *ex hypothesis* the NO BIT principle holds in our world. But then, given this hypothesis, one question that arises naturally is: why do superpositions remain *un*observed in macroscopic massive physical systems?

This question is in fact a variant on the so-called measurement problem in the quantum theory of measurement. In this context the problem arises as a straightforward consequence of applying the Schrödinger linear and deterministic dynamics to the measurement interaction. As is well known the Schrödinger dynamics results for a generic measurement in a superposition of the form

$$|\Psi\rangle = \sum_i \mu_i |\psi_i\rangle \otimes |\varphi_i\rangle, \qquad (1)$$

where the kets $|\psi_i\rangle$ represent some suitably defined pointer states of the measuring apparatus (typically, the $|\psi_i\rangle$ are eigenstates of the pointer position), and the $|\varphi_i\rangle$ are some states of the system. The problem is that in states of the form (1) the measurement has no definite outcome (except in the special case where all but one of the μ_i are zero), since the final (reduced) state of the apparatus cannot in general be described in terms of an ensemble of systems in a classical mixture (in which the $|\mu_i|^2$ represent the probabilities for each $|\psi_i\rangle$ to actually occur).

The information–theoretic approach addresses this problem by appealing to models of decoherence in which the interaction of relatively massive systems with their environment brings about a so-called *effective* collapse onto the eigenstates of some preferred observables (typically, position).[14] According to this approach (called by Bub 1997 some years ago the 'new orthodoxy') macroscopic entangled states in position exist, but for all practical purposes they are unobserved because we have no control over the states of the environment, so that the reduced state of a decohering system is practically indistinguishable from a classical mixture. That this is no solution to the measurement problem was argued in the past by many,[15] including Bub himself in the context of (constructive) hidden variables theories.[16] However, Bub seems to think that the objection does not apply to his own principle information–theoretic approach. We disagree and will argue for this in Section 4. But before doing so, we wish to consider here the constructive collapse theory by GRW that gives a clear and distinct solution to the measurement problem and explains the *un*observability of macroscopic spatial superpositions in the most straightforward way.

---

[14] For standard models of decoherence, see Joos *et al.* (2003) and references therein.

[15] E.g., Bell (1990), and recently Adler (2003).

[16] As Bub puts it in his (2000, 90–91): the fact that the 'effective' quantum state—an improper mixture described by the reduced density operator (obtained by tracing out the degrees of freedom of the environment)—is diagonal with respect to properties associated with some pointer basis "not only fails to account for the occurrence of just one of these [properties] but is actually inconsistent with such occurrence", since taking into account the environment gives us back the pure state from which the mixture was derived, and this state is inconsistent with the occurrence of events associated with definite properties.

3.1 The GRW theory

The GRW theory (formulated for non–relativistic quantum mechanics) explains the *un*observability of some macroscopic superpositions of *position* states by modifying the Schrödinger linear dynamics in such a way that given the new dynamics such superpositions are overwhelmingly likely to collapse at every moment of time, and in this sense they are highly unstable. The Schrödinger equation is changed by adding to it a non-linear and stochastic term that induces the so-called *jump* or collapse of the wavefunction. The jump is supposed to occur on occasion in position space and its postulated frequency is proportional roughly, to the mass density of the system (or in Bell's (1987) model on the number of particles described by the wavefunction). For our purposes it is enough to sketch Bell's (1987) version of the elementary and non–relativistic theory of GRW. This goes roughly as follows.

Consider the quantum mechanical wavefunction of a composite system consisting of N particles:

$$\psi(t, \mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N). \tag{2}$$

The time evolution of the wavefunction usually (at almost all times) satisfies the deterministic Schrödinger equation. But sometimes *at random* the wavefunction collapses or *jumps*) onto a wavefunction ψ_ℓ localized in position of the (normalized) form

$$\psi_\ell = \frac{j(\mathbf{x} - \mathbf{r}_n)\ \psi(t, \mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)}{R_n(\mathbf{x})}, \tag{3}$$

where $\mathbf{r}_n$ in the jump factor $j(\mathbf{x} - \mathbf{r}_n)$ (which is normalized) is randomly chosen from the arguments $\mathbf{r}_1, ..., \mathbf{r}_n$ of the wavefunction immediately before the jump, and $R_n(\mathbf{x})$ is a suitable renormalization term. For j, GRW suggest the Gaussian:

$$j(\mathbf{x}) = K\ \exp(-\mathbf{x}^2/2\Delta^2), \tag{4}$$

where the width Δ of the Gaussian is supposed to be a new constant of nature: $\Delta \approx 10^{-5}$cm.

Probabilities enter the theory twice. First, the probability that the *collapsed* wavefunction ψ_ℓ after a jump is centered around the point $\mathbf{x}$ is given by

$$\mathrm{d}^3\mathbf{x}\, |R_n(\mathbf{x})|^2\, . \tag{5}$$

This probability distribution, as can be seen, is proportional to the standard quantum mechanical probability given by the Born rule for a position measurement on a system with the wavefunction $\psi(t, \mathbf{r})$ just prior to the jump. Second, the probability in a unit time interval for a GRW jump is

$$\frac{N}{\tau}, \tag{6}$$

where N is the number of arguments in the wavefunction (i. e., in Bell's model it may be interpreted as the number of particles), and τ is, again, a new constant of nature ($\tau \approx 10^{15}\,\mathrm{sec} \approx 10^{8}$ year). Note that the expression (6) does not depend on the quantum wave function, but only on N. This is essentially the whole theory.

