# Toy model for local and deterministic wave-function collapse

Sandro Donadi and Sabine Hossenfelder
*Frankfurt Institute for Advanced Studies and Ruth-Moufang-Strasse 1, D-60438 Frankfurt am Main, Germany*

(Received 24 December 2021; revised 17 June 2022; accepted 8 August 2022; published 19 August 2022)

A local, deterministic toy model for quantum mechanics is introduced and discussed. It is demonstrated that the model makes the same predictions as quantum mechanics when averaged over the hidden variables. In this model the dynamics depends on the settings of the measurement device at the time of detection but not on how those settings were chosen. As a consequence, the model has a future input dependence and violates statistical independence but is local. We further show that the model is neither fine tuned nor allows for superluminal signaling and discuss the causality relations.

DOI: 10.1103/PhysRevA.106.022212

## I. INTRODUCTION

A century ago one might have hoped that quantum mechanics would one day be replaced by a theory compatible with the deterministic and local prequantum ontology. No such theory is known today: Neither Bohmian mechanics [1], nor models with spontaneous collapse [2,3], nor stochastic quantum mechanics [4] are local, and the latter two are not deterministic either. Bell's theorem [5] proves that such a local and deterministic replacement of quantum mechanics—if it is to reproduce observations—must violate statistical independence. This option has so-far not been very much explored.

Our aim here is to put forward a toy model that is local and deterministic, but violates statistical independence, and furthers the scientific discussion by elucidating its properties, especially its causal relations. Eventually, the goal of developing such a model is to remove the instantaneous measurement update and hence make it easier to combine quantum mechanics with general relativity. By no means do we claim that we will achieve this goal here. For example, the model we will discuss below still starts with a division between system and measurement devices and hence does not fully solve the measurement problem. Rather, the purpose of our toy model is to demonstrate that the most often raised objection to violations of statistical independence—that they require an unscientific "conspiracy"—is unfounded. This, our modest hope, will encourage further work on the approach.

Several toy models which violate statistical independence have previously been proposed [6–11] (for a review, see [12]). The one introduced here has the advantage that it can be applied to any kind of quantum-mechanical system with a finite-dimensional Hilbert space. This model should not be taken too seriously as a viable description of nature. It is neither pretty nor does it make sense as a fundamental theory for reasons that will be discussed later. However, it will allow us to study the consequences of requiring both locality and determinism.

Our model achieves its goal—of being local yet reproducing the predictions of quantum mechanics without fine tuning—by relying on a boundary condition to the future of preparation, namely, at the time of measurement. Using boundary conditions at several times has a long tradition, see, e.g., [13–22]. We will discuss this property in more detail in Sec. IV.

We will use units in which $c = \hbar = 1$.

## II. THE TOY MODEL

Suppose we have a quantum-mechanical system described by a state $|\Psi\rangle$ and a Hamiltonian $H$ in an $N$-dimensional Hilbert space. We want to measure an observable with $N$ eigenstates $|I\rangle$, $I \in \{1, .., N\}$, each one corresponding to a different state of the detector at the time of detection.[1] The system could be a composite system, and the detector could be distributed over different locations; the following will not depend on these details. If there are multiple detectors, we choose the space-time slicing so that the measurement takes place at the same coordinate time for all detectors.[2]

We then postulate that all the states in the Hilbert space except the eigenstates of the measurement observable are unstable under any kind of disturbance that comes from the hidden variables. That is, we will introduce a deviation from quantum mechanics quantified by a parameter $\kappa > 0$, which collapses arbitrary initial states into detector eigenstates. The measurement outcome will then be determined by the initial state of the system and the value of the hidden variables. When we average over these hidden variables, only the dependency on the initial state remains and the probabilistic outcome agrees with that of quantum mechanics. The reader can think of these variables as encoding the detailed degrees of freedom

[1]Strictly speaking, of course, detection does not happen in a single moment but during a finite time interval. This distinction, however, does not matter for the purposes of this toy model, as will become clear later.

[2]If there is no spacelike slice that can accommodate all the detectors, then they are causally related, in which case the latter detectors can be ignored because the state is already "collapsed."

 

of the detector because that is the most minimal possibility, but it could be more complicated.

In the next section we will start with the case $N = 2$ to introduce the general idea and then generalize to larger $N$.

### A. $N = 2$

For $N = 2$ we have

$$|\Psi(t)\rangle := a_1(t)|1\rangle + a_2(t)|2\rangle\ , \tag{1}$$

with complex factors $a_1(t)$ and $a_2(t)$ that fulfill $|a_1(t)|^2 + |a_2(t)|^2 = 1$. We will denote the time of system preparation with $t_\mathrm{p}$ and that of measurement with $t_\mathrm{d}$. Let us further denote with $|\Psi^*(t)\rangle$ the solution to the normal Hamiltonian evolution

$$|\Psi^*(t)\rangle := \exp[-\mathrm{i}H(t - t_\mathrm{p})]|\Psi(t_\mathrm{p})\rangle. \tag{2}$$

It will also be handy to define the coefficients of the eigenvectors at the time of measurement under the usual Hamiltonian evolution:

$$\alpha_I := \langle I|\Psi^*(t_\mathrm{d})\rangle. \tag{3}$$

We know that in quantum mechanics the probability of measuring the $I$th outcome is $|\alpha_I|^2$, so the superdeterministic theory must reproduce this on the average. These numbers can be calculated from $H$ alone, so the underlying dynamics that we are about to construct is unnecessary to make the probabilistic predictions, but this is not the point. The point of this model is to get rid of the measurement update.

As hidden variables we use complex numbers that are uniformly distributed inside the complex unit circle. The radius of the distribution could be chosen differently for each of these variables, but this would just add unnecessary parameters. We will denote the random variable as $\lambda_2$, where the index 2 on $\lambda_2$ refers to $N = 2$. These hidden variables are part of what fully specifies the initial state.

With that, we construct, as usual, the density matrix $\rho(t) := |\Psi(t)\rangle\langle\Psi(t)|$ but change the dynamical law to

$$\begin{aligned}\partial_t\rho(t) = &-\mathrm{i}[H, \rho(t)]\\ &+\kappa\left(L_{21}\rho(t)L_{21}^\dagger - \tfrac{1}{2}\{\rho_t, L_{21}^\dagger L_{21}\}\right).\end{aligned} \tag{4}$$

Here $\kappa$ is a constant of dimension energy (for the interpretation see Sec. III D), the curly brackets denote the anticommutator, and

$$L_{21} := \theta(\sigma_2)|\chi_2\rangle\langle\chi_1| + \theta(-\sigma_2)|\chi_1\rangle\langle\chi_2|, \tag{5}$$

$$\sigma_2 := |\lambda_2\beta_2|^2 - (1 - |\lambda_2|^2)|\beta_1|^2, \tag{6}$$

where $|\chi_I\rangle$ is an arbitrary orthonormal basis, and

$$\beta_I := \langle\chi_I|\Psi^*(t_d)\rangle. \tag{7}$$

Again, the indices on $L_{21}$ and $\sigma_2$ refer to the case $N = 2$ and are merely there to make it easier to later generalize to higher $N$.

$\theta(\cdot)$, as usual, denotes the Heaviside function, which we define as $\theta(x) = 0$ for $x \leqslant 0$ and $\theta(x) = 1$ otherwise. It must be emphasized that the $\rho(t)$ in Eq. (4) is *not* the density matrix of quantum mechanics. To get the density matrix of quantum mechanics, we have to average over the random variable $\lambda_2$. We will demonstrate this explicitly later.

