Quantum prediction algorithms

Adrian Kent* and Jim McElwaine†
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, United Kingdom
(Received 15 October 1996)

The consistent histories formulation of the quantum theory of a closed system with a pure initial state defines an infinite number of incompatible consistent sets, each of which gives a possible description of the physics. We investigate the possibility of using the properties of the Schmidt decomposition to define an algorithm which selects a single, physically natural, consistent set. We explain the problems which arise, set out some possible algorithms, and explain their properties with the aid of simple models. Though the discussion is framed in the language of the consistent histories approach, it is intended to highlight the difficulty in making any interpretation of quantum theory based on decoherence into a mathematically precise theory. [S1050-2947(97)05103-2]

PACS number(s): 03.65.Bz, 98.80.Hw

I. INTRODUCTION

It is hard to find an entirely satisfactory interpretation of the quantum theory of closed systems, since quantum theory does not distinguish physically interesting time-ordered sequences of operators. The consistent histories approach to quantum theory was originally developed by Griffiths [1], Omnès [2], and Gell-Mann and Hartle [3]. One of its virtues, in our view, is that it allows the problems of the quantum theory of closed systems to be formulated precisely enough to allow us to explore possible solutions. A natural probability distribution is defined on each consistent set of histories, allowing probabilistic predictions to be made from the initial data. There are infinitely many consistent sets, which are incompatible in the sense that pairs of sets generally admit no physically sensible joint probability distribution whose marginal distributions agree with those on the individual sets. Indeed the standard no-local-hidden-variable theorems show that there is no joint probability distribution defined on the collection of histories belonging to all consistent sets [4,5]. Hence, the set selection problem: probabilistic predictions can only be made conditional on a choice of consistent set, yet the consistent histories formalism gives no way of singling out any particular set or sets as physically interesting.

One possible solution to the set selection problem would be an axiom which identifies a unique physically interesting set, or perhaps a class of such sets, from the initial state and the dynamics. Another would be the identification of a physically natural measure on the space of consistent sets, according to which the physically relevant consistent set is randomly chosen. No workable solution has yet been proposed, however.

The problem remains essentially unaltered if the predictions are conditioned on a large collection of data [5], and even if predictions are made conditional on approximately classical physics being observed [6]. The consistent histories approach thus violates both standard scientific criteria and ordinary intuition [5–10]. In our view, the present version of the consistent histories formalism is too weakly predictive in almost all plausible physical situations to be considered a fundamental scientific theory. Nonetheless, we believe that the consistent histories approach gives a new way of looking at quantum theory which raises intriguing questions and should, if possible, be developed further.

The status of the consistent histories approach remains controversial: much more optimistic assessments of the present state of the formalism can be found, for example, in Refs. [3,11,12]. It is, though, generally agreed that set selection criteria should be investigated. For if quantum theory correctly describes macroscopic physics then, it is believed, real-world experiments and observations can be described by what Gell-Mann and Hartle term *quasiclassical* consistent sets of histories. Roughly speaking, quasiclassical sets are defined by projection operators which involve similar variables at different times and which satisfy classical equations of motion, to a very good approximation, most of the time. No precise definition of quasiclassicality has yet been found, nor is any systematic way known of identifying quasiclassical sets within any given model or theory. Whether Gell-Mann and Hartle's program of characterizing quasiclassical sets is taken as a fundamental problem or a phenomenological one, any solution must clearly involve some sort of set selection mechanism.

In this paper, we consider one particular line of attack on this problem: the attempt to select consistent sets by using the Schmidt decomposition together with criteria intrinsic to the consistent histories formalism. The paper is exploratory in spirit: our aims here are to point out obstacles, raise questions, set out some possible selection principles, and explain their properties.

Our discussion is framed in the language of the consistent histories approach to quantum theory, but we believe it is of wider relevance. Many modern attempts to provide an interpretation of quantum theory rely, ultimately, on the fact that quantum subsystems decohere. Subsystems considered include the brains of observers, the pointers of measuring devices, and abstractly defined subspaces of the total Hilbert

*Electronic address: apak@damtp.cambridge.ac.uk
†Electronic address: jnm11@damtp.cambridge.ac.uk

space. Whichever, the moral is intended to be that decoherence selects the projection operators, or space-time events, or algebras of observables which characterize the physics of the subsystem as it is experienced or observed. There is no doubt that understanding the physics of decoherence *does* provide a very good intuitive grasp of how to identify operators from which our everyday picture of real-world quasiclassical physics can be constructed and this lends some support to the hope that a workable interpretation of quantum theory — a plausible successor to the Copenhagen interpretation — *could* possibly be constructed along the lines just described.

However, it seems to us that the key question is whether such an interpretation can be made mathematically precise.[1] That is, given a decohering subsystem, can we find general rules which precisely specify operators (or other mathematical objects) which allow us to recover the subsystem's physics as we experience or observe it? From this point of view, we illustrate below how one might go about setting out such rules, and the sort of problems which arise.

A. Consistent histories

We use a version of the consistent histories formalism in which the initial conditions are defined by a pure state, the basic objects of the formalism are branch-dependent sets of projections, and consistency is defined by Gell-Mann and Hartle's decoherence criterion Eq. (1.3).

Let $|\psi\rangle$ be the initial state of a quantum system. A *branch-dependent set of histories* is a set of products of projection operators indexed by the variables $\alpha=\{\alpha_n,\alpha_{n-1},\ldots,\alpha_1\}$ and corresponding time coordinates $\{t_n,\ldots,t_1\}$, where the ranges of the α_k and the projections they define depend on the values of $\alpha_{k-1},\ldots,\alpha_1$, and the histories take the form

$$C_\alpha=P^n_{\alpha_n}(t_n;\alpha_{n-1},\ldots,\alpha_1)\times P^{n-1}_{\alpha_{n-1}}(t_{n-1};\alpha_{n-2},\ldots,\alpha_1)\ldots P^1_{\alpha_1}(t_1). \quad (1.1)$$

Here, for fixed values of $\alpha_{k-1},\ldots,\alpha_1$, the $P^k_{\alpha_k}(t_k;\alpha_{k-1},\ldots,\alpha_1)$ define a projective decomposition of the identity[2] indexed by α_k, so that $\Sigma_{\alpha_k}P^k_{\alpha_k}(t_k;\alpha_{k-1},\ldots,\alpha_1)=1$ and

$$P^k_{\alpha_k}(t_k;\alpha_{k-1},\ldots,\alpha_1)P^k_{\alpha_k'}(t_k;\alpha_{k-1},\ldots,\alpha_1)=\delta_{\alpha_k\alpha_k'}P^k_{\alpha_k}(t_k;\alpha_{k-1},\ldots,\alpha_1). \quad (1.2)$$

The set of histories is *consistent*,[3] if and only if,

$$D_{\alpha\beta}=\langle\psi|C^\dagger_\beta C_\alpha|\psi\rangle=\delta_{\alpha\beta}p(\alpha), \quad (1.3)$$

in which case $p(\alpha)$ is interpreted as the probability of the history α. D is the decoherence matrix. Here and later, though, we use the compact notation α to refer to a history, we intend the individual projection operators and their associated times to define the history. The histories of nonzero probability in a consistent set thus correspond precisely to the nonzero vectors $C_\alpha|\psi\rangle$. According to the standard view of the consistent histories formalism, which we adopt here, it is only consistent sets which are of physical relevance. The dynamics are defined purely by the Hamiltonian, with no collapse postulate, but each projection in the history can be thought of as corresponding to a historical event, taking place at the relevant time. If a given history is realized, its events correspond to extra physical information, neither deducible from the state vector nor influencing it.

Most projection operators involve rather obscure physical quantities, so that it is hard to interpret a general history in familiar language. However, given a sensible model, with Hamiltonian and canonical variables specified, one can construct sets of histories which describe familiar physics and check that they are indeed consistent to a very good approximation. For example, a useful set of histories for describing the solar system could be defined by projection operators whose nonzero eigenspaces contain states in which a given planet's center of mass is located in suitably chosen small volumes of space at the relevant times, and one would expect a sensible model to show that this is a consistent set and that the histories of significant probability are those agreeing with the trajectories predicted by general relativity.

More generally, Gell-Mann and Hartle [3] introduce the notion of a *quasiclassical domain*: a consistent set which is complete — so that it cannot be nontrivially consistently extended by more projective decompositions — and is defined by projection operators which involve similar variables at different times and which satisfy classical equations of motion, to a very good approximation, most of the time. The notion of a quasiclassical domain seems natural, though presently imprecisely defined. Its heuristic definition is motivated by the familiar example of the hydrodynamic variables — densities of chemical species in small volumes of space, and similar quantities — which characterize our own quasiclassical domain. Here the branch dependence of the formalism plays an important role, since the precise choice of variables (most obviously, the sizes of the small volumes) we use depends on earlier historical events. The formation of our Galaxy and Solar System influences all subsequent local physics; even present-day quantum experiments have the potential to do so significantly, if we arrange for large macroscopic events to depend on their results.

---

[1] Even those who believe that an interpretation relying on intuitive ideas or verbal prescriptions is acceptable would, we hope, concede that it is interesting to ask whether those ideas and prescriptions *can* be set out mathematically.

[2] For brevity, we refer to projective decompositions of the identity as projective decompositions hereafter.

[3] Several different consistency criteria are used in the literature, all of which are believed to be compatible with the standard quasiclassical descriptions of realistic physical examples. This particular criterion is generally known as *medium consistency* or *medium decoherence*; it will be used throughout the paper.

It should be stressed that, according to all the developers of the consistent histories approach, quasiclassicality and related properties are interesting notions to study within, not defining features of, the formalism. In the view of the formalism's developers, all consistent sets of histories have the same physical status, though in any realistic example we are likely to be more interested in the descriptions of the physics given by some than by others.

Identifying interesting consistent sets of histories is presently more of an art than a science. One of the original aims of the consistent histories formalism, stressed, in particular, by Griffiths and Omnès, was to provide a theoretical justification for the intuitive language often used, both by theorists and experimenters, in analyzing laboratory setups. Even here, though there are many interesting examples in the literature of consistent sets which give a natural description of particular experiments, no general principles have been found by which such sets can be identified. Identifying interesting consistent sets in quantum cosmological models or in real-world cosmology seems to be still harder, although some interesting criteria stronger than consistency have recently been proposed [10,13].

B. The Schmidt decomposition

We consider a closed quantum system with a pure initial-state vector $|\psi(0)\rangle$ in a Hilbert space $\mathcal{H}$ with Hamiltonian H. We suppose that $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$; we write $\dim(\mathcal{H}_j)=d_j$ and we suppose that $d_1\leqslant d_2<\infty$. With respect to this splitting of the Hilbert space, the *Schmidt decomposition* of $|\psi(t)\rangle$ is an expression of the form

$$|\psi(t)\rangle=\sum_{i=1}^{d_1}[p_i(t)]^{1/2}|w_i(t)\rangle_1\otimes|w_i(t)\rangle_2, \qquad (1.4)$$

where the *Schmidt states* $\{|w_i\rangle_1\}$ and $\{|w_i\rangle_2\}$ form, respectively, an orthonormal basis of $\mathcal{H}_1$ and part of an orthonormal basis of $\mathcal{H}_2$, the functions $p_i(t)$ are real and positive, and we take the positive square root. For fixed time t, any decomposition of the form Eq. (1.4) then has the same list of probability weights $\{p_i(t)\}$, and the decomposition (1.4) is unique if these weights are all different. These probability weights are the eigenvalues of the reduced density matrix.

This simple result, proved by Schmidt in 1907 [14], means that at any given time there is, generically, a natural decomposition of the state vector relative to any fixed split $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$, which defines a basis on the smaller space $\mathcal{H}_1$ and a partial basis on $\mathcal{H}_2$. The decomposition has an obvious application in standard Copenhagen quantum theory where, if the two spaces correspond to subsystems undergoing a measurement-type interaction, it describes the final outcomes [15].

It has more than once been suggested that the Schmidt decomposition *per se* might define a fundamental interpretation of quantum theory. According to one line of thought, it defines the structure required in order to make precise sense of Everett's ideas [16]. Another idea which has attracted some attention is that the Schmidt decomposition itself defines a fundamental interpretation [17–20]. Some critical comments on this last program, motivated by its irreconcilability with the quantum history probabilities defined by the decoherence matrix, can be found in Ref. [21].

Though a detailed critique is beyond our scope here, it seems to us that any attempt to interpret quantum theory which relies solely on the properties of the Schmidt decomposition must fail, even if some fixed choice of $\mathcal{H}_1$ and $\mathcal{H}_2$ is allowed. The Schmidt decomposition seems inadequate as, although it allows a plausible interpretation of the quantum state at a single fixed time, its time evolution has no natural interpretation consistent with the predictions of Copenhagen quantum theory.