For microscopic systems GRW collapses have extremely low probability to occur, and the quantum mechanical Schrödinger equation turns out to be effectively true at almost all times in just the way that no collapse quantum mechanics predicts (and experiment confirms). However, for massive macroscopic systems (e. g., for systems with 10^{23} particles) the GRW collapses are highly probable at all times. In measurement situations the GRW theory implies that superpositions of macroscopically distinguished pointer states of the form (1) collapse with extremely high probability onto the localized states $|\psi_i\rangle$ on time scales that are much faster than measurement times. In particular, the probability that the wavefunction of the composite of system plus apparatus will stay in the superposition (1) for more than a fraction of a second (e. g., by the time the measurement is complete) vanishes exponentially. So the GRW jumps reduce wavefunctions of the form (1) to one of the components $|\psi_i\rangle \otimes |\varphi_i\rangle$, in which the pointer is in the *localized* state $|\psi_i\rangle$, where the probability for a collapse onto the i-th term (see equation (5)) is given as usual by the squared amplitude $|\mu_i|^2$. This means that in a sequence of quantum mechanical measurements the GRW jumps result in definite outcomes with frequencies that are (approximately) equal to the Born-rule probabilities $|\mu_i|^2$. The measurement problem is solved by this construction as long as measurements involve a macroscopic recording of the measurement outcome in *position* (e. g., a moving pointer of a measuring device, particles impinging on macroscopically separated regions of a computer screen, etc.). Note that, and this is important for our discussion below, the GRW jumps are designed to be extremely effective *only* for macroscopic superpositions of *position* states (and to any other states that are *coupled* to positions), but *not* to arbitrary superpositions.

There are well known physical weaknesses in the GRW theory. In the non–relativistic case the problem seems to be how to avoid the accumulated violations of conservation of energy induced by the jumps.[17]. But perhaps the main problem is to write down a relativistic formulation of the GRW theory. Although GRW give the same predictions as standard quantum mechanics with respect to the nonlocal correlations in EPR–Bell–type experiments, the problem is that it is not clear whether in a relativistic version of the GRW dynamics the jumps could be made Lorentz–covariant.[18] Also, it has been argued against GRW that the theory appears to be strongly *ad hoc* as it allows free adjustments of constants in order to be in agreement with experiments. Although we believe that these complaints are largely unjustified, here we set this issue aside, because we are not concerned with the GRW theory *per se*, nor with its comparison with other constructive theories. Rather, our focus is whether Bub's principle approach which seems to *rule out* the GRW theory on the basis of information–theoretic constraints is acceptable.

3.2 Constructing the principles

How do the information–theoretic constraints of the CBH theorem bear on the GRW theory? The first thing to say is that the GRW theory does not violate the first two constraints (i.e., no signaling and no broadcasting) because of its stochastic dynamics (see, e. g., Gisin 1989). But, we conjecture (following Bub 2005) that the GRW theory does violate the NO BIT principle in the following sense. The GRW dynamics is devised in such a way that superpositions of macroscopically distinguishable *position* states decay extremely quickly by the equations of motion into the corresponding classical-like mixtures. Obviously, no superselection rules are introduced by the GRW theory. That is, all the usual quantum mechanical observables are permissible in the theory. And so *any* pure quantum state of any system and *any* entangled and nonlocal state (say, of the EPR–Bell–type) in the Hilbert space of a composite system is permissible too. In particular, one may prepare, for example, a superposition of highly entangled spin states even for a macroscopic system that will be quite stable over a significant time interval (e. g., as in the spin–echo experiments, where the spins get, as it were, in and out of macroscopically entangled states during an appreciable time interval; see

---

[17] See Ghirardi (2000) and references therein; Some progress is reported in Bassi *et al.* (2005b)

[18] But see Ghirardi (2000) and Myrvold (2002).

Hemmo and Shenker 2005). Such states are perfectly compatible with the GRW theory and might even be quite stable over time. Moreover, two massive molecules can be made to enter with certainty a quantum state in which their position degrees of freedom will be highly entangled as in EPR–type states. But, given the GRW dynamics, states of this latter sort would be highly unstable due to the extremely fast rates of the GRW jumps in *position*. More generally, it is an empirical prediction of the GRW theory that superpositions of entangled position states of spatially separated systems are highly unstable if the systems are massive enough. Hence, the theory predicts that observable effects of such superpositions are hard to come by, and under normal circumstances, remain *un*observed.

Therefore, we conjecture, following Bub (2005), that in the GRW theory an unconditionally secure bit commitment could always be made possible in principle (modulo the complexity of the encryption code) via a set up that requires Alice to encode her bit commitment in the position state of a massive enough system (relative to the GRW parameters that fix the probability for spontaneous localization). The NO BIT constraint is violated in the GRW theory in the sense that spatially separated macro–systems cannot be stably entangled in their positions. That is, in the GRW theory it is *not true* that for *any* mixed state that Alice and Bob can in principle prepare by following some (bit commitment) protocol, there exists a corresponding nonlocal entangled state that can be physically and *stably* occupied by Alice's and Bob's systems. That is, it is a straightforward consequence of the GRW dynamics that some nonlocal entangled states are highly unstable, in the sense that these states—although permissible in principle—decay extremely quickly to mixed states that essentially correspond to classical mixtures). Thus we take the NO BIT condition (following Bub 2005) as a *dynamical* condition on the stability of quantum states, hence as filtering alternative *quantum* Schrödinger–type theories such as the GRW theory.[19] This, of course, is not sufficient to entail that in the GRW theory a secure bit commitment protocol is feasible, but the security of Alice's protocol will depend *only* on the computational complexity

---

[19] Timposn (2004, ch. 9) overlooks this point. For the purpose of filtering out only *classical* theories, the NO BIT condition is redundant, since it is implied (as Timpson argues) by no signaling and no broadcasting. But to rule out GRW–type theories further dynamical conditions are required. For this purpose the NO BIT condition seems to be necessary and (together with the other two conditions) sufficient. Perhaps, it could be narrowed down and translated into an empirical constraint (say, in the form of a Bell-like inequality) for each collapse dynamics.

of the encryption code.

However, the introduction of the unconditional NO BIT constraint into quantum mechanics implicitly presupposes that as a matter of fact the entangled states shared between Alice and Bob are states of *micro*systems, say, over spin degrees of freedom. For such states we have ample experimental confirmation that the NO BIT constraint is, indeed, satisfied. But for such states *also* the GRW theory satisfies the NO BIT constraint with extremely high probability. It violates the NO BIT constraint in the above sense only with respect to superpositions of macroscopically distinguishable position states. But for such *macro*states also standard no collapse quantum theory predicts an *effective* violation of the NO BIT constraint in the same sense due to environmental decoherence (i. e., such states effectively collapse also in the standard theory as all models of environmentally induced decoherence show; see e. g., Zurek 1991, Joos *et al.* 2003).