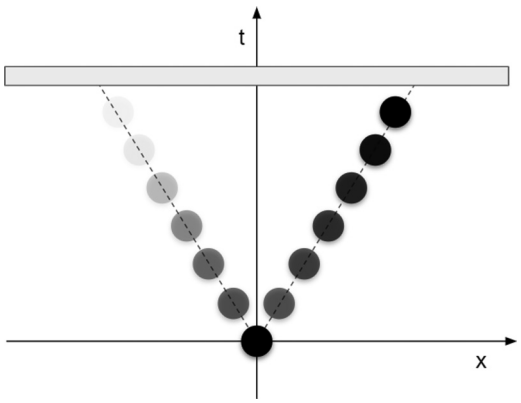


FIG. 1. Schematic illustration of the collapse described by the toy model. Shades of gray indicate absolute values of amplitudes, where darker (lighter) shades are larger (smaller) amplitudes. Just after the superposition is created (e.g., by a beam splitter), the two branches have the same amplitude, but as they approach the detector (gray bar on top), one branch becomes dominant over the other one.

The final ingredient to the model is now the future input: We postulate that the evolution is optimal if and only if $|\chi_I\rangle = |I\rangle$, so that $\alpha_I = \beta_I$. It might seem somewhat perplexing that we introduced the $|\chi_I\rangle$ only to then remove them. We will comment on this later. There are a variety of functions that, when optimized, would spit out this requirement, and it is rather unnecessary here to write down a specific one.[3]

The dynamical law (4) has the Lindblad form [23,24]. It is the master equation for one of the most common examples of decoherence, that of amplitude damping in a two-level system [25]. We will, in the following, work in the limit where $\kappa$ is much larger than the typical energies of the system, so that the collapse dynamics is the dominant effect. In this regime it is then easy to explicitly solve (4) and the solution has the following asymptotic behavior (for details see Sec. III E):

$$\lim_{\kappa t\to\infty}\rho(t) = \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} \quad \text{for } \sigma_2 > 0, \tag{8}$$

$$\lim_{\kappa t\to\infty}\rho(t) = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} \quad \text{for } \sigma_2 < 0. \tag{9}$$

In Fig. 1 we illustrate the collapse for the simplest case, that of a two-state superposition with equal amplitudes. Contrary to what happens in standard quantum mechanics, the collapse in this model is gradual, and it starts before the particle reaches the proximity of the detector. This toy model, importantly, is not a model of spontaneous collapse [3] or, more generally, a model of a stochastic process [26], because there is no stochasticity in the dynamics. The dynamics is determined by the time-independent random variable and not a random walk. If one selects a specific $\lambda_2$, then one knows for certain what $\sigma_2$ is and one knows what the outcome of the evolution is.

[3]For example, $\sum_I(\langle\chi_I|I\rangle + \langle I|\chi_I\rangle)$, when maximized, will yield the constraint $|\chi_I\rangle = |I\rangle$.

We now have to show that the probabilities for the state to evolve to $|1\rangle$ and $|2\rangle$ come out correctly. As we just saw, in the limit where $\kappa \gg E_{\max}$ an initial state $|\Psi(t)\rangle$ will go to $|2\rangle$ if $\sigma_2 > 0$. We therefore have

$$P(|2\rangle) = P(\sigma_2 > 0), \tag{10}$$

where the $P$ here and in the following denotes a probability. In terms of the random variable $\lambda_2$, we can write the condition $\sigma_2 > 0$ as

$$\begin{aligned} & |\lambda_2\alpha_2|^2 > (1-|\lambda_2|^2)|\alpha_1|^2 \\ \Leftrightarrow \quad & |\lambda_2\alpha_2|^2 > (1-|\lambda_2|^2)(1-|\alpha_2|^2) \\ \Leftrightarrow \quad & |\lambda_2|^2 > 1-|\alpha_2|^2. \end{aligned} \tag{11}$$

Since $\lambda_2$ is uniformly distributed in the unit circle, this means we are asking for the area of the ring between the radii $\sqrt{1-|\alpha_2|^2}$ and 1 relative to the area of the whole disk. The area of the ring and the disk scale with the square of the radius, so the fraction is $1-1+|\alpha_2|^2 = |\alpha_2|^2$, and we have

$$P(|2\rangle) = |\alpha_2|^2. \tag{12}$$

It follows that the probability of the state to evolve to $|1\rangle$ is $|\alpha_1|^2$. This is exactly Born's rule.

### B. $N > 2$

To generalize from $N = 2$ to larger $N$, it helps to have a physical description of the process that the dynamical law (4) is commonly used to describe. For the case $\sigma_2 > 0$, the state $|\chi_1\rangle$ is unstable and decays to $|\chi_2\rangle$, while for $\sigma_2 < 0$ it is the other way around, $|\chi_2\rangle$ decays to $|\chi_1\rangle$. The constant $\kappa$ determines the decay time. Provided that the contribution from the Hamiltonian can be neglected, this leads asymptotically to the limits in (9).

For $N > 2$ one then iteratively adds more random vectors $\lambda_N$, each of which is independently uniformly distributed on the unit disk and defines

$$\sigma_N := |\beta_N\lambda_N|^2 - \sum_{I=1}^{N-1} |\beta_I|^2(1-|\lambda_N|^2). \tag{13}$$

With that, we construct the Lindblad operators recursively for higher $N$ by drawing on the comparison with decaying states. In the step from $N-1$ to $N$, we introduce $N-1$ new operators:

$$L_{NM} := \theta(\sigma_N)|\chi_N\rangle\langle\chi_M| + \theta(-\sigma_N)|\chi_M\rangle\langle\chi_N| \tag{14}$$

for $M \in \{1..N-1\}$. This means that we have in total $N(N-1)/2$ Lindblad operators $L_{KM}$ for $N$ dimensions, and by convention we have labeled them so that the second index is always strictly smaller than the first. With that, the master equation can be written as

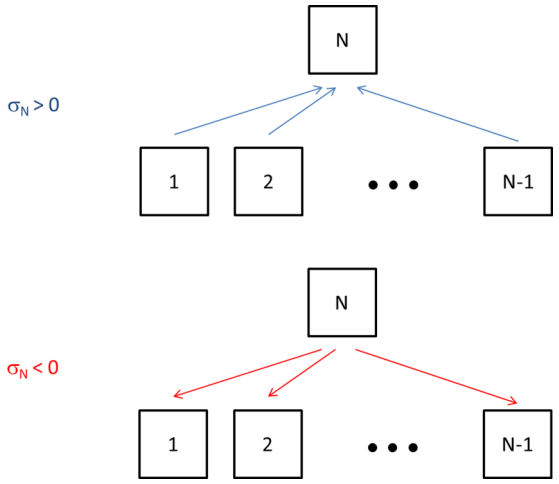


FIG. 2. The dynamics induced by the Lindblad terms involving the operators $L_{NM}$ with $M = 1, 2, ...N-1$. For $\sigma_N > 0$ each state will evolve to the state $|\chi_N\rangle$. For $\sigma_N < 0$, the state $|\chi_N\rangle$ will evolve evenly into the states $|\chi_1\rangle, ...|\chi_{N-1}\rangle$.

$$\partial_t\rho(t) = -\mathrm{i}[H,\rho(t)] + \kappa \sum_{K>M=1}^{N} L_{KM}\rho(t)L_{KM}^\dagger - \kappa \sum_{K>M=1}^{N} \frac{1}{2}\{\rho(t), L_{KM}^\dagger L_{KM}\}. \tag{15}$$

As above for the case $N = 2$, we obtain the density matrix of quantum mechanics by averaging the $\rho$ from Eq. (15) over the hidden variables.

Before we move on, let us briefly interpret the structure of the Lindblad operators that we have constructed. The $N-1$ new operators that we add in each step $N-1$ to $N$ have the property that they either take any initial state $|\chi_K\rangle$ with $K < N$ and asymptotically convert it to $|\chi_N\rangle$ (if $\sigma_N > 0$) or they take the initial state $|\chi_N\rangle$ and distribute it evenly on the states $|\chi_K\rangle$ with $K < N$ (if $\sigma_N < 0$). This is illustrated in in Fig. 2.