Many studies have been made of the behavior of the Schmidt decomposition during system-environment interactions. In developing the ideas of this paper, we were influenced, in particular, by Albrecht's investigations [22,23] of the behavior of the Schmidt decomposition in random Hamiltonian interaction models and the description of these models by consistent histories.

C. Combining consistency and the Schmidt decomposition

The idea motivating this paper is that the combination of the ideas of the consistent histories formalism and the Schmidt decomposition might allow us to define a mathematically precise and physically interesting description of the quantum theory of a closed system. The Schmidt decomposition defines four natural classes of projection operators, which we refer to collectively as *Schmidt projections*. These take the form

$$P_i^1(t)=|w_i(t)\rangle_1\langle w_i(t)|_1\otimes I_2 \quad\text{and}\quad \overline{P^1}=I_1\otimes I_2-\sum_i P_i^1(t),$$

$$P_i^2(t)=I_1\otimes|w_i(t)\rangle_2\langle w_i(t)|_2 \quad\text{and}\quad \overline{P^2}=I_1\otimes I_2-\sum_i P_i^2(t),$$

$$P_i^3(t)=|w_i(t)\rangle_1\langle w_i(t)|_1\otimes|w_i(t)\rangle_2\langle w_i(t)|_2$$

$$\text{and}\quad \overline{P^3}=I_1\otimes I_2-\sum_i P_i^3(t),$$

$$P_{ij}^4(t)=|w_i(t)\rangle_1\langle w_i(t)|_1\otimes|w_j(t)\rangle_2\langle w_j(t)|_2$$

$$\text{and}\quad \overline{P^4}=I_1\otimes I_2-\sum_{ij} P_{ij}^4(t). \qquad (1.5)$$

If $\dim\mathcal{H}_1=\dim\mathcal{H}_2$ the complementary projections $\overline{P^1}$, $\overline{P^2}$, and $\overline{P^4}$ are zero.

Since the fundamental problem with the consistent histories approach seems to be that it allows far too many consistent sets of projections, and since the Schmidt projections appear to be natural dynamically determined projections, it seems sensible to explore the possibility that a physically sensible rule can be found which selects a consistent set or sets from amongst those defined by Schmidt projections.

The first problem in implementing this idea is choosing the split $\mathcal{H}=\mathcal{H}_1\otimes\mathcal{H}_2$. In analyzing laboratory experiments, one obvious possibility is to separate the system and apparatus degrees of freedom. Other possibilities of more general application are to take the split to correspond to more fundamental divisions of the degrees of freedom — fermions and bosons, or massive and massless particles, or, one might speculate, the matter and gravitational fields in quantum gravity. Some such division would necessarily have to be introduced if this proposal were applied to cosmological models.

Each of these choices seems interesting to us in context, but none, of course, is conceptually cost free. Assuming a division between system and apparatus in a laboratory experiment seems to us unacceptable in a fundamental theory, reintroducing, as it does, the Heisenberg cut which post-Copenhagen quantum theory aims to eliminate. It seems justifiable, though, for the limited purpose of discussing the consistent sets which describe physically interesting histories in laboratory situations. It also allows useful tests: if an algorithm fails to give sensible answers here, it should probably be discarded; if it succeeds, applications elsewhere may be worth exploring.

Postulating a fundamental split of Hilbert space, on the other hand, seems to us acceptable in principle. If the split chosen were reasonably natural, and if it were to produce a well-defined and physically sensible interpretation of quantum theory applied to closed systems, we would see no reason not to adopt it. This seems a possibility especially worth exploring in quantum cosmology, where any pointers towards calculations that might give new physical insight would be welcome.

Here, though, we leave aside these motivations and the conceptual questions they raise, as there are simpler and more concrete problems which first need to be addressed. Our aim in this paper is simply to explain the problems which arise in trying to define consistent set selection algorithms using the Schmidt decomposition, to set out some possibilities, and to explain their properties, using simple models of quantum systems interacting with an idealized experimental device or with a series of such devices.

The most basic question here is precisely which of the Schmidt projections should be used. Again, our view is pragmatic: we would happily adopt any choice that gave physically interesting results. Where we discuss the abstract features of Schmidt projection algorithms below, the discussion is intended to apply to all four choices. When we consider simple models of experimental setups, we take $\mathcal{H}_1$ to describe the system variables and $\mathcal{H}_2$ the apparatus or environment. Here we look for histories which describe the evolution of the system state, tracing over the environment, and so discuss set selection algorithms which use only the first class of Schmidt projections: the other possibilities are also interesting, but run into essentially the same problems. Thus, in the remainder of the paper, we use the term Schmidt projection to mean the system space Schmidt projections denoted by P_i^1 and $\overline{P^1}$, defined in Eq. (1.5).

In most of the following discussion, we consider algorithms which use only the properties of the state vector $|\psi(t)\rangle$ and its Schmidt decomposition to select a consistent set. However, we will also consider later the possibility of reconstructing a branching structure defined by the decomposition

$$|\psi(t)\rangle = \sum_{i=1}^{N(t)} |\psi_i(t)\rangle, \tag{1.6}$$

in which the selected set is branch dependent and the distinct orthogonal components $|\psi_i(t)\rangle$ correspond to the different branches at time t. In this case, we will consider the Schmidt decompositions of each of the $|\psi_i(t)\rangle$ separately. Again, it will be sufficient to consider only the first class of Schmidt projections. In fact, for the branch-dependent algorithms we consider, all of the classes of Schmidt projection select the same history vectors and, hence, select physically equivalent consistent sets.

II. APPROXIMATE CONSISTENCY AND NONTRIVIALITY

In realistic examples it is generally difficult to find simple examples of physically interesting sets that are exactly consistent. For simple physical projections, the off-diagonal terms of the decoherence matrix typically decay exponentially. The sets of histories defined by these projections separated by times much larger than the decoherence time, are thus typically very nearly but not precisely consistent [24–35]. Histories formed from Schmidt projections are no exception: they give rise to exactly consistent sets only in special cases, and even in these cases the exact consistency is unstable under perturbations of the initial conditions or the Hamiltonian.

The lack of simple exactly consistent sets is not generally thought to be a fundamental problem *per se*. According to one controversial view [3], probabilities in any physical theory need only be defined, and need only satisfy sum rules, to a very good approximation, so that approximately consistent sets are all that is ever needed. Incorporating pragmatic observation into fundamental theory in this way clearly, at the very least, raises awkward questions. Fortunately, it seems unnecessary. There are good reasons to expect [5] to find exactly consistent sets very close to a generic approximately consistent set, so that even if only exactly consistent sets are permitted the standard quasiclassical description can be recovered. Note, though, that none of the relevant exactly consistent sets will generally be defined by Schmidt projections.

It could be argued that physically reasonable set selection criteria should make predictions which vary continuously with structural perturbations and perturbations in the initial conditions, and that the instability of exact consistency under perturbation means that the most useful consistency criteria are very likely to be approximate. Certainly, there seems no reason, in principle, why a precisely defined selection algorithm, which gives physically sensible answers, should be rejected if it fails to exactly respect the consistency criterion. For, once a single set has been selected, there seems no fundamental problem in taking the decoherence functional probability weights to represent precisely the probabilities of its fine-grained histories and the probability sum rules to *define* the probabilities of coarse-grained histories. On the other hand, allowing approximate consistency raises new difficulties in identifying a single natural set selection algorithm, since any such algorithm would have — at least indirectly — to specify the degree of approximation tolerated.

These arguments over fundamentals, though, go beyond our scope here. Our aim below is to investigate selection rules which might give physically interesting descriptions of quantum systems, whether or not they produce exactly consistent sets. As we will see, it seems surprisingly hard to find good selection rules even when we follow the standard procedure in the decoherence literature and allow some degree of approximate decoherence.

Mathematical definitions of approximate consistency were first investigated by Dowker and Halliwell [27], who proposed a simple criterion — the Dowker-Halliwell criterion, or DHC — according to which a set is approximately consistent to order ϵ if the decoherence functional

$$D_{\alpha\beta}=\langle\psi|C_{\beta}^{\dagger}C_{\alpha}|\psi\rangle \tag{2.1}$$

satisfies the equation

$$|D_{\alpha\beta}|\leqslant\epsilon(D_{\alpha\alpha}D_{\beta\beta})^{1/2},\quad \forall\alpha\neq\beta. \tag{2.2}$$

Approximate consistency criteria were analyzed further in Ref. [36]. As Refs. [27,36] explain, the DHC has natural physical properties and is well adapted for mathematical analyses of consistency. We adopt it here, and refer to the largest term,

$$\max\{|D_{\alpha\beta}|(D_{\alpha\alpha}D_{\beta\beta})^{-1/2}:\alpha,\beta\in S,\alpha\neq\beta,\ \text{and } D_{\alpha\alpha},D_{\beta\beta}\neq 0\}, \tag{2.3}$$

of a (possibly incomplete) set of histories S as the Dowker-Halliwell parameter, or DHP.

A *trivial* history α is one whose probability is zero, $C_{\alpha}|\psi\rangle=0$. Many of the algorithms we discuss involve, as well as the DHP, a parameter which characterizes the degree to which histories approach triviality. The simplest nontriviality criterion would be to require that all history probabilities must be greater than some parameter δ, i.e., that

$$D_{\alpha\alpha}>\delta\quad\text{for all histories }\alpha. \tag{2.4}$$

As a condition on a particular extension $\{P_i:i=1,2,\ldots\}$ of the history α this would imply that $\|P_iC_{\alpha}|\psi\rangle\|^2>\delta$ for all i. This, of course, is an absolute condition, which depends on the probability of the original history α rather than on the relative probabilities of the extensions and which implies that once a history with probability less than 2δ has been selected any further extension is forbidden.

It seems to us more natural to use criteria, such as the DHC, which involve only relative probabilities. It is certainly simpler in practice: applying absolute criteria strictly would require us to compute from first cosmological principles the probability to date of the history in which we find ourselves. We therefore propose the following relative nontriviality criterion: an extension $\{P_i:i=1,2,\ldots\}$ of the nontrivial history α is nontrivial to order δ, for any δ with $0<\delta<1$, if

$$\|P_iC_{\alpha}|\psi\rangle\|^2\geqslant\delta\|C_{\alpha}|\psi\rangle\|^2\quad\text{for all } i. \tag{2.5}$$

We say that a set of histories S, which may be branch dependent, is nontrivial to order δ if every set of projections, considered as an extension of the histories up to the time at which it is applied, is nontrivial to order δ. In both cases we refer to δ as the nontriviality parameter, or NTP.

An obvious disadvantage of applying an absolute nontriviality criterion to branch-independent consistent sets is that, if the set contains one history of probability less than or equal to 2δ, no further extensions are permitted.

Once again, though, our approach is pragmatic, and in order to cover all the obvious possibilities we investigate below absolute consistency and nontriviality criteria as well as relative ones.

III. REPEATED PROJECTIONS AND CONSISTENCY

One of the problems which arises in trying to define physically interesting set selection algorithms is the need to find a way either of preventing near-instantaneous repetitions of similar projections or of ensuring that such repetitions, when permitted, do not prevent the algorithm from making physically interesting projections at later times. It is useful, in analyzing the behavior of repeated projections, to introduce a version of the DHC which applies to the coincident time limit of sets of histories defined by smoothly time-dependent projective decompositions.

To define this criterion, fix a particular time t_0, and consider class operators C_{α} consisting of projections at times $\mathbf{t}=(t_1,\ldots,t_n)$, where $t_n>t_{n-1}>\cdots>t_1>t_0$. Define the *normalized histories* by

$$|\hat{\alpha}\rangle=\lim_{\mathbf{t}'\to\mathbf{t}}\frac{C_{\alpha}(\mathbf{t}')|\psi\rangle}{\|C_{\alpha}(\mathbf{t}')|\psi\rangle\|}, \tag{3.1}$$

where the limits are taken in the order $t_1'\to t_1$ then $t_2'\to t_2$ and so on, whenever these limits exist. Define the limit DHC between two normalized histories $|\hat{\alpha}\rangle$ and $|\hat{\beta}\rangle$ as

$$\langle\hat{\alpha}|\hat{\beta}\rangle\leqslant\epsilon. \tag{3.2}$$

This, of course, is equivalent to the limit of the ordinary DHC when the limiting histories exist and are not null. It defines a stronger condition when the limiting histories exist and at least one of them is null, since in this case the limit of the DHC is automatically satisfied.

If a set of histories is defined by a smoothly time-dependent projective decomposition applied at two nearby times, it will contain many nearly null histories, since $P_mP_n=0$ for all $n\neq m$. Clearly, in the limit as the time separation tends to zero, these histories become null, so that the limit of the ordinary DHC is automatically satisfied. When do the normalized histories satisfy the stronger criterion (3.2)?