So practically, no collapse quantum mechanics and the GRW theory agree on the NO BIT constraint for all cases in which there are good empirical reasons to believe it is true. And the two theories seem to disagree about the NO BIT constraint only with respect to those macro superpositions about which we do not know whether or not they in fact exist in our world. But of course this is no surprise since this disagreement is located precisely where the two theories differ in their empirical predictions. So we are back to square one! The CBH theorem therefore provides a clear cut principle that distinguishes between the empirical predictions of theories of genuine collapse of the GRW type and theories of effective collapse (as in models of environmental decoherence). But given what we know empirically about the world, there seem to be no grounds for adopting the NO BIT principle as an unrestrictive constraint on theory choice.[20]

4 Inconsistent Predictions?

We now proceed to argue for our thesis III, namely that information–theoretic approaches to quantum mechanics are *incomplete* and need be supplemented by further axioms that exceed information–theoretic principles such as those

---

[20] Note that the point made here is quite general. It would be applicable to *any* information–theoretic characterization of the mathematical structure of quantum theory (e. g., Spekkens 2004) and to any constructive alternative of it, provided the latter is empirically well confirmed and *not* empirically equivalent to standard quantum mechanics.

suggested by the CBH theorem, or by subjectivist Bayesian approaches to quantum mechanics (see e. g., Fuchs and Peres 2000, Fuchs 2002, Caves *et al.* 2002, Spekkens 2004, and references therein). For convenience, we shall frame our discussion in the context of the Bayesian approach. After presenting our argument in the form of a thought experiment, we shall make the link to Bub's principle approach.

The Bayesian approach to quantum theory is based on an epistemic attitude according to which the quantum state does not represent a real physical state of a system, but instead supplies an observer with statistical information concerning all possible distributions of measurement results. The probabilities computed by the standard Born rule are understood as probabilities of *finding* the system on measurement in some specific state. Applying von Neumann's projection postulate to the quantum state (or more generally applying Lüder's rule), under this account, is just an adjustment of subjective probabilities, conditionalizing on newly discovered results of measurement, i. e., it is merely a change in the observer's knowledge, or probability assignments. By contrast, the unitary and linear quantum mechanical dynamics (i. e., the Schrödinger equation in the non-relativistic case) describes the observer–independent and in this sense objective time evolution of the quantum probabilities when no measurement takes place. Hence, in this approach measurements can be treated operationally as 'black boxes' and require no further theoretical analysis.

4.1 A thought experiment

We now turn to our thought experiment which is a variant on Wigner's friend. Consider the following set–up in which an observer A measures the z–spin of a spin-half particle P by means of a Stern–Gerlach apparatus (which, to keep things simple, we omit from our description below). The quantum state of $P + A$ initially is

$$|\Psi_0\rangle = (\alpha|+_z\rangle + \beta|-_z\rangle)|\psi_0\rangle_A \tag{7}$$

where $|\alpha|^2 + |\beta|^2 = 1$ $(\alpha, \beta \neq 0)$, the kets $|\pm_z\rangle$ are the $z-$spin eigenstates and $|\psi_0\rangle$ is the initial *ready* state of A. After the measurement, in a no collapse theory, the quantum mechanical state of $P + A$ is the superposition:

$$|\Psi_1\rangle = \alpha|+_z\rangle|\text{see up}\rangle_A + \beta|-_z\rangle|\text{see down}\rangle_A \tag{8}$$

where $|\text{see up}\rangle_A$ and $|\text{see down}\rangle_A$ are, say, the brain states of A corresponding to her perceptions and memories of the two possible outcomes of the measurement. By contrast, in a collapse theory of the GRW kind, the state (8) is highly unstable (assuming that the chain of interactions leading to A's different memory states involves macroscopically distinguishable position states), so that by the time the measurement is complete this state collapses onto one of its components.

Consider now an observable $\hat{O}$ of the composite system $P+A$ of which the state (8) is an eigenstate with some definite eigenvalue, say $+1$.[21] Suppose that the composite system $P+A$ is completely isolated from the environment, and that a measurement of $\hat{O}$ is about to be carried out on $P+A$ immediately after the state (8) obtains. According to no collapse quantum mechanics the measurement of $\hat{O}$, under these circumstances, is completely non-disturbing in the sense that after the measurement the state of $P+A$ remains precisely as in (8). One may think of $\hat{O}$ as an observable that is maximally sensitive to whether or not the interference terms between the different components of (8) exist. In other words, the measurement of $\hat{O}$ on $P+A$ if the state (8) is the true state of $P+A$ is a non-demolition measurement that, as it were, passively *verifies* whether or not $P+A$ is in fact in that state.

Note that $\hat{O}$ commutes neither with the $z-$spin nor with A's perceptions and memories of the outcomes of the $z-$spin measurement. This surely raises interesting questions about the status of the uncertainty relations in this set up and about the reliability of A's memories of the outcome of her spin measurement in the event that A measures $\hat{O}$ just after her spin measurement. However, no matter what happens during the measurement of $\hat{O}$ (to A's memory of the outcome of her spin measurement, or to the $z-$spin values themselves) quantum mechanics implies that the *correlations* between the $z-$spin of P and A's memories must remain exactly the same as they were before the $\hat{O}$–measurement. Moreover, in a no collapse theory, the state of $P+A$ immediately after the $\hat{O}$–measurement will be, with complete certainty, just:

$$|\Psi_2\rangle = \alpha|+_z\rangle|\text{see up}\rangle|\text{see }\hat{O}=+1\rangle + \beta|-_z\rangle|\text{see down}\rangle|\text{see }\hat{O}=+1\rangle \qquad (9)$$

where $|\text{see }\hat{O}=+1\rangle$ is the state corresponding to perceiving the result of

---

[21] Observables like $\hat{O}$ are defined in the tensor product Hilbert space $\mathcal{H}_P \otimes \mathcal{H}_A$ unless superselection rules are introduced. For our purposes think of $\hat{O}$ as an observable that pertains to P's spin degree of freedom and the relevant degrees of freedom of A's sense organs, perceptions, memory, etc.

the $\hat{O}$–measurement.

By contrast, in a collapse theory of the GRW kind, the state of $P + A$ *immediately* after the $\hat{O}$–measurement will be given by *one* of the eigenstates of $\hat{O}$, where the probability that it will be state (8), and therefore that the outcome $\hat{O} = +1$ will be obtained, is only $|\alpha|^2$. Note that even if that outcome will obtain, the state (8) will extremely quickly collapse, again, onto one of the components of state (8) with probabilities that are given by $|\alpha|^2$ and $|\beta|^2$. So, in the GRW theory, the final value of the $z-$spin and the spin memory of A might be different before and after the $\hat{O}$–measurement.