So if $\sigma_N > 0$, we know that the result will be $|\chi_N\rangle$. Let us then look at the case $\sigma_N \leqslant 0$. Once the state $|\chi_N\rangle$ is cleared out and distributed over the subspace spanned by $|\chi_1\rangle...|\chi_{N-1}\rangle$, the same thing happens again for the remaining states, but this time depending on the sign of $\sigma_{N-1}$: If $\sigma_{N-1} > 0$, then all remaining states go to $|\chi_{N-1}\rangle$, if $\sigma_{N-1} < 0$, then they will be distributed over the subspace spanned by $|\chi_1\rangle...|\chi_{N-2}\rangle$, and so on.

Having said that, we can now show that the generalization to $N > 2$ still fulfills Born's rule. First, we postulate as previously that the optimal evolution is the one with $|\chi_I\rangle = |I\rangle$, so let us then look at the eigenstate $|N\rangle$. As just explained, if $\sigma_N > 0$, then the operators $L_{NK}$ will convert any initial state $|K\rangle$ with $K < N$ into the state $|N\rangle$. One can calculate the probability for this the same way as above in (11) and easily find

$$P(|N\rangle) = P(\sigma_N > 0) = |\alpha_N|^2. \tag{16}$$

Correspondingly, the probability that the state goes to any one of the other $N-1$ eigenstates is $1-|\alpha_N|^2$.

To obtain the probability that an initial state asymptotically goes to $|N-1\rangle$, one then needs to calculate the probability for

$$P(|N-1\rangle) = P(\sigma_N < 0 \wedge \sigma_{N-1} > 0) \tag{17}$$

$$= P(\sigma_N < 0)P(\sigma_{N-1} > 0), \tag{18}$$

because the random variables are by assumption uncorrelated.

For this, one has to take into account that for $N-1$ basis states $\sum_{i=1}^{N-1} |\alpha_i|^2 \leqslant 1$, and therefore the probability for $\sigma_{N-1}$ to be larger or smaller than 1 has to be divided by this sum. One sees this by rewriting the requirement $\sigma_{N-1} > 0$ as

$$|\lambda_{N-1}\alpha_{N-1}|^2 > (1-|\lambda_{N-1}|^2)\sum_{I=1}^{N-2}|\alpha_I|^2, \tag{19}$$

which is equivalent to

$$|\lambda_{N-1}|^2 \frac{|\alpha_{N-1}|^2}{\sum_{I=1}^{N-1}|\alpha_I|^2} > (1-|\lambda_{N-1}|^2)\left(1-\frac{|\alpha_{N-1}|^2}{\sum_{I=1}^{N-1}|\alpha_I|^2}\right)$$

$$\Leftrightarrow \qquad |\lambda_{N-1}|^2 > 1-\frac{|\alpha_{N-1}|^2}{\sum_{I=1}^{N-1}|\alpha_I|^2}. \tag{20}$$

So, by the same argument as before, the probability for $\sigma_{N-1} > 0$ is now $|\alpha_{N-1}|^2/\sum_{i=1}^{N-1}|\alpha_i|^2$ (instead of just $|\alpha_{N-1}|^2$, as it was for $N-1$ states). Hence we have

$$P(|N-1\rangle) = (1-|\alpha_N|^2)\frac{|\alpha_{N-1}|^2}{\sum_{i=1}^{N-1}|\alpha_i|^2} = |\alpha_{N-1}|^2, \tag{21}$$

which is what Born's rule requires.

It works the same way for all remaining eigenstates. For example, the probability of an initial state to go to $|N-2\rangle$ is given by

$$P(|N-2\rangle) = P(\sigma_N < 0 \wedge \lambda_{N-1} < 0 \wedge \sigma_{N-2} > 0)$$
$$= P(\sigma_N < 0)P(\sigma_{N-1} < 0)P(\sigma_{N-2} > 0), \tag{22}$$

which is

$$(1-|\alpha_N|^2)\left(1-\frac{|\alpha_{N-1}|^2}{\sum_{i=1}^{N-1}|\alpha_i|^2}\right)\frac{|\alpha_{N-2}|^2}{\sum_{i=1}^{N-2}|\alpha_i|^2} = |\alpha_{N-2}|^2. \tag{23}$$

In general, then, the probability for any initial state to asymptotically come out as $|K\rangle$ is

$$P(|K\rangle) = P(\sigma_N < 0 \wedge \ldots \wedge \sigma_{K+1} < 0 \wedge \sigma_K > 0)$$
$$= (1-|\alpha_N|^2)\prod_{J=N-1}^{K+1}\left(1-\frac{|\alpha_J|^2}{\sum_{I=1}^{J}|\alpha_I|^2}\right)$$
$$\times\left(\frac{|\alpha_K|^2}{\sum_{I=1}^{K}|\alpha_I|^2}\right) = |\alpha_K|^2. \tag{24}$$

This completes our proof that Born's rule is fulfilled for all $N$.

The way we have defined the $N$-dimensional case depends on the order of the basis states. Since the order is arbitrary, this seems unphysical. It does not really matter for the purposes of this model, because the resulting probability distribution is independent of the ordering. However, one could make the model explicitly independent of the ordering by summing over all possible permutations.

One may note that this distribution fulfils the requirements for Born's rule laid out in [27].

## III. PROPERTIES OF THE TOY MODEL

### A. Superdeterminism

It might seem perplexing why we didn't just put the measurement basis into the evolution law right away.[4] The reason is that this way it is easier to see what are the "hidden variables" in Bell's sense. Since we already used $\lambda$, we will denote Bell's hidden variables with $\Lambda$. To fully specify the evolution of the prepared state, we need both the $\lambda_I$ and the basis $|\chi_I\rangle$. Hence Bell's hidden variables are the combination of both $\Lambda = (\lambda_I, |\chi_I\rangle)$. This makes clear that once we assume $|\chi_I\rangle = |I\rangle$, the hidden variables are $\Lambda = (\lambda_I, |I\rangle)$.

If one does not first introduce the $|\chi_I\rangle$ basis, it is easy to confuse oneself over what the hidden variables are, because Bell lists the detector settings separately. Introducing the $|\chi_I\rangle$ makes clear that the detector settings are part of the full specification of beables that determines the measurement outcome. Since the variables $\sigma_I$ are functions of the $\lambda_I$ and $|I\rangle$, one can alternatively write $\Lambda = (\sigma_I, |I\rangle)$.

This means that if we consider a Bell-type test with two detectors that have settings **a** and **b**, then the condition

$$p(\Lambda|\mathbf{a},\mathbf{b}) = p(\Lambda) \tag{25}$$

is not fulfilled. Such theories are known as superdeterministic [28]. They are the only known consistent, local, and deterministic completion of quantum mechanics. If one mistakes the $\lambda_I$ alone for Bell's hidden variables, then one erroneously arrives at the conclusion that the model respects statistical independence (because the $\lambda$s do not depend on the detector settings) but violates local causality (since one needs the detector settings to calculate the probabilities).

As is well known [29], one can construct models with nonlocal correlations that are even stronger than those of quantum mechanics. We therefore want to stress here that our model is explicitly constructed to reproduce standard quantum mechanics. But superdeterminism in general, or our model in particular, does not explain why the correlations of quantum mechanics are what they are and not any stronger.

Our model is deterministic in the sense that if one is given the values of the hidden variables and the initial state of the wave function, then the outcome of the measurement is entirely determined. It is unpredictable, however, to use the distinction proposed in [30], because the values of the hidden variables are unknowable. The only predictions one obtains from this model are the same statistical predictions as that of standard quantum mechanics (our toy model is in this respect similar to Bohmian mechanics).