Let $P(t)$ be a projection operator with a Taylor series at $t=0$,

$$P(t)=P+t\dot{P}+\tfrac{1}{2}t^2\ddot{P}+O(t^3), \tag{3.3}$$

where $P=P(0)$, $\dot{P}=dP(t)/dt|_{t=0}$ and $\ddot{P}=d^2P(t)/dt^2|_{t=0}$. Since $P^2(t)=P(t)$ for all t

$$
\begin{aligned}
P+t\dot{P}+\tfrac{1}{2}t^2\ddot{P}+O(t^3)&=[P+t\dot{P}+\tfrac{1}{2}t^2\ddot{P}+O(t^3)][P+t\dot{P}\\
&\quad+\tfrac{1}{2}t^2\ddot{P}+O(t^3)]\\
&=P+t(P\dot{P}+\dot{P}P)\\
&\quad+\tfrac{1}{2}t^2(P\ddot{P}+\ddot{P}P+2\dot{P}^2)+O(t^3).
\end{aligned}
\tag{3.4}
$$

This implies that

$$\dot{P}=P\dot{P}+\dot{P}P, \tag{3.5}$$

and

$$\tfrac{1}{2}\ddot{P}=\tfrac{1}{2}P\ddot{P}+\tfrac{1}{2}\ddot{P}P+\dot{P}^2. \tag{3.6}$$

Now consider a projective decomposition $\{P_k\}$ and the matrix element

$$
\begin{aligned}
\langle\psi|P_mP_k(t)P_n|\psi\rangle&=\langle\psi|P_mP_kP_n|\psi\rangle+t\langle\psi|P_m\dot{P}_kP_n|\psi\rangle\\
&\quad+\tfrac{1}{2}t^2\langle\psi|P_m\ddot{P}_kP_n|\psi\rangle+O(t^3).
\end{aligned}
\tag{3.7}
$$

Now $P_mP_kP_n=P_k\delta_{km}\delta_{kn}$, since the projections are orthogonal, and

$$
\begin{aligned}
P_m\dot{P}_kP_n&=\delta_{km}(1-\delta_{kn})\dot{P}_kP_n+\delta_{kn}(1-\delta_{km})P_m\dot{P}_k\\
&=\delta_{km}\dot{P}_kP_n+\delta_{kn}P_m\dot{P}_k-\delta_{km}\delta_{kn}\dot{P}_k,
\end{aligned}
\tag{3.8}
$$

since $\dot{P}_kP_n=P_k\dot{P}_kP_n$ if $k\neq n$ and $\dot{P}_kP_k=(1-P_k)\dot{P}_k$. (No summation convention applies throughout this paper.) From Eq. (3.6) we have that

$$\tfrac{1}{2}P_m\ddot{P}_kP_n=\tfrac{1}{2}(\delta_{mk}+\delta_{nk})P_m\ddot{P}_kP_n+P_m\dot{P}_k^2P_n. \tag{3.9}$$

Equation (3.7) can now be simplified. To leading order in t it is

$$\langle\psi|P_k|\psi\rangle+O(t)\quad\text{if}\ \ k=m=n, \tag{3.10}$$

$$t\langle\psi|\dot{P}_kP_n|\psi\rangle+O(t^2)\quad\text{if}\ \ k=m,\ \ k\neq n, \tag{3.11}$$

$$t\langle\psi|P_{mP_k|\psi}\rangle+O(t^2)\quad\text{if}\ \ k\neq m,\ \ k=n, \tag{3.12}$$

and

$$t^2\langle\psi|P_m\dot{P}_k^2P_n|\psi\rangle+O(t^3)\quad\text{if}\ \ k\neq m,\ \ k\neq n. \tag{3.13}$$

Now consider a smoothly time-dependent projective decomposition, $\sigma(t)=\{P(t),\overline{P}(t)\}$, defined by a time-dependent projection operator and its complement. Write $P=P(0)$, and consider a state $|\phi\rangle$ such that $P|\phi\rangle\neq 0$ and $\overline{P}|\phi\rangle\neq 0$. We consider a set of histories with initial projections $P,\overline{P}$, so that the normalized history states at $t=0$ are

$$\left\{\frac{P|\phi\rangle}{\|P|\phi\rangle\|},\frac{\overline{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|}\right\}, \tag{3.14}$$

and consider an extended branch-dependent set defined by applying $\sigma(t)$ on one of the branches — say, the first — at a later time t.

The new normalized history states are

$$\left\{\frac{P(t)P|\phi\rangle}{\|P(t)P|\phi\rangle\|},\frac{\overline{P(t)}P|\phi\rangle}{\|\overline{P(t)}P|\phi\rangle\|},\frac{\overline{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|}\right\}. \tag{3.15}$$

We assume now that $\dot{P}P|\phi\rangle\neq 0$, so that the limit of these states as $t\to 0$ exists. We have that

$$\lim_{t\to 0}\frac{(\overline{P}-t\dot{P})P|\phi\rangle}{(t^2\langle\phi|P\dot{P}^2P|\phi\rangle)^{1/2}}=\frac{-\dot{P}P|\phi\rangle}{\|\dot{P}P|\phi\rangle\|}, \tag{3.16}$$

so that the limits of the normalized histories are

$$\left\{\frac{P|\phi\rangle}{\|P|\phi\rangle\|},\frac{-\dot{P}P|\phi\rangle}{\|\dot{P}P|\phi\rangle\|},\frac{\overline{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|}\right\}. \tag{3.17}$$

The only possibly nonzero terms in the limit DHC are

$$-\frac{\langle\phi|\overline{P}\dot{P}P|\phi\rangle}{\|\overline{P}|\phi\rangle\|\|\dot{P}P|\phi\rangle\|}=-\frac{\langle\phi|\overline{P}\dot{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|\|\overline{P}\dot{P}|\phi\rangle\|}, \tag{3.18}$$

which generically do not vanish.

Consider instead extending the second branch using $P(t)$ again. This gives the set

$$\left\{\frac{P|\phi\rangle}{\|P|\phi\rangle\|},\frac{\overline{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|},\frac{-P(t)\dot{P}P|\phi\rangle}{\|P(t)\dot{P}P|\phi\rangle\|},\frac{-\overline{P}(t)\dot{P}P|\phi\rangle}{\|\overline{P}(t)\dot{P}P|\phi\rangle\|}\right\}. \tag{3.19}$$

Since $P\dot{P}P=0$ the limit $t\to 0$ exists and is

$$\left\{\frac{P|\phi\rangle}{\|P|\phi\rangle\|},\frac{\overline{P}|\phi\rangle}{\|\overline{P}|\phi\rangle\|},\frac{-\dot{P}^2P|\phi\rangle}{\|\dot{P}^2P|\phi\rangle\|},\frac{-\dot{P}P|\phi\rangle}{\|\dot{P}P|\phi\rangle\|}\right\}. \tag{3.20}$$

The DHC term between the first and third histories is

$$-\frac{\langle\phi|P\dot{P}^2P|\phi\rangle}{\|P|\phi\rangle\|\|\dot{P}^2P|\phi\rangle\|}=-\frac{\|\dot{P}P|\phi\rangle\|^2}{\|P|\phi\rangle\|\|\dot{P}^2P|\phi\rangle\|}. \tag{3.21}$$

This is always nonzero since $P\dot{P}|\phi\rangle\neq 0$.

For the same reason, extending the first branch again, or the third branch, violates the limit DHC. Hence, if projections are taken from a continuously parametrized set, and the limit DHC is used, multiple reprojections will generically be forbidden.

The assumption that $\dot{P}P|\phi\rangle\neq 0$ can be relaxed. It is sufficient, for example, that there is some k such that $\|P^{(j)}\|=0$ for all $j<k$ and that $P^{(k)}P|\phi\rangle\neq 0$, where $P^{(j)}=d^jP(t)/dt^j|_{t=0}$.

Note, finally, that it is easy to construct examples in which a single reprojection is consistent. For instance, let

$$P=\begin{pmatrix} I_{d_1} & 0 \\ 0 & 0 \end{pmatrix}, \quad \overline{P}=\begin{pmatrix} 0 & 0 \\ 0 & I_{d_2} \end{pmatrix}, \quad \dot{P}=\begin{pmatrix} 0 & A^{\dagger} \\ A & 0 \end{pmatrix},$$

$$|\phi\rangle=\begin{pmatrix} \sqrt{q}\hat{\mathbf{x}} \\ \sqrt{1-q}\hat{\mathbf{y}} \end{pmatrix}, \tag{3.22}$$

where $\hat{\mathbf{x}}$ is a unit vector in C^{d_1}, $\hat{\mathbf{y}}$ a unit vector in C^{d_2}, and A a $d_2\times d_1$ complex matrix. $\|P|\phi\rangle\|\neq 0,1$ implies that $q\neq 0,1$ and $\dot{P}P|\phi\rangle\neq 0$ implies that $A\hat{\mathbf{x}}\neq 0$. So from Eq. (3.18) the DHC term is

$$-\frac{\hat{\mathbf{y}}^{\dagger}A\hat{\mathbf{x}}}{\|A\hat{\mathbf{x}}\|}. \tag{3.23}$$

If $d_2\geqslant 2$ then $\hat{\mathbf{y}}$ can be chosen orthogonal to $A\hat{\mathbf{x}}$ and then Eq. (3.23) is zero. The triple projection term, however, Eq. (3.21), is

$$-\frac{\|A\hat{\mathbf{x}}\|^2}{\|A^2\hat{\mathbf{x}}\|}, \tag{3.24}$$

which is never equal to 0 since $A\hat{\mathbf{x}}\neq 0$.

IV. SCHMIDT PROJECTION ALGORITHMS

We turn now to the problem of defining a physically sensible set selection algorithm which uses Schmidt projections, starting in this section with an abstract discussion of the properties of Schmidt projection algorithms.

We consider here dynamically generated algorithms in which initial projections are specified at $t=0$, and the selected consistent set is then built up by selecting later projective decompositions, whose projections are sums of the Schmidt projection operators, as soon as specified criteria are satisfied. The projections selected up to time t thus depend only on the evolution of the system up to that time. We will generally consider selection algorithms for branch-independent sets and add comments on related branch-dependent selection algorithms.

We assume that there is a set of Heisenberg picture Schmidt projection operators $\{P_n(t)\}$ with continuous time dependence, defined even at points where the Schmidt probability weights are degenerate, write P_n for $P_n(0)$, and let I be the index set for projections which do not annihilate the initial state, $I=\{n:P_n|\psi\rangle\neq 0\}$.

We consider first a simple algorithm, in which the initial projections are fixed to be the P_n for $n\in I$ together with their complement $(1-\Sigma_n P_n)$, and which then selects decompositions built from Schmidt projections at the earliest possible time, provided they are consistent. More precisely, suppose that the algorithm has selected a consistent set S_k of projective decompositions at times $t_0,t_1,\ldots,t_k$. It then selects the earliest time $t_{k+1}>t_k$ such that there is at least one consistent extension of the set S_k by a projective decomposition formed from sums of Schmidt projections at time t_{k+1}. In generic physical situations, we expect this decomposition to be unique. However, if more than one such decomposition exists, the one with the largest number of projections is selected; if more than one decomposition has the maximal number of projections, one is randomly selected.

Though the limit DHC (3.2) can prevent trivial projections, it does not generically do so here. The limit DHC terms between histories m and n for an extension involving P_k $(k\notin I)$ are

$$\lim_{t\to 0}\frac{|\langle\psi|P_m P_k(t)P_n|\psi\rangle|}{\|P_m|\psi\rangle\|\|P_k(t)P_n|\psi\rangle\|}=t\frac{|\langle\psi|P_m\dot{P}_k^2P_n|\psi\rangle|}{\|P_m|\psi\rangle\|\|P_kP_n|\psi\rangle\|}=0, \tag{4.1}$$

whenever $\|P_m|\psi\rangle\|$ and $\|\dot{P}_kP_n|\psi\rangle\|$ are both nonzero. The first is nonzero by assumption; the second is generically nonzero. Thus the extension of all histories by the projections $P_k(k\notin I)$ and $\Sigma_{n\in I}P_n$ satisfies the limit DHC.

Hence, if the initial projections do not involve all the Schmidt projections, and if the algorithm tolerates any degree of approximate consistency, whether relative or exact, then the DHC fails to prevent further projections arbitrarily soon after $t=0$, introducing histories with probabilities arbitrarily close to zero. Alternatively, if the algorithm treats such projections by a limiting process, then generically all the Schmidt projections at $t=0$ are applied, producing histories of zero probability. Similar problems would generally arise with repeated projections at later times, if later projections occur at all.

There would be no compelling reason to reject an algorithm which generates unexpected histories of arbitrarily small or zero probability, so long as physically sensible histories, of total probability close to one, are also generated. However, as we note in Sec. IV B below and will see later in the analysis of a physical example, this is hard to arrange. We therefore also consider below several ways in which small probability histories might be prevented: (1) The initial state could be chosen so that it does not precisely lie in the null space of any Schmidt projection. (See Sec. IV A.) (2) An initial set of projections could somehow be chosen, independent of the Schmidt projections, and with the property that for every Schmidt projection at time zero there is at least one initial history not in its null space. (See Sec. IV C.) (3) The algorithm could forbid zero probability histories by fiat and require that the selected projective decompositions form an exactly consistent set. It could then prevent small probability histories from occurring by excluding any projective decomposition $\sigma(t)$ from the selected set if $\sigma(t)$ belongs to a continuous family of decompositions, defined on some semiopen interval $(t-\epsilon,t]$, which satisfy the other selection criteria. (See Sec. IV D.)