Note further that we deliberately do not specify here *who* carries out the $\hat{O}$–measurement (i. e., in which degree of freedom the outcome $\hat{O} = +1$ is recorded). It may be carried out by A or by some other observer B external to A's laboratory. As can be seen from our notation in (9), we have assumed that the outcomes of A's spin measurement and of the $\hat{O}$–measurement are recorded in *separate* degrees of freedom. Quantum mechanics imposes no restrictions whatsoever on the way in which the outcomes of these measurements are recorded, except that they cannot be recorded *simultaneously* in the same degree of freedom (since $\sigma_z \otimes \mathbf{1}$ as well as A's memory observable are incompatible with $\hat{O}$). No further restrictions are imposed by quantum mechanics (with or without collapse) on the *identity* of the observers who may carry out $\hat{O}$–type measurements (we return to this point below).[22]

However, consider what would happen in our scenario if A were to communicate her outcome to B. As long as B does not know the outcome of A's spin measurement, B's predictions for the $\hat{O}$ measurement will conform to option (b) (but see below). So if A and B don't communicate, B's predictions are *not* ambiguous (unlike A's predictions). What about A's predictions? Once A communicates her spin outcome to B, A's predictions about the $\hat{O}$–measurement would no longer be ambiguous, since the reduced state of $P+A$ after A's communication with B would be given by the mixture corresponding to the state (8) (and the total state of $P+A+B$ would no longer be an eigenstate of $\hat{O}$). In fact, the predictions of A and B would coincide in this case; that is, both A and B would predict probability of 1/2 to the outcome $\hat{O} = +1$. However, quantum mechanics entails the existence of *other* $\hat{O}$–type observables (e. g., any observable $\hat{O}_2$ of which the state of $P+A+B$

---

[22] See Albert (1983) for an extended discussion of $\hat{O}$–type measurements and their implications. Aharonov and Albert (1981) use $\hat{O}$–type measurements in their discussion of the collapse of the quantum state in a relativistic setting.

immediately after the interaction of A and B is an eigenstate) with respect to which we can run the *same* argument again, this time on the predictions of both A *and* B. And likewise *ad infinitum.* But now the crucial point is that whether or not A communicates her outcome to B becomes irrelevant. Suppose that A doesn't. Then, A and B differ in their predictions about the outcome of the $\hat{O}$–measurement. And so one of them at least must be wrong, or so it seems.

To make things simple, let us suppose that the $\hat{O}$–measurement is to be carried out by the external observer B, but let us consider A's *predictions* of the probabilities of the outcomes of the $\hat{O}$–measurement.[23] Here we encounter a problem in the information–theoretic approach since it doesn't tell us on what quantum state A should base her predictions. In order to calculate her expected probabilities A may choose one of the following two options.

(a) Update her quantum state in accordance with the outcome of the spin measurement she actually observed, either $|\text{see up}\rangle_A$ or $|\text{see down}\rangle_A$. In this case, she would collapse the state (8) onto one of its spin+memory components. Applying the Born rule to this state, she will predict that the result of the $\hat{O}$–measurement will be $+1$ with probability $|\alpha|^2$.

(b) Ignore the outcome of the spin measurement she actually observed, and conditionalize her probabilities on the *uncollapsed* state as in (8). In this case, since the state in (8) is an eigenstate of $\hat{O}$ with eigenvalue $+1$, she will predict that the result of the $\hat{O}$–measurement will be $+1$ with *certainty* (i. e., probability 1).

But by actually performimg a series of repeated $\hat{O}$–measurements on identically prepared systems, all in state (8), we can in principle distinguish experimentally between the two predictions. So, on pain of inconsisteny, no theory can accept both predictions as true.

The point to be made here, however, is that the information–theoretic approach gives no plausible account of *which* option, (a) or (b), is the correct one. On the one hand, the full information about the lab available to A before the $\hat{O}$–measurement is given by the *collapsed* state, and this implies that option (a) is true. On the other hand if the dynamics of quantum states is invariably given by the Schrödinger equation, then (assuming that this

---

[23] Clearly, in quantum mechanics the quantum state assigned to a system is supposed to give the probabilities of the outcomes for *all* possible measurements.

information is admissible to A), her predictions ought to be guided by the *uncollapsed* state as in (8). And so this means that option (b) is true. And so we seem to have a straightforwrd inconsistency.

We have formulated the above argument in terms of the Bayesian information–theoretic approach. The link to Bub's principle approach goes as follows. In both approaches quantum mechanical measurements are construed *operationally* as 'black boxes' with no further analysis. This is precisely the sense intended by Bub (2005, Sec. 4) of taking quantum information as a primitive and irreducible physical concept. As argued by Bub, once we accept the three constraints suggested by the CBH theorem, it follows that measurements are to be treated operationally and measuring apparatuses as black boxes. But, any 'black box' theory of measurement is bound to be incomplete in accounting for Wigner'stype scenarios as spelled out above. Moreover, as long as the mathematical formalism of quantum mechanics remains intact, any operational approach that is not committed to either a collapse–type theory or a no–collapse theory of a sort (on this point, see also the next subsection) will necessarily run into circumstances in which the assignment of quantum states will become ambiguous, and as we just argued this ambiguity will lead to inconsistent probability assignments to measurement outcomes! If quantum theory is, indeed, on the right track (e.g., if there no unknown superselection rules), then the decision between collapse and no–collapse dynamics is forced upon us by *empirical* constraints. In fact, we take this simply to mean that the measurement problem in quantum mechanics by *constructive* theories[24]

4.2 Some resolutions

In what follows we examine various possible objections to our criticism (above).