### B. Locality

To discuss the notion of locality, we will, as usual, assume that the preparation of the state as well as the measurements take place in regions localized in space-time. Let us denote the measurement outcome at the detector with setting **a** as $O_{\mathbf{a}}$ and

[4] Indeed, we did this in the first version of this paper.

that at the detector with setting $\mathbf{b}$ as $O_\mathbf{b}$. Then Bell's notion of locality is the condition that [31]

$$p(O_\mathbf{b}|O_\mathbf{a}, \mathbf{a}, \mathbf{b}, \Lambda) = p(O_\mathbf{b}|\mathbf{b}, \Lambda), \tag{26}$$

where we have omitted a further possible dependence on the preparation procedure because our model obviously does not have this dependence. Strictly speaking, condition (26) is a combination of two conditions:

$$p(O_\mathbf{b}|O_\mathbf{a}, \mathbf{a}, \mathbf{b}, \Lambda) = p(O_\mathbf{b}|O_\mathbf{a}, \mathbf{b}, \Lambda), \tag{27}$$

$$p(O_\mathbf{b}|O_\mathbf{a}, \mathbf{a}, \mathbf{b}, \Lambda) = p(O_\mathbf{b}|\mathbf{a}, \mathbf{b}, \Lambda). \tag{28}$$

While condition (27) can readily be identified as a locality requirement (the probability for the outcome at one detector should not depend on the setting at the other detector), condition (28) is instead a requirement of completeness of the hidden variables $\Lambda$: When (28) is violated, it means that the hidden variables $\Lambda$ are not enough to determine the outcome at one detector (given the setting), but further information of the outcome on the other detector is required. This condition can also be violated in classical systems and has little to do with locality.

The condition (26) is fulfilled because in our toy model $\Lambda$ contains the information about $\mathbf{a}$ and $\mathbf{b}$ and it predicts both measurement outcomes. For this reason the settings are redundant for the calculation of the probability. From (26) follows Bell's condition of factorization:

$$p(O_\mathbf{a}, O_\mathbf{b}|\mathbf{a}, \mathbf{b}, \Lambda) = p(O_\mathbf{a}|\mathbf{a}, \Lambda)p(O_\mathbf{b}|\mathbf{b}, \Lambda), \tag{29}$$

which is the starting point for the derivation of Bell's inequality.

### C. Causality

For his notion of "local causality," Bell adds the following criterion [32]:

> "Correlations between events in different spacetime regions should be explicable in terms of physical events in the overlap of the backward lightcones."

In our toy model the correlations between measurements in different space-time regions ($\mathbf{D_1}$ and $\mathbf{D_2}$) are explained by the hidden variables ($\Lambda$) of the prepared state ($\mathbf{P}$). This prepared state is necessarily in the overlap of the backward light cones because otherwise the prepared state could not reach the detector (see Fig. 3). Formally, therefore, our model fulfils Bell's notion of local causality.

There are several reactions one can have to this. First, one may argue that, leaving aside what Bell wrote, it may not have been what he meant. One may ask, for example, just exactly what "explicable" means anyway. In our model the correlations between events can be calculated from the information in the overlap of the backward light cone. They are hence explicable in the context of the model. The model does not explain why this initial state is the way it is. But any model necessarily requires some assumptions that are unexplained within the model, and none of our current theories "explain" their own initial state. Hence, requiring that our model explains its initial state would be holding us to a higher standard than everybody else.

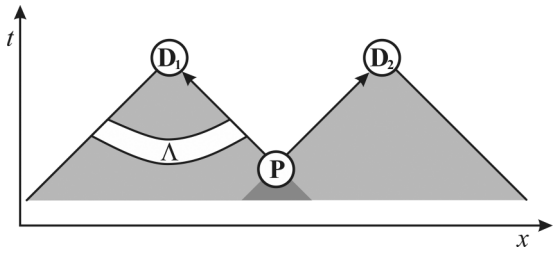


FIG. 3. Sketch to illustrate local causality. A state is prepared in region $\mathbf{P}$ and measured at two detectors $\mathbf{D}_1$ and $\mathbf{D}_2$. The hidden variables $\Lambda$ are present already at $\mathbf{P}$ and therefore in the backward light cones of both detectors.

Second, one may ask what counts as a "physical event." Maybe the preparation of a state does not count as a physical event. We have nothing insightful to say about this other than that the term is not well defined and it is therefore hard to assess whether the criterion is fulfilled.

The third reaction one may have is questioning whether it matters what Bell said to begin with. After all, Bell's writing is not holy scripture, and our task is not to interpret what he may have meant but to explain observations. Maybe we should therefore apply a different notion of causality. The problem with this is that various different terms have been proposed in the literature.

It has been suggested, e.g., to distinguish retrocausal approaches that violate statistical independence from those that do not [33]. Unfortunately, the term "retrocausality" is itself not well defined. If one uses the space-time notion of causality (in which causes are always in the past), then retrocausality simply does not exist. This is why our model is locally causal according to Bell's definition, because Bell explicitly (by referring to backward light cones) uses the space-time notion of causality. One might instead use the interventionist notion of causality [34,35], but then one needs to explain why this notion should be relevant for the foundations of physics. The situation is not helped by the unclear relation between the two notions of causality that were confused in [36], as pointed out in [37]. For an attempt at clarifying this, see [38].

For this reason we think the most appropriate description is that our model has a "future input dependence," which we will elaborate on in more detail in Sec. IV. Note, however, that [39] equates this term with retrocausality, using the interventionist notion of causality, which threatens to bring back the confusion that we try to avoid. Maybe it would be better to instead use the term proposed by Adlam [40,41], who coins models with a global constraint like ours as "atemporal" or "all-at-once" approaches. While this seems like an apt description, it merely answers the question of causality with a shrug.

In the end these are all just words which change nothing about the properties of the model. The reader is welcome to form their own opinion about what notion of causality our model fulfils based on the above elaboration on the difficulty. We want to emphasize, though, that for the goal of making quantum mechanics compatible with general relativity, the question in just which way the model is causal is not relevant. It is only relevant that the model respects the causal structure

of space-time, and this requirement does not single out one direction of time as preferred.

The problem with Bell's notion of causality ultimately stems from his definition of the "hidden variables." The hidden variables by definition are that what determines the measurement outcome. This, of course, must include an evolution law for the hidden variables, because an initial state without an evolution law does not determine the final outcome [42]. But assigning a location to a parameter in an evolution law makes no sense. One therefore should not be too surprised about the here discussed problem of evaluating the causality of a hidden variables model in Bell's sense.

### D. Fine tuning, Absence of

Theories that violate statistical independence are usually dismissed as "conspiracy theories" that require "fine tuning." Those arguments, which originated in the 1970s [43,44], were previously addressed in [28,37], and our toy model demonstrates just exactly why these fine-tuning arguments are misleading.

What we mean by fine tuning here is that the model requires one to specify so many details to make a prediction that it becomes scientifically useless. It is already clear from the above that the toy model is not fine tuned. It gives us exactly the same predictions as quantum mechanics without the need to specify many details. But just why are the fine-tuning arguments wrong?

These arguments have it that there are many different ways to choose detector settings, and the model must work for all of them. This, it is then argued, requires that the model specifies what happens for a huge number of different cases and must still reproduce quantum mechanics on the average.

The problem with this argument is that there is no reason why, in a model that violates statistical independence, the measurement outcome should depend on how the detector settings were chosen any more than this is the case in quantum mechanics. The settings at the time of measurement alone are enough to get the correct average predictions. What happens in the brain of the experimenter (or with any other mechanism used to determine detector settings [45–47]) does not matter; the only thing that matters is the setting of the measurement device at the time of measurement. It follows that the average outcome is insensitive to the details and is hence not fine tuned. We already know this is possible, because it's what we do in quantum mechanics itself: we predict average measurement outcomes from the detector settings alone.