(4) A parametrized nontriviality criterion could be used. (See Sec. IV E.) (5) Some combination of parametrized criteria for approximate consistency and nontriviality could be used. (See Sec. IV F.) We will see though, in this section and the next, that each of these possibilities leads to difficulties.

A. Choice of initial state

In the usual description of experimental situations, $\mathcal{H}_1$ describes the system degrees of freedom, $\mathcal{H}_2$ those of the apparatus (and/or an environment), and the initial state is a pure Schmidt state of the form $|\psi\rangle=|\psi_1\rangle_1\otimes|\psi_2\rangle_2$. According to this description, probabilistic events occur only after the entanglement of system and apparatus by the measurement interaction. It could, however, be argued that, since states can never be prepared exactly, we can never ensure that the system and apparatus are precisely uncorrelated, and the initial state is more accurately represented by $|\psi\rangle = |\psi_1\rangle_1 \otimes |\psi_2\rangle_2 + \gamma|\phi\rangle$, where γ is small and $|\phi\rangle$ is a vector in the total Hilbert space chosen randomly subject to the constraint that $\langle\psi|\psi\rangle = 1$. A complete set of Schmidt projections $\{P_n\}$, with $P_n|\psi\rangle \neq 0$ for all n, is then generically defined at $t=0$, and any Schmidt projection algorithm which begins by selecting all initial Schmidt projections of nonzero probability will include all of the P_n.

An obvious problem here, if relative criteria for approximate consistency and nontriviality are used to identify subsequent projections, is that the small probability initial histories constrain the later projections just as much as the large probability history which corresponds, approximately, to the Schmidt state $|\psi_1\rangle_1 \otimes |\psi_2\rangle_2$ and which is supposed to reproduce standard physical descriptions of the course of the subsequent experiment. If a branch-dependent selection algorithm is used, a relative nontriviality criterion will not cause the small probability initial histories to constrain the projections selected later on the large probability branch, but a relative approximate consistency criterion still will.

There seems no reason to expect the projections which reproduce standard descriptions to be approximately consistent extensions of the set defined by the initial Schmidt projections, and, hence, no reason to expect to recover standard physics from a Schmidt projection algorithm. When we consider a simple model of a measurement interaction in the next section we will see that, indeed, the initial projections fail to extend to a physically natural consistent set.

If absolute criteria are used, on the other hand, we would expect either that essentially the same problem arises, or that the small probability histories do not constrain the projections subsequently allowed and, hence, in particular, do not solve the problems associated with repeated projections, depending whether the probability of the unphysical histories is large or small relative to the parameters δ and ϵ^2.

B. Including null histories

If the initial state is Schmidt pure, or more generally does not define a maximal rank Schmidt decomposition, a full set of Schmidt projections can, nonetheless, generically be defined at $t=0$ — which we take to be the start of the interaction — by taking the limit of the Schmidt projections as $t \to 0^+$. The normalized histories corresponding to the projections of zero probability weight can then be defined as above, if the relevant limits exist, and used to constrain the subsequent projections in any algorithm involving relative criteria. Again, though, there seems no reason to expect these constraints to be consistent with standard physical descriptions.

C. Redefining the initial conditions

The projections selected at $t=0$ could, of course, be selected using quite different principles from those used in the selection of later projections. By choosing initial projections which are not consistently extended by any of the decompositions defined by Schmidt projections at times near $t=0$, we can certainly prevent any immediate reprojection occurring in Schmidt selection algorithms. We know of no compelling theoretical argument against incorporating projections into the initial conditions, but have found no natural combination of initial projections and a Schmidt projection selection algorithm that generally selects physically interesting sets.

D. Exact consistency and a nontriviality criterion

Since many of the problems above arise from immediate reprojections, it seems natural to look at rules which prevent zero probability histories. The simplest possibility is to impose precisely this constraint, together with exact consistency and the rules that (i) only one decomposition can be selected at any given time and (ii) no projective decomposition can be selected at time t if it belongs to a continuous family of projections $\sigma(t)$, whose members would, but for this rule, be selected at times lying in some interval $(t-\epsilon,t]$. This last condition means that the projections selected at $t=0$ are precisely those initially chosen and that no further projections occur in the neighborhood of $t=0$. Unfortunately, as the model studied later illustrates, it also generally prevents physically sensible projective decompositions being selected at later times. If it is abandoned, however, and if the initial state $|\psi\rangle$ is a pure Schmidt state, then further projections will be selected as soon as the interaction begins: in other words, at times arbitrarily close to $t=0$. Again, these projections are generally inconsistent with later physically natural projections. On the other hand, if $|\psi\rangle$ is Schmidt impure, this is generally true of the initial projections. All of these problems also arise in the case of branch-dependent set selection algorithms.

E. Exact consistency and a parametrized nontriviality criterion

Another apparently natural possibility is to require exact consistency together with one of the parametrized nontriviality criteria (2.4) or (2.5), rather than simply forbidding zero probability histories. *A priori*, there seem no obvious problems with this proposal but, again, we will see that it gives unphysical answers in the model analyzed below, whether branch-dependent or branch-independent selection algorithms are considered.

F. Approximate consistency and a parametrized nontriviality criterion

There are plausible reasons, apart from the difficulties of other proposals, for studying algorithms which use approximate consistency and parametrized nontriviality. The following comments apply to both branch-dependent and branch-independent algorithms of this type.

Physically interesting sets of projective decompositions — for example, those characterizing the pointer states of an apparatus after each of a sequence of measurements — certainly form a set which is consistent to a very good approximation. Equally, in most cases successive physically interesting decompositions define nontrivial extensions of the set defined by the previous decompositions: if the probability of a measurement outcome is essentially zero then, it might plausibly be argued, it is not essential to include the outcome in the description of the history of the system. Moreover, a finite nontriviality parameter δ ensures that, after a Schmidt projective decomposition is selected at time t, there is a finite time interval $[t,t+\Delta t]$ before a second decomposition can be chosen. One might hope that, if the parameters are well chosen, the Schmidt projective decompositions at the end of and after that interval will no longer define an approximately consistent extension unless and until they correspond to what would usually be considered as the result of a measurement-type interaction occurring after time t. While, on this view, the parameters ϵ and δ are artificial, one might also hope that they might be eliminated by letting them tend to zero in a suitable limit.

However, as we have already mentioned, in realistic physical situations we should not necessarily expect any sequence of Schmidt projective decompositions to define an exactly consistent set of histories. When the Schmidt projections correspond, say, to pointer states, the off-diagonal terms of their decoherence matrix typically decay exponentially, vanishing altogether only in the limit of infinite time separation [24–35]. An algorithm which insists on exact consistency, applied to such situations, will fail to select any projective decompositions beyond those initially selected at $t=0$ and so will give no historical description of the physics. We therefore seem forced, if we want to specify a Schmidt projection set selection algorithm mathematically, to introduce a parameter ϵ and to accept sets which are approximately consistent to order ϵ and then, in the light of the preceding discussion, to introduce a nontriviality parameter δ in order to try to prevent unphysical projective decompositions being selected shortly after $t=0$. This suggests, too, that the best that could be expected in practice from an algorithm which uses a limit in which ϵ and δ tend to zero is that the resulting set of histories describes a series of events whose time separations tend to infinity.

A parameter-dependent set selection algorithm, of course, leaves the problem of which values the parameters should take. One might hope, at least, that there is a range of values for ϵ and δ over which the selected set varies continuously and has essentially the same physical interpretation. An immediate problem here is that, if the first projective decomposition selected after $t=0$ defines a history which only just satisfies the nontriviality condition, the decomposition will, once again, have no natural physical interpretation and will generally be inconsistent with the physically natural decompositions which occur later. We will see that, in the simple model considered below, this problem cannot be avoided with an absolute consistency criterion.

Suppose now that we impose the absolute nontriviality condition that all history probabilities must be greater than δ together with the relative approximate consistency criterion that the modulus of all DHC terms is less than ϵ. The parameters ϵ and δ must be chosen so that these projections stop being approximately consistent before they become nontrivial, otherwise projections will be made as soon as they produce histories of probability exactly δ, in which case the nontriviality parameter, far from eliminating unphysical histories, would be responsible for introducing them.

Let t_ϵ denote the latest time that the extension with projection $P_k(t)$ is approximately consistent and t_δ the earliest time at which the extension is nontrivial. We see from Eq. (4.1) that, to lowest order in t,

$$t_\delta = \sqrt{\delta}\|\dot{P}_k P_n|\psi\rangle\|^{-1}, \tag{4.2}$$

$$t_\epsilon = \epsilon \frac{\|P_m|\psi\rangle\|\|\dot{P}_k P_n|\psi\rangle\|}{|\langle\psi|P_m \dot{P}_k^2 P_n|\psi\rangle|}. \tag{4.3}$$

$t_\delta > t_\epsilon$ implies

$$\sqrt{\delta}|\langle\psi|P_m \dot{P}_k^2 P_n|\psi\rangle| > \epsilon\|P_m|\psi\rangle\|\|\dot{P}_k P_n|\psi\rangle\|^2. \tag{4.4}$$

Thus we require $\delta > \epsilon^2$, up to model-dependent numerical factors: this, of course, still holds if we use a relative nontriviality criterion rather than an absolute one.

This gives, at least, a range of parameters in which to search for physically sensible consistent sets, and over which there are natural limits — for example, $\lim_{\delta\to 0}\lim_{\epsilon\to 0}$. We have, however, as yet only looked at some model-independent problems which arise in defining suitable set selection rules. In order to gain some insight into the physical problems, we look next at a simple model of system-environment interactions.

V. A SIMPLE SPIN MODEL

We now consider a simple model in which a single spin-half particle, the system, moves past a line of spin-half particles, the environment, and interacts with each in turn. This can be understood as modeling either a series of measurement interactions in the laboratory or a particle propagating through space and interacting with its environment. In the first case the environment spin-half particles represent pointers for a series of measuring devices, and in the second they could represent, for example, incoming photons interacting with the particle.

Either way, the model omits features that would generally be important. For example, the interactions describe idealized sharp measurements — at best a good approximation to real measurement interactions, which are always imperfect. The environment is represented initially by the product of N-particle states, which are initially unentangled either with the system or each other. The only interactions subsequently considered are between the system and the environment particles, and these interactions each take place in finite time. We assume too, for most of the following discussion, that the interactions are distinct: the kth is complete before the $(k+1)$th begins. It is useful, though, even in this highly idealized example, to see the difficulties which arise in finding set selection algorithms: we take the success of a set selection algorithm here to be a necessary, but not sufficient, condition for it to be considered as a serious candidate.

A. Definition of the model

We use a vector notation for the system states, so that if $\hat{\mathbf{u}}$ is a unit vector in R^3 the eigenstates of $\boldsymbol{\sigma}\cdot\hat{\mathbf{u}}$ are represented by $|\pm\hat{\mathbf{u}}\rangle$. With the pointer state analogy in mind, we use the basis $\{|\uparrow\rangle_k, |\downarrow\rangle_k\}$ to represent the kth environment particle state, together with the linear combinations $|\pm\rangle_k = (|\uparrow\rangle_k \pm |\downarrow\rangle_k)/\sqrt{2}$. We compactify the notation by writing environment states as single kets, so that, for example, $|\uparrow\rangle_1 \otimes \cdots \otimes |\uparrow\rangle_n$ is written as $|\uparrow_1 \cdots \uparrow_n\rangle$, and we take the initial state $|\psi(0)\rangle$ to be $|\mathbf{v}\rangle \otimes |\uparrow_1 \cdots \uparrow_n\rangle$.

The interaction between the system and the kth environment particle is chosen so that it corresponds to a measurement of the system spin along the $\hat{\mathbf{u}}_k$ direction, so that the states evolve as follows:

$$|\hat{\mathbf{u}}_k\rangle\otimes|\uparrow\rangle_k\rightarrow|\mathbf{u}_k\rangle\otimes|\uparrow\rangle_k, \tag{5.1}$$

$$|-\hat{\mathbf{u}}_k\rangle\otimes|\uparrow\rangle_k\rightarrow|-\hat{\mathbf{u}}_k\rangle\otimes|\downarrow\rangle_k. \tag{5.2}$$

A simple unitary operator that generates this evolution is

$$U_k(t)=P(\hat{\mathbf{u}}_k)\otimes I_k+P(-\hat{\mathbf{u}}_k)\otimes e^{-i\theta_k(t)F_k}, \tag{5.3}$$

where $P(\hat{\mathbf{x}})=|\hat{\mathbf{x}}\rangle\langle\hat{\mathbf{x}}|$ and $F_k=i|\downarrow\rangle_k\langle\uparrow|_k-i|\uparrow\rangle_k\langle\downarrow|_k$. Here $\theta_k(t)$ is a function defined for each particle k, which varies from 0 to $\pi/2$ and represents how far the interaction has progressed. We define $P_k(\pm)=|\pm\rangle_k\langle\pm|_k$, so that $F_k=P_k(+)-P_k(-)$.