(I) $\hat{O}$–type measurements are not feasible. As a matter of fact, the $\hat{O}$ measurement is physically impossible to carry out due to decoherence and the complexity of the set–up under consideration. As far as our best neurophysiological theories tell us, the sensory apparatus and brain processes of a human observer involved in typical perception and memory processes are

---

[24] This includes also the many worlds theory in all its variants. Of course, one may interpret the NO BIT condition as a dynamical constraint, which *a forteriori*, implies a no–collapse theory. But then the issue of a principle approach vs. a constructive one is off the table. See also footnote 18.

macroscopic and subject to continuous decoherence (induced by interactions inand outside of the brain) that cannot be screened off. Consequently, $\hat{O}$ will become extremely fast the wrong observable to measure in order to detect the interference terms in states of the form of (8). Moreover, even if one could identify the right observable to measure at a given time we cannot expect to have control over all the relevant degrees of freedom in and outside the observer's brain.

Reply (I) Obviously, $\hat{O}$–type measurements are extremely hard to carry out and need be continuously protected against decoherence. This is much beyond our experimental reach. In fact, even if we set decoherence aside, $\hat{O}$–type measurements are not quite feasible in microscopic superpositions over, say, only spin degrees of freedom (since we need to measure total spin without measuring the spin components separately in order for the measurement to be non-disturbing). But ways to overcome such problems may be found. For example, in spin–echo experiments we know today by means of macroscopic manipulations only how to screen–off the effects of decoherence for appreciable time intervals (see Hemmo and Shenker 2005). Moreover, feasibility considerations are quite beside the point in the present context. Quantum mechanics allows $\hat{O}$–type measurements, and the above ambiguity must be resolved independently of whether we can or cannot practically translate it into an experimental context. Obviously, it may turn out that, on the basis of some new physics, e.g., quantum gravity, the $\hat{O}$–measurement would be impossible *in principle* on pain of violating certain new physical laws, but as far as standard quantum theory is concerned, this is not the case.

(II) Option (b) is wrong. A should stick to option (a) and use her spin–memory eigenstate in order to calculate her probabilities (likewise for the external observer B). By construction, the above set–up stipulates that A carried out a *measurement*, and therefore from A's point of view the state in (8) has collapsed. As to B, although he doesn't know which outcome A has observed in the lab, he knows that she has carried out a measurement. Therefore, he ought to condition his predictions on the mixture of components in (8) rather than the full superposition. On this view the predictions of A and B for the $\hat{O}$ measurement coincide: they both give probability of one–half to $\hat{O} = +1$ despite the fact that their predictions are conditional on different quantum states.

Reply (II) The above argument is wrong–headed in the context of the information–theoretic approach, since it seems to imply that a measurement induces a stochastic transition from a pure to a mixed state. This, however, is exactly what's argued by collapse theories which aim at *constructing* a dynamical theory in which such transitions can be accounted for. To put matters differently, option (b) cannot in general be rejected unless one is committed to a genuine collapse theory. If the mathematical fomalism of quantum mechanics is not changed (e. g., by adding some suitable superselection rules), then given the standard ways of thinking about it, it is a fact that no collapse theories that employ dynamics which only induces effective collapses (somehow within the global envelope of the unitary dynamics of the quantum state) prescribe different predictions for $\hat{O}$–type measurements than those of genuine collapse theories, precisley because of the difference in their dynamical laws for the quantum state.

(III) Options (a) and (b) are compatible. Which option, (a) or (b), will turn out to give the right predictions for the $\hat{O}$–measurement depends on *who* carries out the measurement, A or B. The outcome $+1$ of the $\hat{O}$–measurement is assigned probability 1 only if A carries out the measurement, but probability $1/2$ if B carries out the measurement. There is no inconsistency, because the different probabilities are assigned on the basis of incompatible quantum states (i.e. eigenstates of incompatible—maximally non-commuting—observables) which cannot be identified with a single system (or a single point of view). In order for the $\hat{O}$–measurement to pick up the interference terms, and thus yield the $+1$ outcome with probability 1 in accordance with B's prediction, the apparatus must be kept completely isolated from the content of A's laboratory. If, before the $\hat{O}$–measurement but just after A's spin measurement, the apparatus gets coupled to even a single air molecule or a photon that interacted with the z–spin measuring device or with A's relevant degrees of freedom, its initial quantum state will 'split', and following the $\hat{O}$–measurement it will split again relative to each component of the state (8) exhibiting both ± 1 outcomes (i. e., the ± 1 outcomes will be probabilistic). It is highly plausible due to standard decoherence effects (with a realistic Hamiltonian) that if A carries out the measurement herself she will have to interact with the $\hat{O}$–apparatus, if only to switch it on, in just this way. Therefore, there is a fundamental distinction between A–systems, i. e., systems that are coupled to the content of A's laboratory, and B–systems, i. e., systems that are not so coupled, and the probability distribution of the outcomes of the $\hat{O}$–measurement depends on which kind of system actually carries out the measurement.[25]

Reply (III) The above argument requires a constructive notion of an observer or of measurement in order to make sense of the fundamental distinction between the A– and the B– systems. In this sense it does not fit a "black-box" treatment of measurement. Here are some problems that such a constructive view will have to resolve. First, it is not at all clear what is it that makes the $A-B$ distinction so fundamental? In other words, How can a physically meaningful distinction between A– and B– systems be made on the basis of quantum mechanics as we know it, or some other known physical principles? For example, A might prepare *in advance* the $\hat{O}$–apparatus in its ready state, and we can just assume that the $\hat{O}$–measurement immediately and extremely quickly follows the spin measurement. In that case, A would be acting as a B–system, but she would also *know* the outcome of the spin measurement. Second, why can't our scenario be compatible with the case in which A stores her memory of the spin outcome in some perfectly *isolated* degree of freedom in her brain (i. e., her information is completely private!), so that when performing the $\hat{O}$–measurement she is actually acting as a B–system? It doesn't seem to us that physical laws as we know them dictate that A's information about the $\hat{O}$–outcome need be public in the sense that leads to decoherence. But if no restriction of this kind can be imposed here, then the inconsistency isn't resolved: what would be in this case A's predictions about an upcoming $\hat{O}$–measurement?