A few words are in order here about the notion of fine tuning that we are using. The expression "fine tuning" is used in many different areas of physics, but with slightly different meanings [48,49]. Most of these notions of fine tuning are metaphysical, that is, they are assumptions which physicists make that have no empirical support but are merely used to select one model out of many possible models. That an assumption is metaphysical does not make it wrong, but it means it is optional. Such assumptions are in many cases used out of tradition, not because of empirical necessity.

Metaphysical fine-tuning arguments typically analyze the behavior of the predictions of a model under a variation of input parameters which are not known to be variable. This is the case, for example, in cosmology when constants of nature are varied, or in particle physics, when perturbations in variables (which are physically possible) are conflated with perturbations in parameters that cannot change in any physically possible process. It is also the case in the arguments put forward in [36,50,51].

These analyses, while technically entirely correct, are based on assumptions about probability measures of input parameters which are empirically unknown (or even, qua assumption, unknowable). The output of a model might very well sensitively depend on some input parameter, yet if that input parameter cannot vary, or its variations are constrained, this dependence is irrelevant for how predictive a model is.

This is also the case for initial values, which are yet another type of input for a model. Unless there is empirical evidence that an initial value for a variable can indeed take on different values with a certain frequency, assumptions about the distribution of such initial values are metaphysical. This is especially true for initial values for the state of the universe for which we can never empirically determine the distribution over initial values. Our universe had one set of initial values and that's the one and only instance we can observe.

Metaphysical arguments are the subject of philosophy. This does not mean they are irrelevant per se, but since metaphysical fine-tuning arguments about superdeterminism were previously addressed in [28,37], they do not concern us here.

In our analysis, in contrast, we take a strictly instrumental approach. A model is not scientifically viable if it is impossible to make predictions with it or if making predictions with it requires more information as input than the model delivers as output. Such a model, that requires one to specify a lot of detail to make a prediction, is what we call fine tuned. This is *not* the definition of fine tuning used in [36,50,51]. This is because the definition used in these other works does not tell us what it actually takes to make predictions with a model. The notion of fine tuning we use here is the notion that underlies the criticism first raised in [43,44] and later repeated many times (for more references see [28,52]), according to which superdeterminism is supposedly unscientific.

It would be possible to quantify the predictability of our model, but this is both unnecessary and uninstructive because any such quantifier would be arbitrary anyway. We have instead compared our model to standard quantum mechanics and demonstrated that they come out in a tie.

### E. Superluminal Signaling, Absence of

In this section we want to explicitly show why the model does not lead to faster-than-light signaling. We could simply argue that since the toy model was constructed to reproduce the predictions of quantum mechanics, faster-than-light signaling is impossible with the toy model just because it is impossible with standard quantum mechanics. One may further notice that the master equation (15) is linear in the density matrix: the model hence does not suffer from the problem discussed in [53,54].[5] However, while correct, these

[5]It should be mentioned that Refs. [53,54] were concerned with a different question than the one investigated here, namely, whether

arguments are not very illuminating. Just exactly how does it happen that superluminal signaling is prevented even though the model is deterministic? The calculation in this section will demonstrate explicitly how the average over the hidden variables removes the possibility for superluminal signaling. It will also be instructive to see how the model works in practice.

We will concretely look at a singlet state, shared between two, potentially very distant, observers, Alice ($A$) and Bob ($B$), who measure the spin of a particle, which is either up ($\uparrow$) or down ($\downarrow$) in their basis. We will chose the weights of the entangled state to be $q$ and $\sqrt{1-q^2}$, respectively (instead of using $1/\sqrt{2}$ for both), to make sure cancellations between terms do not occur because of the symmetry of the state.

The total dimension of the Hilbert space is $N = 2^2 = 4$. We now assume the optimization condition is fulfilled $|\chi_I\rangle = |I\rangle$ and, to connect with the notation of the previous section, we define

$$|1\rangle := |\downarrow_A\rangle \otimes |\uparrow_B\rangle \equiv |\downarrow_A \uparrow_B\rangle, \tag{30}$$

$$\begin{aligned} |2\rangle &:= |\uparrow_A\rangle \otimes |\downarrow_B\rangle \equiv |\uparrow_A \downarrow_B\rangle, \\ |3\rangle &:= |\downarrow_A\rangle \otimes |\downarrow_B\rangle \equiv |\downarrow_A \downarrow_B\rangle, \end{aligned} \tag{31}$$

$$|4\rangle := |\uparrow_A\rangle \otimes |\uparrow_B\rangle = |\uparrow_A \uparrow_B\rangle. \tag{32}$$

The prepared state at time $t = t_\mathrm{p}$ is

$$\begin{aligned} |\Psi(t_\mathrm{p})\rangle &= q|\uparrow_A \downarrow_B\rangle - \sqrt{1-q^2}|\downarrow_A \uparrow_B\rangle \\ &= q|1\rangle - \sqrt{1-q^2}|2\rangle, \end{aligned} \tag{33}$$

and has the density matrix

$$\begin{aligned} \rho(t_\mathrm{p}) = q^2|1\rangle\langle 1| + (1-q^2)|2\rangle\langle 2| \\ - q\sqrt{1-q^2}(|1\rangle\langle 2| + |2\rangle\langle 1|). \end{aligned} \tag{34}$$

Our task is now to solve the master equation (15) for the density matrix with the initial value (34) and then take the average over the hidden variables to get

$$\overline{\rho}(t) := \int \rho(t) \prod_{I=2}^{4} d\lambda_I. \tag{35}$$

For a four-dimensional Hilbert space, we have six Lindblad operators, one for each entry above the diagonal. For the singlet state under consideration, $\alpha_1^2 = q^2$, $\alpha_2 = 1-q^2$, $\alpha_3^2 = \alpha_4^2 = 0$, so we have $\sigma_2 = |\lambda_2|^2 - q^2$, $\sigma_3 = |\lambda_3|^2 - 1$, and $\sigma_4 = |\lambda_4|^2 - 1$. Since the $\lambda'$s all lie within the unit circle, we see that $\sigma_3$ and $\sigma_4$ are both $\leqslant 0$. This reflects the fact that the probability to end up in state $|3\rangle$ or $|4\rangle$ should be zero.

One can then integrate the master equation (see Appendix), using the initial state (34), from which one obtains the evolution for the matrix elements $\rho_{ij}(t) := \langle i|\rho(t)|j\rangle$:

$$\begin{aligned} \rho_{11}(t) &= \theta(\sigma_2)q^2 e^{-\kappa t} + \theta(-\sigma_2)[1-(1-q^2)e^{-\kappa t}], \\ \rho_{22}(t) &= \theta(\sigma_2)(1-q^2e^{-\kappa t}) + \theta(-\sigma_2)(1-q^2)e^{-\kappa t}, \\ \rho_{12}(t) &= \rho_{22}(t) = q\sqrt{1-q^2}e^{-\kappa t/2}, \end{aligned} \tag{36}$$

and all the other matrix elements are equal to zero. Note that $\mathrm{Tr}(\rho) = 1$ for all times.