The Hamiltonian for this interaction is thus

$$H_k(t)=i\dot{U}_k(t)U_k^{\dagger}(t)=\dot{\theta}_k(t)P(-\hat{\mathbf{u}}_k)\otimes F_k, \tag{5.4}$$

in both the Schrödinger and Heisenberg pictures. We write the extension of U_k to the total Hilbert space as

$$V_k=P(\hat{\mathbf{u}}_k)\otimes I_1\otimes\cdots\otimes I_n+P(-\hat{\mathbf{u}}_k)\otimes I_1\otimes\cdots\otimes I_{k-1}\otimes e^{-i\theta_k(t)F_k}\otimes I_{k+1}\otimes\cdots\otimes I_n. \tag{5.5}$$

We take the system particle to interact initially with particle 1 and then with consecutively numbered ones, and there is no interaction between environment particles, so that the evolution operator for the complete system is

$$U(t)=V_n(t)\cdots V_1(t), \tag{5.6}$$

with each factor affecting only the Hilbert spaces of the system and one of the environment spins.

We suppose, finally, that the interactions take place in disjoint time intervals and that the first interaction begins at $t=0$, so that the total Hamiltonian is simply

$$H(t)=\sum_{k=1}^{n}H_k(t), \tag{5.7}$$

and we have that $\theta_1(t)>0$ for $t>0$ and that, if $\theta_k(t)\in(0,\pi/2)$, then $\theta_i(t)=\pi/2$ for all $i<k$ and $\theta_i(t)=0$ for all $i>k$.

B. Classification of Schmidt projection consistent sets in the model

For generic choices of the spin measurement directions, in which no adjacent pair of the vectors $\{\hat{\mathbf{v}},\hat{\mathbf{u}}_1,\ldots,\hat{\mathbf{u}}_n\}$ is parallel or orthogonal, the exactly consistent branch-dependent sets defined by the Schmidt projections onto the system space can be completely classified in this model. The following classification theorem is proved in Ref. [37]: *Theorem*. In the spin model defined above, suppose that no adjacent pair of the vectors $\{\hat{\mathbf{v}},\hat{\mathbf{u}}_1,\ldots,\hat{\mathbf{u}}_n\}$ is parallel or orthogonal. Then the histories of the branch-dependent consistent sets defined by Schmidt projections take one of the following forms. (i) A series of Schmidt projections made at times between the interactions — i.e., at times t such that $\theta_k(t)=0$ or $\pi/2$ for all k. (ii) A series as in (i), made at times $t_1,\ldots,t_n$, together with one Schmidt projection made at any time t during the interaction immediately preceding the last projection time t_n. (iii) A series as in (i), together with one Schmidt projection made at any time t during an interaction taking place after t_n. Conversely, any branch-dependent set, each of whose histories takes one of the forms (i)–(iii), is consistent. We assume below that the set of spin measurement directions satisfies the condition of the theorem: since this can be ensured by an arbitrarily small perturbation, this seems physically reasonable. The following sections explain, with the aid of this classification, the results of various set selection algorithms applied to the model.

VI. APPLICATION OF SELECTION ALGORITHMS TO THE SPIN MODEL

We can define a natural consistent set which reproduces the standard historical account of the physics of the separated interaction spin model by selecting the Schmidt projections at all times between each successive spin measurement. A set of this type ought to be produced by a good set selection algorithm, either as the selected set itself or, perhaps, a subset. Sections VI A, VI B, and VI C describe the results actually produced by various set selection algorithms applied to the spin model. All of these algorithms are dynamical, in the sense that the decision whether to select projections at time t, and if so which, depends only on the evolution of the state vector up to time t. Sections VI D and VI E discuss how these results are affected by altering the initial conditions of the model. In Sec. VI F we consider a selection algorithm which is quasidynamical, in the sense that the decisions at time t depend on the evolution of the state vector up to and just beyond t. We summarize our conclusions in Sec. VI G.

A. Exact limit DHC consistency

Since any projective decomposition at time t defines an exactly consistent set when there is only one history up to that time, a Schmidt projection selection algorithm without a nontriviality criterion will immediately make a projection. The normalized histories are defined as

$$\lim_{t\rightarrow 0}P_{\pm}(t)|\psi\rangle/\|P_{\pm}(t)|\psi\rangle\|, \tag{6.1}$$

where $P_{\pm}(t)$ denotes the Schmidt projections at time t. The Schmidt states to first order in $\omega=\theta_1(t)$ are

$$|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle-i\omega/2(1-\hat{\mathbf{u}}_1\cdot\hat{\mathbf{v}})|\hat{\mathbf{v}}\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle \tag{6.2}$$

and

$$|\hat{\mathbf{u}}_1\wedge\hat{\mathbf{v}}||-\hat{\mathbf{v}}\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle+i\omega/2\sqrt{\left(\frac{1-\hat{\mathbf{u}}_1\cdot\hat{\mathbf{v}}}{1+\hat{\mathbf{u}}_1\cdot\hat{\mathbf{v}}}\right)^{1/2}}|-\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle, \tag{6.3}$$

so the normalized histories are

$$\{|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\uparrow_2\cdots\uparrow_n\rangle,|-\hat{\mathbf{v}}\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle\}. \tag{6.4}$$

The limit DHC term for one projection at time 0 and another during interaction k at time t is

$$\begin{aligned} &\cos\phi \quad \text{for } k=1,\\ &\frac{\sin^2\phi|\hat{\mathbf{u}}_1\cdot\hat{\mathbf{u}}_2||\hat{\mathbf{v}}\wedge(\hat{\mathbf{u}}_1\wedge\hat{\mathbf{u}}_2)|}{N_2(\phi)[1-(\hat{\mathbf{v}}\cdot\hat{\mathbf{u}}_1)^2N_2^2(\phi)]^{1/2}} \quad \text{for } k=2,\\ &\frac{\lambda_{2(k-1)}N_k(\phi)|\hat{\mathbf{v}}\wedge(\hat{\mathbf{u}}_1\wedge\hat{\mathbf{u}}_2)|}{[1-\lambda_{0(k-1)}^2N_k^2(\phi)]^{1/2}} \quad \text{for } k>2, \end{aligned} \tag{6.5}$$

where $\phi=\theta_k(t)$. Here we define

$$\lambda_{ij}=\prod_{k=i}^{j-1}|\hat{\mathbf{u}}_k\cdot\hat{\mathbf{u}}_{k+1}|, \tag{6.6}$$

with the convention that $\lambda_{ij}=1$ for $j\leqslant i$, and

$$N_k(\phi)=|A_k(\phi)\hat{\mathbf{u}}_{k-1}|, \tag{6.7}$$

where

$$A_k(\phi)=P(\hat{\mathbf{u}}_k)+\cos\phi\overline{P(\hat{\mathbf{u}}_k)}, \tag{6.8}$$

where $P(\hat{\mathbf{u}}_k)$ is the projection onto the vector $\hat{\mathbf{u}}_k$ in R^3, and $\overline{P(\hat{\mathbf{u}}_k)}$ its complement.

Whether the algorithm is taken to be branch-dependent or branch-independent, the only future Schmidt projections which are consistent with the initial projections are thus those between the first and second interactions, and the projections selected will be at the end of the first interaction. The state at this time is

$$\begin{aligned}|\psi(1)\rangle=&|\hat{\mathbf{u}}_1\rangle\langle\hat{\mathbf{u}}_1|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle+|-\hat{\mathbf{u}}_1\rangle\langle-\hat{\mathbf{u}}_1|\mathbf{v}\rangle\\ &\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle.\end{aligned} \tag{6.9}$$

The time evolved histories are

$$\begin{aligned}|h_1(t)\rangle=&|\hat{\mathbf{u}}_1\rangle\langle\hat{\mathbf{u}}_1|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle+|-\hat{\mathbf{u}}_1\rangle\langle-\hat{\mathbf{u}}_1|\hat{\mathbf{v}}\rangle\\ &\otimes|\uparrow_1\ldots\uparrow_n\rangle\end{aligned} \tag{6.10}$$

$$\begin{aligned}|h_2(t)\rangle=&|\hat{\mathbf{u}}_1\rangle\langle\hat{\mathbf{u}}_1|-\hat{\mathbf{v}}\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle-|-\hat{\mathbf{u}}_1\rangle\langle-\hat{\mathbf{u}}_1|-\hat{\mathbf{v}}\rangle\\ &\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle,\end{aligned} \tag{6.11}$$

so the new normalized histories are

$$\{|\hat{\mathbf{u}}_1\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle,|\hat{\mathbf{u}}_1\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle, \tag{6.12}$$

$$|-\hat{\mathbf{u}}_1\rangle\otimes|\uparrow_1\rangle\cdots\uparrow_n\rangle,|-\hat{\mathbf{u}}_1\rangle\otimes|\downarrow_1\uparrow_2\cdots\uparrow_n\rangle\}. \tag{6.13}$$

Since no future Schmidt projections are consistent with those selected, the algorithm clearly fails to produce the correct set.

B. Exact consistency and nontriviality

Suppose that, instead of using the limit DHC, we consider only sets defined by decompositions at different times and require exact consistency. As explained earlier, without a nontriviality criterion this leads to an ill-defined algorithm: the initial projections at $t=0$ produce a null history, and the Schmidt projections at all times greater than zero are consistent with these initial projections, so that no minimal nonzero time is selected by the algorithm.

Introducing a nontriviality criterion removes this problem. Suppose, for example, we impose the absolute criterion $D_{\alpha\alpha}\geqslant\delta$ for all histories α. Since any physically reasonable δ would have to be extremely small, let us assume $\delta\ll|\hat{\mathbf{u}}_i\wedge\hat{\mathbf{u}}_j|$. The first projections after $t=0$ are then selected at the first time when $D_{\alpha\alpha}=\delta$, which occurs during the first interaction. Whether or not branch-dependent projections are allowed, the only other Schmidt projections which can consistently be selected then take place at the end of the first interaction, and it again follows from the classification theorem that no further projections can take place. Again, by making projections too early, this algorithm fails to produce the correct consistent set.

A suitably large value of δ could ensure that no extension will occur until later interactions but, generically, the first extension made after $t=0$ will take place during an interaction rather than between interactions, and the classification theorem ensures that no more than four histories will ever be generated.

The same problems arise if the nontriviality criterion is taken to be relative rather than absolute. It is possible to do better by fine tuning the parameters: for example, if branch-independent histories are used, a relative nontriviality criterion is imposed and $\delta=(1-|\hat{\mathbf{u}}_k\cdot\hat{\mathbf{u}}_{k+1}|)/2$ for all $k=0,\ldots,n-1$, then projections will occur at the end of each interaction producing the desired set of histories. This, though, is clearly not a satisfactory procedure.

C. Approximate consistency and nontriviality

One might wonder if these problems can be overcome by relaxing the standards of consistency, since a projection at a very small time will be approximately consistent — according to absolute measures of approximate consistency, at least — with projections at the end of the other interactions. However, this approach too runs into difficulties, whether relative or exact criteria are used.

Consider first a branch-dependent set selection algorithm which uses the absolute nontriviality criterion $D_{\alpha\alpha}\geqslant\delta$ for all α, and the absolute criterion for approximate consistency $|D_{\alpha\beta}|\leqslant\epsilon$ for all $\alpha\neq\beta$. No history with probability less than 2δ will thus be extended, since if it were one of the resultant histories would have probability less than δ.

Any history α with a probability less than or equal to ϵ^2 will automatically be consistent with any history β according to this criterion, since $|D_{\alpha\beta}|\leqslant(D_{\alpha\alpha}D_{\beta\beta})^{1/2}\leqslant(\epsilon^2\times1)^{1/2}=\epsilon$. Therefore if $\delta\leqslant\epsilon^2$ then histories of probability δ will be consistent with all other histories. The first projection after $t=0$ will be made as soon as the nontriviality criterion permits, when the largest Schmidt eigenvalue is $1-\delta$. Other projections onto the branch defined by the largest probability history will follow similarly as the Schmidt projections evolve. The final set of histories after n projections will thus consist of one history with probability $1-n\delta$ and n histories with probability δ — clearly far from the standard picture.