(IV) Option (a) is wrong. Only the full quantum state in (8) is the right one for conditionalization (for both A and B), because the $\hat{O}$–measurement involves interference between the two components of the state in (8). A's information about the outcome of the spin measurement becomes completely irrelevant due to the *objective* (observer-independent) features of the dynamics of the quantum state involved in the $\hat{O}$–measurement itself. One might even argue that this is the 'flipped-side' of the uncertainty relations: since the value of $\hat{O}$ is known in advance with certainty, A's information about the spin outcome is unreliable, and this is manifest by the fact that A's 'memory' of the spin value (i. e., her record observable) and $\hat{O}$ are maximally non–commuting. Moreover, as suggested by Bub (2005, Sec. 4), the emergence of classical information is explained only by decoherence. But,

---

[25] Note that this dependence of the actual probability distribution of the $\hat{O}$–measurement on the A– or B– identity is absolutely crucial. Otherwise, the probability assignments will be straightforwardly inconsistent, as we argue above.

by construction, the $\hat{O}$–measurement requires *recoherence* of the components in (8). Therefore, whatever information these components carry ought to be disregarded (by both A and B) in the face of the $\hat{O}$–measurement. Hence option (a) is no option, and the above inconsistency is resolved.

Reply (IV) This argument, again, implicitly relies on some sort of a constructive theory of measurement. Consider the claim that A's memory of her spin outcome is *unreliable* in the face of the $\hat{O}$–measurement. In so far as standard (operational) quantum mechanics goes, we are faced with a situation in which the measurement of spin on P is followed by a measurement of $\hat{O}$ on $P+M$. Since $\sigma_z \otimes \mathbf{1}$ and $\hat{O}$ are (maximally) non–commuting, by the uncertainty relations, the outcomes of the spin measurement only determine the probabilities of the outcomes of the $\hat{O}$–measurement (in accordance with the Born rule). That's about all we can say in an operational "black box" description of measurement. Whether or not A's memories are reliable *during* the measurement of $\hat{O}$ is a question that cannot be settled by a quantum theory that treats measurements as "black boxes" nor by the CBH constrains on information transfer. It cannot even be settled by appealing to the fact that perceptions and memories need be modelled by decohering systems. The above argument requires additional laws (over and above the laws of quantum mechanics) about the dynamical behavior of A's memory during the interference involved in the $\hat{O}$–measurement.[26] But, strictly speaking, such laws are nothing but laws about the behavior of hidden (or extra) variables of a sort. And so accepting (IV) above presupposes that quantum mechanics is *incomplete.*

In the above discussion we've tried to counter various objections to the $\hat{O}$–scenario. It is sometimes claimed that such a scenario is, for some reason,

---

[26] In Bohm's theory, for example, A's memory of the spin is perfectly reliable since the trajectories given by Bohm's deterministic guidance equation cannot cross each other. By contrast, in modal interpretations where the dynamics of the extra values is stochastic, A's memory of the spin might flip during the $\hat{O}$–measurement. Anyway, on either theory the analysis of $\hat{O}$–type measurements is a straightforward physical analysis, and neither says that such measurements are *un*physical or cannot be carried out, or what have you. There is a subtle issue concerning $\hat{O}$–type measurements in decoherence-based many worlds theories, where the branching is defined by environmental decoherence, since $\hat{O}$–type measurements involve recoherence of the branches associated with A's memories of her spin measurement. So it is not clear whether in our scenario the branches associated with A's memories are well-defined, and therefore whether or how these memories might play a role in A's predictions (e. g., of the outcome of the $\hat{O}$–measurement). But we go no further on this issue here.

excluded *as a matter of principle* by quantum mechanics. For example, it is claimed that a single observer (say A) cannot carry out both the spin measurement and the $\hat{O}$–measurement, since $\hat{O}$ is an observable of $P + A$, and therefore measuring it by A entails some problematic form of *self-reference*. It is also claimed that because $\hat{O}$ and A's spin memories (as we said above) don't commute, no single observer can ever be in a position to know (with complete certainty) the values of both σ_z and $\hat{O}$ (as manifested by the uncertainty relations). So if A knows the value of σ_z (as we assume) she cannot possibly know simultaneously also the value of $\hat{O}$. Both claims seems to take it that quantum theory itself imposes some sort of physical constraints as to the identity of the observers who may carry out $\hat{O}$–type measurements, or that $\hat{O}$–type measurements are for some reason meaningless. Therefore, the inconsistency of A's predictions is somehow unphysical or cannot in principle be revealed by a single experiment.[27]

However, these claims and various variants thereof are completely off the mark in the context of our thought experiment. First, the question of whether or not the $\hat{O}$-scenario involves self–reference of any sort is *irrelevant*, since the issue boils down in its entirety (as we presented it above) to a straightforward question about the *predictions* of A as to the statistics of the outcomes of the $\hat{O}$–measurement—*no matter who carries out the measurement*! Second, as a matter of fact, our contraption above would involve no problematic form of self-reference even if we were to assume that the $\hat{O}$–measurement is to be carried out by A herself.[28] Third, standard quantum mechanics imposes no physical constraints whatsoever on the identity of the observers who may carry out $\hat{O}$–type measurements. Fourth, our argument above leads to no sort of infinite regress.[29]

Taking stock, by this we have established our thesis III above, namely, that the information-theoretic approach is *incomplete* and that this incompleteness leads to inconsistent statistical predictions. Therefore, the approach must be supplemented by further constructive laws that will remove the inconsistency. We see no way of escaping the conclusion that $\hat{O}$–type scenarios require some form of a constructive theory of measurement. This

---

[27] That is, it is claimed that our argument presupposes some sort of a God's–eye view which, by quantum mechanics, is unavailable.

[28] This is essentially because we could always assume that A's memories of the outcomes of the spin measurement and the $\hat{O}$–measurement are to be associated with separate physical features of, say A's brain. So the scenario requires *no* genuine self-reference.

[29] For a more extensive discussion of some of these issues, see Albert (1983, 1990).

is why we believe that neither information–theoretic approaches nor "black-boxes" approaches to quantum measurement circumvent Bell's (1990) objection that decoherence is not enough in order to make sense of quantum measurement—no matter whether or not 'measurement' is operationally construed.