We already know that

$$P(\sigma_2 < 0) = q^2, P(\sigma_2 > 0) = 1-q^2. \tag{37}$$

So, for the average of the density matrix, we have

$$\begin{aligned} \overline{\rho}_{11}(t) &= (1-q^2)q^2e^{-\kappa t} + q^2(1-(1-q^2)e^{-\kappa t}) \\ &= q^2, \\ \overline{\rho}_{22}(t) &= (1-q^2)(1-q^2e^{-\kappa t}) + q^2(1-q^2)e^{-\kappa t} \\ &= 1-q^2, \\ \overline{\rho}_{12}(t) &= \overline{\rho}_{22}(t) = q\sqrt{1-q^2}e^{-\kappa t/2}. \end{aligned} \tag{38}$$

This means the only thing the Lindblad operators do is to exponentially suppress the off-diagonal terms in the density matrix. We can now take the partial traces

$$\begin{aligned} \overline{\rho}_A &= \mathrm{Tr}_B(\overline{\rho}) \\ &= q^2|\uparrow_A\rangle\langle\uparrow_A| + (1-q^2)|\downarrow_A\rangle\langle\downarrow_A|, \end{aligned} \tag{39}$$

$$\begin{aligned} \overline{\rho}_B &= \mathrm{Tr}_A(\overline{\rho}) \\ &= q^2|\downarrow_B\rangle\langle\downarrow_B| + (1-q^2)|\uparrow_B\rangle\langle\uparrow_B|. \end{aligned} \tag{40}$$

This is exactly the same result as in quantum mechanics. We thus see that if one treats the random variables in the above-discussed toy model as truly random, then superluminal signaling is not possible.

The full calculation of the two-qubit system, which we present in the Appendix, can be applied to generic Bell-type tests and shows that one obtains the same result as quantum mechanics for any initial state and any combination of detector settings. Our model therefore reproduces the same Bell-inequality violations as quantum mechanics.

However, one may object that if the hidden variables are really random, then the model is not superdeterministic because it is not deterministic in the first place. Indeed, in a superdeterministic model, the randomness should not be fundamental but merely a consequence of lacking detailed information—as one expects in a "hidden variables" theory. Thus the distribution of the hidden variables should in certain circumstances deviate from being completely uniform. In this case, taking the average over the hidden variables will not give the same result as in quantum mechanics. This is apparent from Eqs. (36) which, when averaged with probabilities other than (37) will not give (38), but have correction terms in whatever is the deviation of the distribution of the $\lambda'$s from uniformity.

This, per se, does not constitute an opportunity for signaling, because to exploit this channel one of the observers would have to be able to skew the distribution of hidden variables so that the measurement outcome for the other observer would be correlated with information they desire to send. What it takes to do that is, again, a question this toy model cannot address

nonlinear generalizations of the Schrödinger equation allow for superluminal signaling. Their answer is "yes" for deterministic evolutions and "in some cases" for stochastic dynamics. But the equation we use here is neither nonlinear, nor is it a Schrödinger equation.

because it does not specify the origin of the randomness. However, that so far no experiment has revealed any deviations from Born's rule suggests it is not easy to prepare a system in a state where the distribution of hidden variables does not reproduce quantum mechanics.[6]

Nevertheless, one can certainly speculate about different distributions of the random variables. This has been done, for example, for pilot-wave theory [55–57], where deviations from quantum equilibrium might be observable in astrophysical and cosmological tests [58,59].

## IV. ON FUTURE INPUT DEPENDENCE

The key property of the model discussed here is that the evolution of the prepared state depends on what the measurement setting will be at the time of measurement; it has hence a "future input dependence" [21,22,39]. It is easy to see how this can restore locality and determinism if we have a look at the sketch for Bell's notion of local causality (Fig. 3). Here, the **P** indicates the location of preparation, and $\mathbf{D}_1$ and $\mathbf{D}_2$ are the two detectors in a typical Bell-type experiment.

Local causality, in a nutshell, says that if all variables are "fully specified" in the backward light cone of $\mathbf{D}_1$, then whatever happens at $\mathbf{D}_2$ provides no further information (for a discussion of this point, please see [60]). The model we discussed here has the information about both detector settings at **P**. This means also that the information about the detector setting $\mathbf{D}_2$ is already in the backward light cone of $\mathbf{D}_1$ (and vice versa). Hence, whatever happens at $\mathbf{D}_2$ provides no further information. Note that this is a local requirement because the place of preparation is necessarily in causal contact with both detectors.

An often-raised question is how the hidden variables at **P** already "know" the future detector settings at $\mathbf{D}_1$ and $\mathbf{D}_2$. As for any scientific theory, we make assumptions to explain data. The assumptions of a theory are never explained within the theory itself (if they were, we wouldn't need them). Their scientific justification is that they allow one to make correct predictions. The question of how the hidden variables "know" something about the detector setting makes equally little sense as asking in quantum mechanics how one particle in a singlet state "knows" what happens to the other particle. In both cases that is just how the theory works. The difference is that in our toy model the required correlations are within the light cone and are hence compatible with general relativity.

Nevertheless, this is arguably a somewhat unusual feature, so we here want to offer a classical example for future input dependence to illustrate how it can come about by an optimization principle.

### A. Future input dependence and optimization principle: A classical example

A familiar example for an optimization principle is the principle of least action: Given a Lagrangian for a particle, as well as an initial and final position, we can calculate the trajectory by minimizing the action. In this case we can always express the evolution law as an Euler-Lagrange equation. This is not the type of optimization we will will consider here. We are interested instead in the case where we also allow the final point to vary. That is, we evaluate the action for all possible paths for all possible end points.

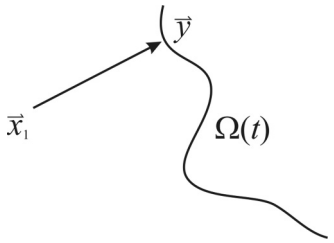


FIG. 4. Classical example for optimization problem that cannot be expressed by Euler-Lagrange equations.

For a quantum system, the optimal evolution should then be one in which the prepared state always ends up in one of the possible detector eigenstates. Note that this is *not* what happens in quantum mechanics. In quantum mechanics the state (generically) ends up in a superposition of detector eigenstates. The wave-function collapse then needs to be added as an *ad hoc* prescription, because the Schrödinger equation alone fails to give the result that we actually observe. To replace the collapse postulate with a principle of least action, one would need to know the right action to vary over. This is beyond the scope of this present paper. However, we can think of the toy model discussed in the previous section, which has the detector settings in the evolution law, as an effective description of an optimization over a final state.

To see how an optimization principle results in future input dependence, consider the following classical optimization problem. A particle with location $\vec{x}(t)$ moves in three-dimensional space in the absence of a potential, so we have the Lagrangian $\mathcal{L} := \dot{\vec{x}}^2/2$ and action $S := \int dt \mathcal{L}(\vec{x}, \dot{\vec{x}})$, where a dot denotes a time derivative, $d/dt$. In the usual principle of least action, we consider a start time, $t_1$, and an end time, $t_2$, and fix the start point and end point to $\vec{x}_1 := \vec{x}(t_1)$ and $\vec{x}_2 := \vec{x}(t_2)$. Then we can derive the Euler-Lagrange equations, which in this case come out to be simply $\ddot{\vec{x}} = 0$.

Now, however, we want to look at a different optimization problem. We start at time $t_1$ at place $x_1$, but at time $t_2$ the particle's location can be any point $\vec{y} \in \Omega$, where $\Omega$ is a two-dimensional surface (see Fig. 4). That is, in addition to the path itself, we also allow the end point to vary. What is the path of least action now? In this case the derivation of the Euler-Lagrange equations cannot be directly applied. We can, however, work around this.

For this, we first keep the end point fixed, as usual, and then integrate the Euler-Lagrange equation once. This gives $\dot{\vec{x}} = (\vec{y} - \vec{x}_1)/\Delta t$, where $\Delta t = t_2 - t_1$. We can insert this into the Lagrangian and integrate it to obtain the action as a function of $\vec{y}$. We can then vary the action again, this time over $\vec{y}$. The result is that the optimal path is the one to the point on $\Omega$ that is closest to $\vec{x}_1$. This is because we have kept the time difference fixed, so a shorter distance means a smaller velocity, which means a smaller action.