Suppose now that $\delta > \epsilon^2$. The probabilities for histories with projection in the first interval, at time t with $\theta_1(t) = \omega$, are

$$\tfrac{1}{2}[1 - \sqrt{1 - \sin^2\omega |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^2}]. \tag{6.14}$$

The first projection will therefore be made when

$$\theta_1(t) = \omega \simeq 2\sqrt{\delta} |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^{-1}, \tag{6.15}$$

producing histories of probabilities δ and $(1-\delta)$. The next projections selected will necessarily extend the history of probability $(1-\delta)$, since the absolute nontriviality criterion forbids further extensions of the other history. We look first at projections taking place at a later time t', with $\theta_1(t') = \phi$, during the first interaction, and define $N_1(\omega) = (1 - \sin^2\omega |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^2)^{1/2}$. Of the probabilities of the extended histories, the smaller is

$$\begin{aligned} &\tfrac{1}{4}[1 + N_1(\omega)]\{1 - N_1^{-1}(\omega) N_1^{-1}(\phi)[(\hat{\mathbf{v}} \cdot \hat{\mathbf{u}}_1)^2 \\ &\quad + \cos\phi \,\cos\omega \,\cos(\phi - \omega)|\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^2]\} \\ &= \tfrac{1}{4}|\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^2 (\omega - \phi)^2 [1 + O(\omega) + O(\phi)]. \end{aligned} \tag{6.16}$$

Therefore this extension will be nontrivial when

$$\phi \simeq \omega + 2\sqrt{\delta} |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^{-1} = 4\sqrt{\delta} |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^{-1} + O(\delta). \tag{6.17}$$

The largest off-diagonal element in the decoherence matrix for this extension is

$$\tfrac{1}{4} N_1^{-1}(\phi) |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|^2 \cos\phi \,\sin\omega \,\sin(\phi - \omega) = \delta + O(\delta^3). \tag{6.18}$$

Unless $\delta > \epsilon$, then, this extension is selected together, again, with a series of further extensions generating small probability histories.

Suppose now that $\delta > \epsilon$. The term on the left-hand side of Eq. (6.18) increases monotonically until $\phi \simeq \pi/4$, and then decreases again as $\phi \to \pi/2$. For $\phi \simeq \pi/2$, it equals

$$\tfrac{1}{2}\sqrt{\delta} \,\cos\phi |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1| |\hat{\mathbf{v}} \cdot \hat{\mathbf{u}}_1|^{-1} [1 + O(\cos\phi)]. \tag{6.19}$$

Hence, the approximate consistency criterion is next satisfied when

$$\phi = \pi/2 - \frac{2\epsilon |\hat{\mathbf{v}} \cdot \hat{\mathbf{u}}_1|}{\sqrt{\delta} |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1|} + O(\epsilon^2/\delta), \tag{6.20}$$

and this extension is also nontrivial unless $\hat{\mathbf{v}}$ and $\hat{\mathbf{u}}_1$ are essentially parallel, which we assume not to be the case. In this case, then, projections are made towards the beginning and towards the end of the first interaction, and a physically reasonable description of the first measurement emerges.

This description, however, cannot generally be consistently extended to describe the later measurements. If we consider the set of histories defined by the Schmidt projections at time t, given by Eq. (6.15) above, together with the Schmidt projections at time t'' such that $\theta_k(t'') = \phi$ for some $k > 1$, we find that the largest off-diagonal decoherence matrix element is

$$\tfrac{1}{2}\sqrt{\delta} \lambda_{2(k-1)} N_k(\phi) |\hat{\mathbf{v}} \wedge \hat{\mathbf{u}}_1| |\hat{\mathbf{v}} \wedge (\hat{\mathbf{u}}_1 \wedge \hat{\mathbf{u}}_2)| [1 + O(\sqrt{\omega})]. \tag{6.21}$$

Since we have chosen $\epsilon < \delta$ to prevent multiple projections, and since the other terms are not small for generic choices of the vectors, the set generally fails to satisfy the criterion for approximate consistency. Note, however, that if all the measurement directions are apart by an angle greater than equal to some $\theta > 0$, then $\lambda_{2(k-1)}$ decreases exponentially with k. After a large enough number [of order $O(-\ln\epsilon)$] of interactions have passed the algorithm will select a consistent extension, and further consistent extensions will be selected at similar intervals. The algorithm does thus eventually produce nontrivial consistent sets, though the sets produced do not vary smoothly with ϵ and do not describe the outcome of most of the spin measurements.

The reason this algorithm, and similar algorithms using approximate consistency criteria, fail is easy to understand. The off-diagonal decoherence matrix component in a set defined by the Schmidt projections at time t together with Schmidt projections during later interactions is proportional to $\sin\omega \cos\omega$, together with terms which depend on the angles between the vectors. The decoherence matrix component for a set defined by the projections at time t, together with Schmidt projections at a second time t' soon afterwards is proportional to $\sin^2(\phi - \omega)$. The obstacle to finding nontriviality and approximate consistency criteria that can prevent reprojections in the first interaction period, yet allow interactions in later interaction periods, is that when $(\phi - \omega)$ is small the second term is generally smaller than the first.

Using a relative nontriviality criterion makes no difference, since the branchings we consider are from a history of probability close to 1, and using the DHC instead of an absolute criterion for approximate consistency only worsens the problem of consistency of later projections, since the DHC alters Eq. (6.21) by a factor of $1/\sqrt{\delta}$, leaving a term which is generically of order unity. Requiring branch independence, of course, only worsens the problems.

D. Nonzero initial Schmidt eigenvalues

We now reconsider the possibility of altering the initial conditions in the context of the spin model. Suppose first that the initial state is not Schmidt degenerate. For example, as the initial normalized histories are $\{|\hat{\mathbf{v}}\rangle \otimes |\uparrow_1 \cdots \uparrow_n\rangle, |-\hat{\mathbf{v}}\rangle \otimes |\downarrow_1 \uparrow_2 \cdots \uparrow_n\rangle\}$ a natural ansatz is

$$|\psi(0)\rangle = \sqrt{p_1} |\hat{\mathbf{v}}\rangle \otimes |\uparrow_1 \ldots \uparrow_n\rangle + \sqrt{p_2} |-\hat{\mathbf{v}}\rangle \otimes |\downarrow_1 \uparrow_2 \cdots \uparrow_n\rangle. \tag{6.22}$$

Consider now a set of histories defined by Schmidt projections at times 0 and a time t during the kth interaction for $k > 2$, so that $\theta_1(t) = \theta_2(t) = \pi/2$. The moduluses of the nonzero off-diagonal elements of the decoherence matrix are

$$\tfrac{1}{2}\sqrt{p_1 p_2} |\hat{\mathbf{v}} \wedge [\hat{\mathbf{u}}_1 \wedge \mathbf{u}_2]| \lambda_{2k}. \tag{6.23}$$

Generically, these off-diagonal elements are not small, so that the perturbed initial conditions prevent later physically sensible projections from being selected.

E. Specifying initial projections

We consider now the consequence of specifying initial projections in the spin model. Suppose the initial projections are made using $P(\pm\hat{\mathbf{h}})\otimes I_E$. The modulus of the nonzero off-diagonal elements of the decoherence matrix for a projection at time t during interaction k, for $k>2$, is

$$\tfrac{1}{4}|\hat{\mathbf{h}}\wedge\hat{\mathbf{v}}||\hat{\mathbf{h}}\wedge\hat{\mathbf{u}}_1|\lambda_{1(k-1)}N_k[\theta_k(t)], \tag{6.24}$$

and again we see that physically natural projections generically violate the approximate consistency criterion.

It might be argued that the choice of initial projections given by $\hat{\mathbf{h}}=\pm\hat{\mathbf{v}}$ is particularly natural. This produces an initial projection on to the initial state, with the other history undefined unless a limiting operation is specified. If the limit of the normalized histories for initial projections $\hat{\mathbf{h}}'\to\hat{\mathbf{h}}$ is taken, the normalized histories are simply $|\pm\hat{\mathbf{h}}\rangle$. If an absolute consistency criterion is used the null history will not affect future projections and the results will be the same as if no initial projection had been made. If, on the other hand, the limit DHC is used then the consistency criterion is the same as for general $\hat{\mathbf{h}}$, that is, $\hat{\mathbf{h}}$ must be parallel to $\hat{\mathbf{u}}_1$. This requires that the initial conditions imposed at $t=0$ depend on the axis of the first measurement, and still fails to permit a physically natural description of later measurements.

F. A quasidynamical algorithm

For completeness, we include here an algorithm which, though not strictly dynamical, succeeds in selecting the natural consistent set to describe the spin model. In the spin model as defined, it can be given branch-dependent or branch-independent form and selects the same set in either case. In the branch-independent version, the Schmidt projections are selected at time t provided that they define an exactly consistent and nontrivial extension of the set defined by previously selected projections *and* that this extension can itself be consistently and nontrivially extended by the Schmidt projections at time $t+\epsilon$ for every sufficiently small $\epsilon>0$.[4] In the branch-dependent version, the second condition must hold for at least one of the newly created branches of nonzero probability in the extended set.

It follows immediately from the classification theorem that no Schmidt projections can be selected during interactions, since no exactly consistent set of Schmidt projections includes projections at two different times during interactions. The theorem also implies that the Schmidt projections are selected at the end of each interval between interactions, so that the selected set describes the outcomes of each of the measurements.

[4]Alternatively, a limiting condition can be used.

G. Comments

The simple spin model used here illustrates the difficulty in encoding our physical intuition algorithmically. The model describes a number of separated interactions, each of which can be thought of as a measurement of the system spin. There is a natural choice of consistent set, given by the projections onto the system spin states along the measured axes at all times between each of the measurements.[5] This set does indeed describe the physics of the system as a series of measurement events and assigns the correct probabilities to those events. Moreover, the relevant projections are precisely the Schmidt projections.

We first considered a series of Schmidt projection set selection algorithms which are dynamical, in the sense that the projections selected at time t depend only on the physics up to that time. Despite the simplifying features of the models, it seems very hard to find a dynamical Schmidt projection set selection algorithm which selects a physically natural consistent set and which is not specifically adapted to the model in question.

It might be argued that the very simplicity of the model makes it an unsuitable testing ground for set selection algorithms. It is certainly true that more realistic models would generally be expected to allow fewer exactly consistent sets built from Schmidt projections: it is not at all clear that any nontrivial exactly consistent sets of this type should be expected in general. However, we see no way in which all the problems encountered in our discussion of dynamical set selection algorithms can be evaded in physically realistic models.

We have, on the other hand, seen that a simple quasidynamical set selection algorithm produces a satisfactory description of the spin model. However, as we explain in Sec. VII, there is another quite general objection which applies both to dynamical set selection algorithms and to this quasidynamical algorithm.

VII. THE PROBLEM OF RECOHERENCE

The set selection algorithms above rely on the decoherence of the states of one subsystem through their interactions with another. This raises another question: what happens when decoherence is followed by recoherence?

For example, consider a version of the spin model in which the system particle initially interacts with a single environment particle as before, and then reencounters the particle, reversing the interaction, so that the evolution takes the form

$$\begin{aligned} &a_1|\hat{\mathbf{u}}\rangle\otimes|\uparrow_1\rangle+a_2|-\hat{\mathbf{u}}\rangle\otimes|\uparrow_1\rangle\\ &\quad\to a_1|\hat{\mathbf{u}}\rangle\otimes|\uparrow_1\rangle+a_2|-\hat{\mathbf{u}}\rangle\otimes|\downarrow_1\rangle\\ &\quad\to a_1|\hat{\mathbf{u}}\rangle\otimes|\uparrow_1\rangle+a_2|-\hat{\mathbf{u}}\rangle\otimes|\uparrow_1\rangle, \end{aligned} \tag{7.1}$$

generated by the unitary operator

[5]Strictly speaking, there are many equivalent consistent sets, all of which include the Schmidt projections at some point in time between each measurement and at no time during measurements, and all of which give essentially the same physical picture.

$$U(t)=P(\hat{\mathbf{u}})\otimes I+P(-\hat{\mathbf{u}})\otimes e^{-i\theta(t)F}, \tag{7.2}$$

where

$$\theta(t)=\begin{cases} t & \text{for } 0\leq t\leq \pi/2 \\ \pi/2 & \text{for } \pi/2\leq t\leq \pi \\ 3\pi/2-t & \text{for } \pi\leq t\leq 3\pi/2. \end{cases} \tag{7.3}$$

We have taken for granted, thus far, that a dynamical algorithm makes selections at time t based only on the evolution of the system up to that time. Thus any dynamical algorithm which behaves sensibly, according to the criteria which we have used so far, will select a consistent set which includes the Schmidt projections at some time between $\pi/2$ and π, since during that interval the projections appear to describe the result of a completed measurement. These projections cannot be consistently extended by projections onto the initial state $a_1|\hat{\mathbf{u}}\rangle+a_2|-\hat{\mathbf{u}}\rangle$ and the orthogonal state $\overline{a_2}|\hat{\mathbf{u}}\rangle-\overline{a_1}|-\hat{\mathbf{u}}\rangle$ at time $3\pi/2$, so that the algorithm will not agree with the standard intuition that at time π the state of the system particle has reverted to its initial state. In particular, if the particle subsequently undergoes interactions of the form (5.1) with other environment particles, the algorithm cannot reproduce the standard description of these later measurements. The same problem afflicts the quasidynamical algorithm considered in Sec. VI F.