5 Conclusion: the Ancilla Argument

In this paper we have criticized the information–theoretic approach to quantum mechanics and Bub's 'principle' reading of the CBH theorem. Agreed, the theorem remarkably demonstrates that certain salient features of quantum mechanics as we know it (by empirical observation) can be expressed very elegantly in information–theoretic terms. But it does not support any preference to the principle view of quantum theory over its constructive counterparts. Moreover (and unlike other constructive no collapse theories), our thought experiment clearly shows that the information–theoretic approach does not address major unresolved interpretational issues of quantum theory some of which await *empirical* resolution.[30]

More generally, the question raised by this paper boils down to the following. Suppose that crucial experiments that are capable of distinguishing between, say the GRW theory and environmentally induced decoherence were to come out (no kidding!) in accordance with the GRW predictions to a very good approximation.[31] Nevertheless, it is often argued that the principle (information–theoretic) approach to quantum mechanics would remain *intact* for the following reason.

Suppose that an *open* system S is subjected to *perfect* decoherence, namely to interactions with some degrees of freedom in the environment E, such that the environment states become strictly orthogonal. Suppose further that we have no access whatsoever (as a matter of either physical fact or law) to these degrees of freedom. In this case, the GRW dynamics for the density operator of S would be indistinguishable from the dynamics of the *reduced* density operator of S obtained by evolving the composite quantum state of $S + E$ *unitarily* and tracing over the inaccessible degrees

---

[30] See also Hagar 2003.

[31] For recent progress see Adler 2005, Adler *et al.* 2005, Bassi *et al.* 2005a. See also Hemmo and Shenker 2005 for crucial experiments between collapse and no collapse quantum theories testing thermodynamical effects.

of freedom of E. It turns out that this feature is mathematically quite general, because the GRW dynamics for the density operator is a completely positive linear map (see Nielsen and Chuang 2000, pp. 353-373, and Simon, Buzek and Gisin 2001, especially fn. 14). From a physical point of view, this means that the GRW theory is empirically equivalent to a quantum mechanical theory with a unitary (and linear) dynamics of the quantum state defined on a *larger* Hilbert space. In other words, one could always introduce a new quantum mechanical *ancilla* field whose degrees of freedom are inaccessible to us, and cook up a unitary dynamics on the larger Hilbert space that would simulate the GRW dynamics on the reduced density operator. Therefore, experimental results that might seem to confirm GRW-like dynamics could always be *re-interpreted* as confirming quantum mechanical no collapse theories on larger Hilbert spaces. In particular, such a theory could always be made to satisfy the three CBH constraints, and thus save the principle information–theoretic approach.

In so far as information–theoretic approaches are concerned it seems to us that the above argument is quite premature. As the above criticism of the ancilla argument is meant to show, it seems to us that in the current state of quantum mechanics we are far from being able to pinpoint a principle theory. In fact, we believe that in general a principle approach is inferior relative to a constructive one, whenever the latter is possible.[32] The ancilla field in the above argument has, by construction, no observable effects (see, e.g., Diosi 1989) and this amounts to introducing extra variables (or more appropriately, a new 'quantum ether') into standard quantum mechanics, whose sole theoretical role is to save some principles against (putative) empirical refutation. If we are willing to accept *ad hoc* such an argument in the context of non–relativistic quantum mechanics, why should we reject similar 'ether'–like approaches in the context of relativity theory, or hidden variables theories in the context of elementary quantum mechanics? One of the main reasons that such latter approaches are sweepingly rejected (by information theorists, among others) is that their complex underlying structure doesn't translate into new empirical predictions. But, by construction, the ancilla field in the above argument doesn't translate into new empirical predictions either. Moreover, although the ancilla theory could always be made to satisfy

---

[32] This seems to be Einstein's later view despite his introduction of special relativity as a principle theory; see Brown and Timpson (2006) and references therein for an extended account of this issue and of Einstein's views on this matter.

the three CBH principles (in particular the NO BIT condition) on the larger Hilbert space, unconditionally secure bit commitment would in principle be possible (in the sense spelled out in Section 3.2, in so far as our experimental capacities are concerned (as imagined in our scenario above), via protocols that require Alice or Bob to access the ancilla field (which *ex hypothesis* is inaccessible). So, the NO BIT principle as a constraint on the feasible flow of information in the above story becomes quite idle.

More generally, we accept that one *might* have good reasons to protect unitary quantum mechanics against what would seem as a straightforward empirical refutation. From a theoretical point of view it might turn out that both collapse and hidden variables theories could not be made compatible with some fundamental physical principles which we cannot give up without giving up some significant chunk of contemporary theoretical physics (conservation of energy or Lorentz–covariance might be such examples).[33] In that case a protective argument of the kind suggested above might be understandable. But, in our view the current theoretical state of quantum mechanics doesn't warrant such a stance. This is mainly because although extremely successful, quantum mechanics itself has quite deep foundational problems not only at its most basic level (i. e., the measurement problem), but also for example in its generalizations to *both* special and general relativity. In such circumstances we believe that the right epistemological stance is to suspend judgment and let alternative theories, such as the GRW theory, flourish.

Acknowledgement

We thank four anonymous referees, and Itamar Pitowsky and Orly Shenker for very helpful comments. We especially thank Guido Bacciagaluppi for reading an earlier draft of this paper and for valuable and thoughtful suggestions and criticism. AH acknowledges financial support from the Alexander von Humboldt Foundation and the German Government through the Sofja Kovalskaja Award through the PPM group in Konstanz.

---

[33] But compare Putnam (2005). We thank Itamar Pitowsky for conversations about this issue.

References

[1] Adler, S. (2003), ‘Why Decoherence Has not Solved the Measurement Problem: A Response to P. W. Anderson’, *Studies in History and Philosophy of Modern Physics*, 34B, 135–142.

[2] Adler, S. (2005), ‘Stochastic Collapse and Decoherence of a Non-Dissipative Forced Harmonic Oscillator’, *Journal of Physics* A38, 2729–2760.

[3] Adler, S., Bassi, A., and Ippoliti, E. (2005), ‘Towards Quantum Superpositions of a Mirror: an Exact Open Systems Analysis—Calculational Details’, *Journal of Physics* A38, 2715.

[4] Aharonov, Y. and Albert, D. (1981), ‘Can We Make Sense of the Measurement Process in Relativistic Quantum Mechanics?’, *Physical Review Letters*, D 24, 359–370

[5] Albert, D.Z. (1983), ‘On Quantum Mechanical Automata’, *Physics Letters*, A 98, 249–252.

[6] Albert, D.Z (1990), ‘The Quantum Mechanics of Self–measurement’, in: W.H. Zurek (editor), *Complexity, Entropy and the Physics of Information*, New York: Addison-Wesley Publishing Company, 471–476.