[6]In the model considered here, this would correspond to a nonuniform distribution of the $\lambda$s.

Let us denote this nearest point on $\Omega$ with $\vec{y}_{\rm n}$. We can then insert this into the once integrated Euler-Lagrange equation, which becomes $\dot{x} = (\vec{y}_{\rm n} - \vec{x}_1)/\Delta t$. The final step is then to allow that $\Omega$ could be a function of time, $\Omega(t)$, in which case the nearest point will also be a function of time $\vec{y}_{\rm n}(t)$. The point that minimizes the action is then the nearest point at time $t_2$. We finally get $\dot{x} = [\vec{y}_{\rm n}(t_2) - \vec{x}_1]/\Delta t$.

Of course it is utterly trivial to integrate this equation, but deriving this equation was not the point. The point was to show that it is possible to rewrite an optimization principle in which the end point is also varied into a differential equation. This equation will depend on information about the optimization constraint *in the future*, which we do not have. This does not mean that one cannot make a prediction. Even if we do not know the time dependence of $\Omega(t)$, we can very well predict that when we detect the particle, it will be at the point that was nearest to $x_1$.

This example is just meant to illustrate the general idea for why there is future input in the evolution law of the toy model. Loosely speaking, the surface $\Omega(t)$ is an analogy to the detector setting. The most important difference between the cases is that in quantum mechanics one cannot make a measurement before the final time, because the final time is, by definition, the time of measurement.

We want to stress that one does not need to know this future input to make predictions with the model. One simply solves the equation for any possible future input and makes a conditional prediction. (If input is this, then the prediction is that.) We do the same in quantum mechanics: When we make a calculation in normal quantum mechanics, we do not know what the measurement setting will be. We simply calculate what happens for any possible measurement setting. This is also what we do in our toy model.

The important difference between the toy model and quantum mechanics is that in quantum mechanics the time evolution of the state is independent of the measurement settings. The settings only enter in the calculation of probabilities at the time of measurement. This is the reason why one needs a nonlocal update of the wave function and why a realist interpretation of the wave function is difficult to reconcile with relativity. In our future-input-dependent toy model, on the other hand, the settings determine the local evolution, which avoids a nonlocal update upon measurements.

## V. DISCUSSION

One should not think of this model as a viable description of nature, because the way that the random variables enter the dynamics has no good motivation. It is a toy model for an effective limit of a more fundamental theory. The parameter $\kappa$, for example, defines a time scale for decoherence and should be understood as induced by the interaction with a larger system that has been integrated out. That is, $\kappa$ scales with the total size of the system, and in the limit when the prepared state has no further interaction, $\kappa = 0$, one just has normal quantum mechanics.

This is also why the model is not Lorenz invariant: the environment defines a preferred frame, yet the environment does not explicitly appear. The model further violates energy conservation. Again, this is because it stands in for an effective description that, among other things, does not take into account the recoil (and resulting entanglement with) parts of the experimental equipment. For the same reason, as mentioned in footnote 2, the time of measurement is not really *the* time of measurement but a time at which we would for all practical purposes say the measurement has been completed.

Furthermore, since the operators $L_{NK}$ are defined in terms of the eigenstates of the observable measured at the time of detection, we have still used an external definition of detection that is not included in the definition of the model. Now, the pointer basis defining these operators can be identified using the einselection methods introduced by Zurek [61] when one includes an environment, so this is not a problem *per se*. However, a more fundamental model is needed to explain how these operators come to couple to the prepared state the way that they do.

This toy model avoids nonlocal interactions by hard coding the dependence on the detector settings into the evolution law. This is another reasons why one should not take this model too seriously: A good, fundamental model should allow us to *derive* that the effective law for the prepared state depends on the detector settings. This requires that the to-be-found fundamental model includes the detectors and the environment and possibly other transformation devices that are part of the experimental setup. This has to be the case, because otherwise we would lack information to define what the detector eigenstates are.

All these issues are resolvable in principle, but given that this model is not intended to make a lot of sense, putting more effort into it seems not a good time investment. The reason for this little exercise was merely to demonstrate that there is nothing to fine tune here. For a Hilbert space of dimension $N$, we have $N - 1$ uniformly distributed, complex, random variables. Picking a set of specific values for these variables will in the limit $\kappa t \to \infty$ lead to one particular detector eigenstate—no collapse postulate necessary. But if we do not know what the values of the random variables are, we average over all possible values and get Born's rule. The model has one free parameter, $\kappa$, but its value does not matter as long as it is much larger than the typical energies of the system, so that the collapse dynamics is dominant. One may not like this construction for one reason or the other, but clearly there is no huge number of details to be specified here.

This works regardless of what the Hamiltonian is, how large the system is, how many detectors there are, or what observable is being measured. It will work for the double-slit,[7] for Stern-Gerlach, and for EPR-type experiments, no matter how the settings were chosen. In addition, the dynamical law that we have made up here is both deterministic and local. It is not like the prepared state actually interacts with the detector before measurement, it just has the required infor-

[7]Double-slit experiments are typically described using infinite Hilbert spaces, which we did not consider here. However, in reality, one never actually measures the particle position with infinite precision; one always has a finite grid of positions. Therefore, while it may be mathematically more convenient to use an infinite-dimensional Hilbert space, the system can in principle be described using a finite Hilbert space.

mation already in the dynamical law. We have found a local, deterministic, hidden-variable model that reproduces quantum mechanics on the average without a hint of conspiracy.

For what the economy of this model is concerned, one may debate whether replacing the collapse postulate with random variables is a good tradeoff. It certainly is a conceptual advantage, because it gives a well-defined probabilistic interpretation to quantum mechanics. However, if we merely count axioms, our toy model has no advantage over quantum mechanics. This is not surprising. If one believes in "shut up and calculate" and a probabilistic prediction is all one wants, then one may as well stick with quantum mechanics. But the here discussed toy model demonstrates that there is nothing wrong with theories that violate measurement independence *per se*, and that they are hence a viable route towards a theory more fundamental than quantum mechanics.

## VI. CONCLUSION

We have shown that fine tuning, which has frequently been used as an argument against superdeterminism, is unnecessary. To make predictions with this model, one does not need to introduce a huge number of delicately chosen parameters. This toy model shows that violating statistical independence is a scientifically sound option. Since violating statistical independence is a starting point for resolving the quantum measurement problem, this provides motivation for the further exploration of models with this property.

The model, when averaged over the uniformly distributed hidden variables $\lambda$, reproduces the quantum-mechanical predictions. If the variables were not uniformly distributed, for example, because the sample space is not large enough, this model would give predictions that differ from quantum mechanics. That is, even in an experiment where one expects to have a good enough sampling to reproduce Born's rule with only small statistical deviations, the observed distribution could still be heavily skewed in the space of hidden variables and hence not approximate Born's rule as expected. This makes clear that superdeterminism is not an interpretation of quantum mechanics; it is a concrete possibility for replacing quantum mechanics with a better theory, and more effort should be made to experimentally test it.

For example, following [62], consider the case in which the values of the hidden variables are determined by the degrees of freedom of the measurement device (other than the measurement setting itself). Then, if we perform multiple measurements of noncommuting observables using the same setup, but the degrees of freedoms of the device which determine the $\lambda$s only change very slowly between the measurements, we might end up always sampling the same region of the parameter space of the $\lambda$s, hence observing results which are more strongly correlated than quantum mechanics predicts.