In principle, then, dynamical set selection algorithms of the type considered so far imply that, following any experiment in which exact decoherence is followed by exact recoherence and then by a probabilistic measurement of the recohered state, the standard quasiclassical picture of the world cannot generally be recovered. If the algorithms use an approximate consistency criterion — as we have argued is necessary for a realistic algorithm — then this holds true for experiments in which the decoherence and recoherence are approximate.

We know of no experiments of precisely this type. Several neutron interferometry experiments have been performed in which one or both beams interact with an electromagnetic field before recombination [38–45] and measurement. In these experiments, though, the electromagnetic field states are typically superpositions of many different number states, and are largely unaffected by the interaction, so that Eq. (7.1) is a poor model for the process.[6] Still, it seems hard to take seriously the idea that if a recoherence experiment were constructed with sufficient care it would jeopardize the quasiclassicality we observe, and we take the recoherence problem as a conclusive argument against the general applicability of the algorithms considered to date.

VIII. RETRODICTIVE ALGORITHMS

We have seen that dynamical set selection algorithms which run forwards in time generally fail to reproduce standard physics. Can an algorithm be developed for reconstructing the history of a series of experiments or, in principle, of the Universe?

[6] See, for example, Ref. [46] for a review and analysis.

A. Retrodictive algorithms in the spin model

We first look at the spin model with separated interactions and initial state

$$|\psi(0)\rangle=|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle, \tag{8.1}$$

and take the first interaction to run from $t=0$ to $t=1$, the second from $t=1$ to $t=2$, and so on. The final state, in the Schrödinger picture, is

$$|\psi(n)\rangle=\sum_{\alpha}\sqrt{p_{\boldsymbol{\alpha}}}|\alpha_n\hat{\mathbf{u}}_n\rangle\otimes|\beta_1\cdots\beta_n\rangle. \tag{8.2}$$

Here $\alpha=\{\alpha_1,\dots,\alpha_n\}$ runs over all strings of n pluses and minuses, we write $\beta_i=\uparrow$ if $\alpha_i=1$ and $\beta_i=\downarrow$ if $\alpha_i=-1$, and

$$p_{\boldsymbol{\alpha}}=2^{-n}(1+\alpha_n\alpha_{n-1}\hat{\mathbf{u}}_n\cdot\hat{\mathbf{u}}_{n-1})\cdots(1+\alpha_1\hat{\mathbf{u}}_1\cdot\hat{\mathbf{u}}_0). \tag{8.3}$$

Consider now a set selection algorithm which begins the selection process at $t=n$ and works backwards in time, selecting an exactly consistent set defined by system space Schmidt projections. The algorithm thus begins by selecting projections onto the Schmidt states $|\pm\hat{\mathbf{u}}_n\rangle$ at $t=n$. The classification theorem implies that any Schmidt projection during the time interval $[n-1,n)$ defines a consistent and nontrivial extension to the set defined by these projections. If the algorithm involves a parametrized nontriviality condition with sufficiently small nontriviality parameter δ, the next projection will thus be made as soon as the nontriviality condition is satisfied, which will be at some time $t=n-\Delta t$, where Δt is small.

If a nontriviality condition is not used but the limit DHC is used instead, then a second projection will be made at $t=n$, but the normalized path projected states will be the same (to lowest order in Δt) as for projection at $t=n-\Delta t$. The classification theorem then implies that the only possible times at which further extensions can consistently be made are $t=n-1,\dots,1$ and, if δ is sufficiently small and the measurement axes are nondegenerate, the Schmidt projections at all of these times will be selected.

In fact, this algorithm gives very similar results whether a nontriviality condition or the limit DHC is used. We use the limit DHC here for simplicity of notation. Since the Schmidt states at the end of the kth interaction are $|\pm\hat{\mathbf{u}}_k\rangle$, the histories of the selected set are indexed by strings $\{\alpha_1,\dots,\alpha_{n+1}\}$ consisting of $n+1$ pluses and minuses. The corresponding class operators are defined in terms of the Heisenberg picture Schmidt projections as

$$P_H^{\alpha_{n+1}}(n)P_H^{\alpha_n}(n)P_H^{\alpha_{n-1}}(n-1)\cdots P_H^{\alpha_1}(1). \tag{8.4}$$

Define $C_{\boldsymbol{\alpha}}=P_H^{\alpha_n}(n)\cdots P_H^{\alpha_1}(1)$. Then

$$P_H^{\alpha_{n+1}}(n)C_\alpha=C_\alpha \quad \text{if } \alpha_{n+1}=\alpha_n,$$

$$P_H^{\alpha_{n+1}}(n)C_\alpha=0 \quad \text{if } \alpha_{n+1}=-\alpha_n, \tag{8.5}$$

and to calculate the limit DHC Eq. (3.2) we note that Eq. (3.5) implies that

$$\begin{aligned}\lim_{\epsilon\to 0}\epsilon^{-1}P_H^{-\alpha_n}(n)P_H^{\alpha_n}(n-\epsilon)\cdots P_H^{\alpha_1}(1)\\ &=P_H^{-\alpha_n}(n)\dot{P}_H^{\alpha_n}(n)\cdots P_H^{\alpha_1}(1)\\ &=\dot{P}_H^{\alpha_n}(n)P_H^{\alpha_n}(n)\cdots P_H^{\alpha_1}(1)\\ &=P_H^{\alpha_n}(n)C_{\boldsymbol{\alpha}}.\end{aligned} \tag{8.6}$$

The complete set of class operators (up to multiplicative constants) is $\{C_{\boldsymbol{\alpha}},\dot{P}_H^{+}(n)C_{\boldsymbol{\alpha}}\}$ and the set of normalized histories is, therefore,

$$\{|\alpha_n\hat{\mathbf{u}}_n\rangle\otimes|\alpha\rangle,|-\alpha_n\hat{\mathbf{u}}_n\rangle\otimes|\alpha\rangle\}. \tag{8.7}$$

Of these histories, the first 2^n have probabilities $p_\alpha=2^{-n}(1+\alpha_n\alpha_{n-1}\hat{\mathbf{u}}_n\cdot\hat{\mathbf{u}}_{n-1})\ldots(1+\alpha_1\hat{\mathbf{u}}_{\mathbf{1}}\cdot\hat{\mathbf{u}}_{\mathbf{0}})$ and have a simple physical interpretation, namely, that the particle was in direction $\alpha_i\hat{\mathbf{u}}_i$ at time $t=i$, for each i from 1 to n, while the second 2^n have zero probability. Thus the repeated projections that the algorithm selects at $t=n$, while nonstandard, merely introduce probability zero histories, which need no physical interpretation. The remaining projections reproduce the standard description so that, in this example, at least, retrodictive algorithms work. While this is somewhat encouraging, the algorithm's success here relies crucially on the simple form of the classification of consistent sets in the spin model, which in turn relies on a number of special features of the model. In order to understand the behavior of retrodictive algorithms in more generality, we look next at two slightly more complicated versions of the spin model.

B. Spin model with perturbed initial state

Consider now the spin model with a perturbed initial state $|\psi\rangle+\gamma|\phi\rangle$. For generic choices of ϕ and γ, there is no nontrivial exactly consistent set of Schmidt projections, but it is easy to check that the set selected in the preceding section remains approximately consistent to order γ, in the sense that the DHC and limit DHC parameters are $O(\gamma)$.

This example, nonetheless, highlights a difficulty with the type of retrodictive algorithm considered so far. Some form of approximate consistency criterion is clearly required to obtain physically sensible sets in this example. However, there is no obvious reason to expect that there should be any parameter ϵ with the property that a retrodictive algorithm which requires approximate consistency (via the limit DHC and DHC) to order ϵ will select a consistent set whose projections are all similar to those of the set previously selected. The problem is that, given any choice of ϵ which selects the right projections at time n, the next projections selected will be at time $(n-1)+O(\gamma)$ rather than at precisely $t=n-1$. The level of approximate consistency then required to select projections at times near $n-2$, $n-3$, and so forth, depends on the projections already selected, and so depends on γ only indirectly and in a rather complicated way.

We expect that, for small γ and generic ϕ, continuous functions $\epsilon_k(\gamma,\phi)$ exist with the properties that $\epsilon_k(\gamma,\phi)\to 0$ as $\gamma\to 0$ and that some approximation to the set previously selected will be selected by a retrodictive algorithm which requires approximate consistency to order $\epsilon_k(\gamma,\phi)$ for the kth projection. Clearly, though, since the aim of the set selection program is to replace model-dependent intuition by a precise algorithmic description, it is rather unsatisfactory to have to fine tune the algorithm to fit the model in this way.

C. Delayed choice spin model

We now return to considering the spin model with an unperturbed initial state and look at another shortcoming. The interaction of the system particle with each successive environment particle takes the form of a spin measurement interaction in which the axis of each measurement, $\{\hat{\mathbf{u}}_i\}$, is fixed in advance. This is a sensible assumption when modeling a natural system-environment coupling, such as a particle propagating past a series of other particles. As a model of a series of laboratory experiments, however, it is unnecessarily restrictive. We can model experiments with an element of delayed choice simply by taking the axis $\{\hat{\mathbf{u}}_i\}$ to depend on the outcome of the earlier measurements.

If we do this, while keeping the times of the interactions fixed and nonoverlapping, the measurement outcomes can still be naturally described in terms of a consistent set built from Schmidt projections onto the system space at times $t=1,2,\ldots,n$, so long as both the Schmidt projections and the consistent set are defined to be appropriately branch dependent. Thus, let

$$|\psi(0)\rangle=|\hat{\mathbf{v}}\rangle\otimes|\uparrow_1\cdots\uparrow_n\rangle \tag{8.8}$$

be the initial state and let $P_H^{\alpha_1}(1)$, for $\alpha_1=\pm$, be the Schmidt projections onto the system space at time $t=1$. We define a branch-dependent consistent set in which these projections define the first branches and consider independently the evolution of the two states $P_H^{+}(1)|\psi(0)\rangle$ and $P_H^{-}(1)|\psi(0)\rangle$ between $t=1$ and $t=2$. These evolutions take the form of measurements about axes $\hat{\mathbf{u}}_{2;\alpha_1}$ which depend on the result of the first measurement. At $t=2$ the second measurements are complete, each branch splits again, and the subsequent evolutions of the four branches now depend on the results of the first two measurements. Similar splittings take place at each time from 1 to n, so that the axis of the mth measurement in a given branch, $\mathbf{u}_{m;\alpha_{m-1},\ldots,\alpha_1}$, depends on the outcomes $\alpha_{m-1},\ldots,\alpha_1$ of the previous $(m-1)$ measurements. Thus, the evolution operator describing the mth interaction is

$$\begin{aligned}V_m(t)=\sum_{\alpha_{m-1},\ldots,\alpha_1}&\{P(\mathbf{u}_{m;\alpha_{m-1},\cdots,\alpha_1})\otimes P_1(\beta_1)\otimes\cdots\\ &\otimes P_{m-1}(\beta_{m-1})\otimes I_m\otimes\cdots\otimes I_n\\ &+P(-\hat{\mathbf{u}}_{m;\alpha_{m-1},\cdots,\alpha_1})\otimes P_1(\beta_1)\\ &\otimes\cdots\otimes P_{m-1}(\beta_{m-1})\otimes e^{-i\theta_m(t)F_m}\otimes I_{m+1}\\ &\otimes\cdots\otimes I_n\}.\end{aligned}$$

Again we take $\beta_i=\uparrow$ if $\alpha_i=+$ and $\beta_i=\downarrow$ if $\alpha_i=-$. The full evolution operator is

$$U(t)=V_n(t)\cdots V_1(t). \tag{8.9}$$

During the interval $(m-1,m)$ we consider the Schmidt decompositions on each of the 2^{m-1} branches defined by the states

$$U(t)P_H^{\alpha_{m-1};\alpha_{m-2},\ldots,\alpha_1}(m-1)\cdots P_H^{\alpha_1}(1)|\psi(0)\rangle$$
$$=V_m(t)[P(\alpha_{m-1}\hat{\mathbf{u}}_{m-1;\alpha_{m-2},\ldots,\alpha_1})\cdots P(\alpha_1\hat{\mathbf{u}}_1)|\hat{\mathbf{v}}\rangle]$$
$$\otimes|\beta_1\cdots\beta_{m-1}\uparrow_m\cdots\uparrow_n\rangle,$$

with $\alpha_1,\ldots,\alpha_{m-1}$ independently running over the values $\pm$. Here

$$P_H^{\alpha_m;\alpha_{m-1},\ldots,\alpha_1}(t)=U^\dagger(t)P(\alpha_m\hat{\mathbf{u}}_{m;\alpha_{m-1},\ldots,\alpha_1})\otimes IU(t), \tag{8.10}$$

that is, the Heisenberg picture projection operator onto the branch-dependent axis of measurement. The branches, in other words, are defined by the branch-dependent Schmidt projections at times from 1 to $m-1$.