[7] Balashov, Y. and Janssen, M. (2003), ‘Presentism and Relativity’, *British Journal of the Philosophy of Science*, 54, 327–346.

[8] Bassi, A., Adler, S., Ippoliti, E. (2005a), ‘Towards Quantum Superpositions of a Mirror: an Exact Open Systems Analysis’, *Physical Review Letters* 94: 030401

[9] Bassi, A., Ippoliti, E., Vacchini, B. (2005b), ‘On the Energy Increase in Space-Collapse Models’, http://xxx.lanl.gov/abs/quant-ph/0506083.

[10] Bell, J.S (1987), *Speakable And Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press.

[11] Bell, J.S. (1990), ‘Against “measurement”’, *Physics world*, 8, 33–40.

[12] Brown, H. (2003), ‘Michelson, FitzGerald and Lorentz: The Origins of Relativity revisited’, http://philsci-archive.pitt.edu/archive/00000987.

[13] Brown, H. (2006) *Physical Relativity*. Oxford: Oxford University Press.

[14] Brown, H. and Timpson, C. (2006), 'Why Special Relativity Should Not Be a Template for a Fundamental Reformulation of Quantum Mechanics', http://xxx.lanl.gov/abs/quant-ph/0601182.

[15] Bub, J. (1997), *Interpreting the Quantum World*, Cambridge: Cambridge University Press.

[16] Bub, J. (2000), 'Quantum Mechanics As A Principle Theory', *Studies in the History and Philosophy of Modern Physics*, 31B, 75–94.

[17] Bub, J. (2004), 'Why the Quantum?' *Studies in the History and Philosophy of Modern Physics*, 35B, 241–266.

[18] Bub, J. (2005), 'Quantum Theory Is About Quantum Information', *Foundations of Physics*, 35(4), 541–560.

[19] Clifton, R., Bub, J. and , Halvorson, H. (2003), 'Characterizing Quantum Theory in Terms of Information–Theoretic Constraints', *Foundations of Physics*, 33(11), 1561–1591.

[20] Caves, C., Fuchs. C. Schack, R. (2002), 'Unknown Quantum States: the Quantum de Finetti Representation', *Journal of Mathematical Physics*, 43, 4537.

[21] Diosi, L. (1989), 'Models for Universal Reduction of Macroscopic Quantum Fluctuations', *Physical Review* A 40, 1165–1174.

[22] Dürr, D., Goldstein, S. and Zanghi, N. (1992), 'Quantum Equilibrium and the Origin of Absolute Uncertainty', *Journal of Statistical Physics*, 67 (5/6), 843–907.

[23] Einstein, A. (1919), 'Time, Space, and Gravitation', *Times* (London). 28 November 1919, 13-14. Reprinted as 'What is the Theory of Relativity?' In Einstein, *Ideas and Opinions*, New York: Bonanza Books. 1954, 227–232.

[24] Fuchs, C. (2002), 'Quantum Mechanics as Quantum Information (And a Little More)', http://xxx.lanl.gov/abs/quant-ph/0205039.

[25] Fuchs, C. and Peres, A. (2000), 'Quantum Theory Needs No Interpretation', *Physics Today*, 3, 70–71.

[26] Ghirardi, G.C. (2000), 'Beyond Conventional Quantum Mechanics', in J. Ellis and D. Amati (eds.), *Quantum Reflections*, Cambridge: Cambridge University Press) 79–116.

[27] Ghirardi, G.C., Rimini, A., and Weber, T. (1980), 'A General Argument Against Superluminal Transmission Through the Quantum Mechanical Measurement Process', *Lettere al Nuovo Cimento*, 27, 293–298.

[28] Ghirardi, G.C., Rimini, A. and Weber, T. (1986), 'Unified Dynamics for Microscopic and Macroscopic Systems', *Physical Review* D34, 470-479.

[29] Gisin, N. (1989), 'Stochastic Quantum Dynamics and Relativity', *Helvetica Physica Acta*, 62, 363-371.

[30] Hagar, A. (2003) 'A Philosopher Looks at Quantum Information Theory' *Philosophy of Science* 70(3), 752–775.

[31] Harman, P.M. (2001), *The Natural Philosophy of James Clerk Maxwell*, Cambridge: Cambridge University Press.

[32] Hemmo, M. and Shenker, O. (2005), 'Quantum Decoherence and the Approach to Equilibrium', *Studies in History and Philosophy of Modern Physics.* Forthcoming in December 2005.

[33] Jánossy, L. (1971), *Theory of Relativity Based on Physical Reality*, Budapest: Akadémiai Kiadó.

[34] Joos, Eric, *et al.* (2003), *Decoherence and the Appearance of a Classical World*, Berlin: Springer

[35] Myrvold, W. (2002), 'On Peaceful Coexistence: Is the Collapse Postulate Incompatible With Relativity?' *Studies in the History and Philosophy of Modern Physics*, 33B, 435–466.

[36] Nielsen, M.A. and Chuang I.L. (2000), *Quantum Computation and Quantum Information*, Cambridge: Cambridge University Press.

[37] Putnam, H. (2005), 'A Philosopher Looks at Quantum Mechanics (Again)', *British Journal for the Philosophy of Science* **56**, 615–634.

[38] Schrödinger, E. (1936), ‘Probability Relations Between Separated Systems’, *Proceedings of the Cambridge Philosophical Society*, 32, 446–452.

[39] Spekkens, R. (2004), ‘In Defense of the Epistemic View of Quantum States: a Toy Theory’, http://xxx.lanl.gov/abs/quant-ph/0401052.

[40] Simon, C., Buzek, V., Gisin, N. (2001), ‘No-Signaling Condition and Quantum Dynamics’, *Physical Review Letters* 87, 170405–170408.

[41] Timpson, C. (2004) ‘Quantum Information Theory and the Foundations of Quantum Mechanics’, PhD Dissertation, Oxford University. http://xxx.lanl.gov/abs/quant-ph/0412063.

[42] Valentini, A. (2002), ‘Subquantum Information and Computation’, http://xxx.lanl.gov/abs/quant-ph/0203049.

[43] Zurek, W.H. (1991), ’Decoherence and the Transition From Quantum to Classical’, *Physics Today*, 44, 36–44.