It is not hard, thus, to find a branch-dependent consistent set, built from the branch-dependent Schmidt projections at times 1 through to n, which describes the delayed-choice spin model sensibly.[7] However, since the retrodictive algorithms considered so far rely on the existence of a branch-independent set defined by the Schmidt decompositions of the original state vector, they will not generally reproduce this set (or any other interesting set). Branch-dependent physical descriptions, which are clearly necessary in quantum cosmology as well as in describing delayed-choice experiments, appear to rule out the type of retrodictive algorithm we have considered so far.

IX. BRANCH-DEPENDENT ALGORITHMS

The algorithms we have considered so far do not allow for branch dependence, and, hence, cannot possibly select the right set in many physically interesting examples. We have also seen that it is hard to find good Schmidt projection selection algorithms in which the projections selected at any time depend only on the physics up to that time, and that the possibility of recoherence rules out the existence of generally applicable algorithms of this type.

This suggests that *retrodictive* branch-dependent algorithms should be considered. Such algorithms, however, seem generally to require more information than is contained in the evolution of the quantum state. In the delayed-choice spin model, for example, it is hard to see how the Schmidt projections on the various branches, describing the delayed-choice measurements at late times, could be selected by an algorithm if only the entire state $\psi(t)$ — summed over all the branches — is specified.

The best, we suspect, that can be hoped for in the case of the delayed-choice spin model is an algorithm which takes all the final branches, encoded in the 2^n states $|\pm\mathbf{v}\rangle\otimes|\beta_1\ldots\beta_n\rangle$, where each of the β_i is one of the labels $\uparrow$ or $\downarrow$, and attempts to reconstruct the rest of the branching structure from the dynamics.

One possibility, for example, is to work backwards from $t=n$, and at each time t search through all subsets Q of branches defined at that time, checking whether the sum $|\psi^Q(t)\rangle$ of the corresponding states at time t has a Schmidt decomposition with the property that the Schmidt projections, applied to $|\psi^Q(t)\rangle$, produce (up to normalization) the individual branch states. If so, the Schmidt projections are taken to belong to the selected branch-dependent consistent set, the corresponding branches are unified into a single branch at times t and earlier, and the state corresponding to that branch at time t' is taken to be $U(t')U(t)^\dagger|\psi^Q(t)\rangle$, where U is the evolution operator for the model. Clearly, though, by specifying the final branch states we have already provided significant information — arguably most of the significant information — about the physics of the model. Finding algorithmic ways of supplying the branching structure of a natural consistent set, given all of its final history states, may seem a relatively minor accomplishment. It would obviously be rather more useful, though, if the final history states themselves were specified by a simple rule. For example, if the system and environment Hilbert spaces are both of large dimension, the final Schmidt states would be natural candidates. It would be interesting to explore these possibilities in quantum cosmology.

X. CONCLUSIONS

Bell, writing in 1975, said of the continuing dispute about quantum measurement theory that it ''is not between people who disagree on the results of simple mathematical manipulations. Nor is it between people with different ideas about the actual practicality of measuring arbitrarily complicated observables. It is between people who view with different degrees of concern or complacency the following fact: so long as the wave-packet reduction is an essential component, and so long as we do not know exactly when and how it takes over from the Schrödinger equation, we do not have an exact and unambiguous formulation of our most fundamental physical theory'' [47].

New formulations of quantum theory have since been developed, and the Copenhagen interpretation itself no longer dominates the debate quite as it once did. The language of wave-packet reduction, in particular, no longer commands anything approaching universal acceptance — thanks in large part to Bell's critiques. But the fundamental dispute is still, of course, very much alive, and Bell's description of the dispute still essentially holds true. Many approaches to quantum theory rely, at the moment, on well-developed intuition to explain, case by case, what to calculate in order to obtain a useful description of the evolution of any given physical system. The dispute is not over whether those calculations are correct, or even as to whether the intuitions used are helpful: generally, both are. The key question is whether we should be content with these successes, or whether we should continue to seek to underpin them by an exact and unambiguous formulation of quantum theory.

Consensus on this point seems no closer than it was in 1975. Many physicists take the view that we should not ever expect to find a complete and mathematically precise theory

[7]This sort of branch-dependent Schmidt decomposition could, of course, be considered in the original spin model, where all the axes of measurement are predetermined, but would not affect the earlier analysis, since the Schmidt projections in all branches are identical.

of nature, that nature is simply more complex than any mathematical representation. If so, some would argue, present interpretations of quantum theory may well represent the limit of precision attainable: it may be impossible, in principle, to improve on imprecise verbal prescriptions and intuition. On the other hand, this doubt could be raised in connection with any attempt to tackle any unsolved problem in physics. Why, for example, should we seek a unified field theory, or a theory of turbulence, if we decide *a priori* not to look for a mathematically precise interpretation of quantum theory? Clearly, too, accepting the impossibility of finding a complete theory of nature need not imply accepting that any definite boundary to precision will ever be encountered. One could imagine, for example, that every technical and conceptual problem encountered can eventually be resolved, but that the supply of problems will turn out to be infinite. And many physicists, of course, hope or believe that a complete and compelling theory of nature will ultimately be found, and so would simply reject the initial premise.

Complete agreement on the desiderata for formulations of quantum theory thus seems unlikely. But it ought to be possible to agree whether any given approach to quantum theory actually does supply an exact formulation and, if not, what the obstacles might be. Our aim in this paper has been to help bring about such agreement, by characterizing what might constitute a precise formulation of some of the ideas in the decoherence and consistent histories literature, and by explaining how hard it turns out to be to supply such a formulation.

Specifically, we have investigated various algorithms that select one particular consistent set of histories from among those defined by the Schmidt decompositions of the state, relative to a fixed system-environment split. We give examples of partial successes. There are several relatively simple algorithms which give physically sensible answers in particular models, and which we believe might usefully be applied elsewhere. We have not, though, found any algorithm which is guaranteed to select a sensible consistent set when both recoherence and branch-dependent system-environment interactions are present.

Our choice of physical models is certainly open to criticism. The spin model, for example, is a crudely simplistic model of real-world decoherence processes, which supposes both that perfect correlations are established between system and environment particles in finite time and that these interactions do not overlap. We would not claim, either, that the delayed-choice spin model necessarily captures any of the essential features of the branching structure of quasiclassical domains, though we would be very interested to know whether it might. We suspect that these simplifications should make it easier rather than harder to find set selection algorithms in the models, but we cannot exclude the possibility that more complicated and realistic models might prove more amenable to set selection.

The type of mathematical formulation we have sought is, similarly, open to criticism. We have investigated what seem particularly interesting classes of Schmidt projection set selection algorithms, but there are certainly others which may be worth exploring. There are also, of course, other mathematical structures relevant to decoherence apart from the Schmidt decomposition, and other ways of representing historical series of quantum events than through consistent sets of histories.

Our conclusion, though, is that it is extraordinarily hard to find a precise formulation of nonrelativistic quantum theory, based on the notions of quasiclassicality or decoherence, that is able to provide a probabilistic description of series of events at different points in time sufficiently rich to allow our experience of real world physics to be reconstructed. The problems of recoherence and of branch-dependent system-environment interaction, in particular, seem sufficiently serious that we doubt that the ideas presented in the literature to date are adequate to provide such a formulation. However, we cannot claim to have exhaustively investigated every possibility, and we would like to encourage sceptical readers to improve on our attempts.

ACKNOWLEDGMENTS

We are grateful to Fay Dowker and Trevor Samols for several helpful discussions and to Philip Pearle for raising the question of the implications of recoherence for set selection algorithms at an early stage of this work and for other valuable comments. A.K. was supported by the Royal Society, J.M. by the United Kingdom Engineering and Physical Sciences Research Council.

[1] R. B. Griffiths, J. Stat. Phys. **36**, 219 (1984).
[2] R. Omnès, J. Stat. Phys. **53**, 893 (1988).
[3] M. Gell-Mann and J. B. Hartle, in *Complexity Entropy and the Physics of Information*, Vol. III of *SFI Studies in the Science of Complexity*, edited by W. H. Zurek (Addison-Wesley, Reading, 1990).
[4] S. Goldstein and D. N. Page, Phys. Rev. Lett. **74**, 3715 (1995).
[5] F. Dowker and A. Kent, J. Stat. Phys. **82**, 1575 (1996).
[6] A. Kent, Phys. Rev. A **54**, 4670 (1996).
[7] F. Dowker and A. Kent, Phys. Rev. Lett. **75**, 3038 (1995).
[8] A. Kent, in *Bohmian Mechanics and Quantum Theory: An Appraisal*, edited by J. Cushing, A. Fine, and S. Goldstein (Kluwer Academic, Dordrecht, 1996).
[9] A. Kent, Cambridge University Report No. DAMTP/96-18, 1996 (unpublished).
[10] A. Kent, Cambridge University Report No. DAMTP/96-74, 1996 (unpublished).
[11] R. Omnès, *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton, 1994).
[12] R. B. Griffiths, Phys. Rev. A **54**, 2759 (1996).
[13] M. Gell-Mann and J. B. Hartle, University of California, Santa Barbara Report No. UCSBTH-95-28, 1995 (unpublished).
[14] E. Schmidt, Math. Ann. **83**, 433 (1907).
[15] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).
[16] D. Deutsch, Int. J. Theor. Phys. **24**, 1 (1985).

[17] D. Dieks, Phys. Lett. A **142**, 439 (1989).

[18] B. van Fraassen, in *Current Issues in Quantum Logic*, edited by E. Beltrametti and B. van Fraassen (Plenum, New York, 1981).

[19] R. Healey, *The Philosophy of Quantum Mechanics* (Cambridge University Press, New York, 1989).

[20] S. Kochen, in *Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985).

[21] A. Kent, Phys. Lett. A **196**, 313 (1995).

[22] A. Albrecht, Phys. Rev. D **46**, 5504 (1992).

[23] A. Albrecht, Phys. Rev. D **48**, 3768 (1993).

[24] M. Gell-Mann and J. B. Hartle, Phys. Rev. D **47**, 3345 (1993).

[25] A. O. Caldeira and A. J. Leggett, Physica A **121**, 587 (1983).

[26] E. Joos and H. D. Zeh, Zeit. Phys. B **59**, 223 (1985).

[27] H. F. Dowker and J. J. Halliwell, Phys. Rev. D **46**, 1580 (1992).

[28] H.-J. Pohle, Physica A **213**, 345 (1995).

[29] J. P. Paz, in *The Physical Origins of Time Asymmetry*, NATO, edited by J. J. Halliwell, J. P. Pérez-Mercader, and W. H. Zurek (Cambridge University Press, Cambridge, 1994).

[30] W. H. Zurek, in *The Physical Origins of Time Asymmetry*, NATO, edited by J. J. Halliwell, J. P. Pérez-Mercader, and W. H. Zurek (Cambridge University Press, Cambridge, 1994).

[31] C. Anastopoulos and J. J. Halliwell, Phys. Rev. D **51**, 6870 (1995).

[32] M. Tegmark and H. S. Shapiro, Phys. Rev. E **50**, 2538 (1994).

[33] J. P. Paz and W. H. Zurek, Phys. Rev. D **48**, 2728 (1993).

[34] J. B. Hartle, Phys. Rev. D **44**, 3173 (1991).

[35] W. H. Zurek, Phys. Today **44** (10), 36 (1991).

[36] J. N. McElwaine, Phys. Rev. A **53**, 2021 (1996).

[37] J. N. McElwaine, Ph.D. thesis, DAMTP, Cambridge University (1996).

[38] E. Muskat, D. Dubbers, and O. Schärpf, Phys. Rev. Lett. **58**, 2047 (1987).

[39] H. Weinfurter and G. Badurek, Phys. Rev. Lett. **64**, 1318 (1990).

[40] B. Alefeld, G. Badurek, and H. Rauch, Zeit. Phys. B **41**, 231 (1981).

[41] H. Weinfurter, G. Badurek, H. Rauch, and D. Schwahn, Zeit. Phys. B **72**, 195 (1988).

[42] L. Niel and H. Rauch, Zeit. Phys. B **74**, 133 (1989).

[43] G. Badurek, H. Rauch, and J. Summhammer, Phys. Rev. Lett. **51**, 1015 (1983).

[44] G. Badurek, H. Rauch, and D. Tuppinger, Phys. Rev. A **34**, 2600 (1986).

[45] B. Allman *et al.*, Phys. Rev. Lett. **68**, 2409 (1992).

[46] J. Summhammer, Phys. Rev. A **47**, 556 (1993).

[47] J. S. Bell, Helv. Phys. Acta **48**, 93 (1975).