Exactly when the deviations from quantum mechanics become non-negligible depends on what the hidden variables are; the above-introduced toy model cannot answer this question. The toy model is, therefore, strictly speaking, untestable, because it does not specify where the distribution of hidden variables comes from. But, as pointed out above and in more detail in Refs. [28,62], if the hidden variables are the degrees of freedom of the detector, it is reasonable to expect that minimizing the variation in the detectors' degrees of freedom between consecutive measurements will reveal deviations from Born's rule which cannot be detected by Bell-type experiments [45–47], no matter how ingeniously conceived and executed.

## ACKNOWLEDGMENTS

We gratefully acknowledge support from the Fetzer Franklin Fund. S.H. thanks Tim Palmer for helpful discussion.

## APPENDIX: GENERAL SOLUTION FOR $N = 4$

As previously, we work in the limit where $\kappa$ is much larger than the typical energies of the system; hence in very good approximation we can set $H = 0$. We begin with rewriting Eq. (15) more explicitly, inserting Eq. (14):

$$\begin{aligned}\partial_t \rho(t) &= \kappa \sum_{M>K=1}^{4} \left( L_{MK}\rho(t)L_{MK}^\dagger - \frac{1}{2}\{\rho(t), L_{MK}^\dagger L_{MK}\}\right) \\ &= \kappa \sum_{M>K=1}^{4} \Big[\theta(\sigma_M)\Big( |M\rangle\langle K|\rho(t)|K\rangle\langle M| - \frac{1}{2}\{\rho(t), |K\rangle\langle K|\}\Big) \\ &\quad + \theta(-\sigma_M)\Big( |K\rangle\langle M|\rho(t)|M\rangle\langle K| - \frac{1}{2}\{\rho(t), |M\rangle\langle M|\}\Big)\Big]. \end{aligned} \tag{A1}$$

We now introduce the notation $\rho_{IJ}(t) := \langle I|\rho(t)|J\rangle$ by use of which the master equation takes the form

$$\begin{aligned}\partial_t \rho_{IJ}(t) = \kappa \sum_{M>K=1}^{4} \Big[&\theta(\sigma_M)\Big(\delta_{IM}\delta_{MJ}\rho_{KK}(t) - \frac{1}{2}\delta_{KJ}\rho_{IK}(t) - \frac{1}{2}\delta_{IK}\rho_{KJ}(t)\Big) \\ &+ \theta(-\sigma_M)\Big(\delta_{IK}\delta_{KJ}\rho_{MM}(t) - \frac{1}{2}\rho_{IM}(t)\delta_{MJ} - \frac{1}{2}\delta_{IM}\rho_{MJ}(t)\Big)\Big]. \end{aligned} \tag{A2}$$

To apply the Kronecker delta, we recall that the double sum can be equivalently written as

$$\sum_{M>K}^{4} = \sum_{M=2}^{4}\sum_{K=1}^{M-1} = \sum_{K=1}^{3}\sum_{M=K+1}^{4}. \tag{A3}$$

Using this, we obtain for $I = J$,

$$\begin{aligned}
\partial_t \rho_{JJ}(t)/\kappa &= \sum_{M=2}^{4}\sum_{K=1}^{M-1}\theta(\sigma_M)\delta_{JM}\delta_{MJ}\rho_{KK}(t) - \frac{1}{2}\sum_{K=1}^{3}\sum_{M=K+1}^{4}\theta(\sigma_M)\delta_{KJ}\rho_{JK}(t) \\
&\quad - \frac{1}{2}\sum_{K=1}^{3}\sum_{M=K+1}^{4}\theta(\sigma_M)\delta_{JK}\rho_{KJ}(t) + \sum_{K=1}^{3}\sum_{M=K+1}^{4}\theta(-\sigma_M)\delta_{JK}\delta_{KJ}\rho_{MM}(t) \\
&\quad - \frac{1}{2}\sum_{M=2}^{4}\sum_{K=1}^{M-1}\theta(-\sigma_M)\rho_{JM}(t)\delta_{MJ} - \frac{1}{2}\sum_{M=2}^{4}\sum_{K=1}^{M-1}\theta(-\sigma_M)\delta_{JM}\rho_{MJ}(t) \\
&= (1-\delta_{J1})\theta(\sigma_J)\sum_{K=1}^{J-1}\rho_{KK}(t) - \frac{1}{2}(1-\delta_{J4})\rho_{JJ}(t)\sum_{M=J+1}^{4}\theta(\sigma_M) \\
&\quad - \frac{1}{2}(1-\delta_{J4})\rho_{JJ}(t)\sum_{M=J+1}^{4}\theta(\sigma_M) + (1-\delta_{J4})\sum_{M=J+1}^{4}\theta(-\sigma_M)\rho_{MM}(t) \\
&\quad - \frac{1}{2}(1-\delta_{J1})\sum_{K=1}^{J-1}\theta(-\sigma_J)\rho_{JJ}(t) - \frac{1}{2}(1-\delta_{J1})\sum_{K=1}^{J-1}\theta(-\sigma_J)\rho_{JJ}(t) \\
&= (1-\delta_{J1})\left(\theta(\sigma_J)\sum_{K=1}^{J-1}\rho_{KK}(t) - \sum_{K=1}^{J-1}\theta(-\sigma_J)\rho_{JJ}(t)\right) \\
&\quad + (1-\delta_{J4})\left(\sum_{M=J+1}^{4}\theta(-\sigma_M)\rho_{MM}(t) - \rho_{JJ}(t)\sum_{M=J+1}^{4}\theta(\sigma_M)\right) \\
&= (1-\delta_{J1})\left[\theta(\sigma_J)\sum_{K=1}^{J-1}\rho_{KK}(t) - \theta(-\sigma_J)\rho_{JJ}(t)(J-1)\right] \\
&\quad + (1-\delta_{J4})\sum_{M=J+1}^{4}[\theta(-\sigma_M)\rho_{MM}(t) - \theta(\sigma_M)\rho_{JJ}(t)].
\end{aligned} \tag{A4}$$

This can be written explicitly as

$$\begin{aligned}
\partial_t \rho_{11}(t)/\kappa &= \theta(-\sigma_2)\rho_{22}(t) + \theta(-\sigma_3)\rho_{33}(t) \\
&\quad + \theta(-\sigma_4)\rho_{44}(t) - \sum_{M=2}^{4}\theta(\sigma_M)\rho_{11}(t),
\end{aligned} \tag{A5}$$

$$\begin{aligned}
\partial_t \rho_{22}(t)/\kappa &= \theta(\sigma_2)\rho_{11}(t) - [\theta(-\sigma_2) + \theta(\sigma_3) + \theta(\sigma_4)]\rho_{22}(t) \\
&\quad + \theta(-\sigma_3)\rho_{33}(t) + \theta(-\sigma_4)\rho_{44}(t),
\end{aligned} \tag{A6}$$

$$\begin{aligned}
\partial_t \rho_{33}(t)/\kappa &= \theta(\sigma_3)\rho_{11}(t) + \theta(\sigma_3)\rho_{22}(t) \\
&\quad - [2\theta(-\sigma_3) + \theta(\sigma_4)]\rho_{33}(t) + \theta(-\sigma_4)\rho_{44}(t),
\end{aligned} \tag{A7}$$

$$\begin{aligned}
\partial_t \rho_{44}(t)/\kappa &= \theta(\sigma_4)\rho_{11}(t) + \theta(\sigma_4)\rho_{22}(t) \\
&\quad + \theta(\sigma_4)\rho_{33}(t) - 3\theta(-\sigma_4)\rho_{44}(t).
\end{aligned} \tag{A8}$$

For $I \neq J$, Eq. (A2) becomes

$$\begin{aligned}
\partial_t \rho_{IJ}(t)/\kappa &= \sum_{K=1}^{3}\sum_{M=K+1}^{4}\theta(\sigma_M)\left(-\frac{1}{2}\delta_{KJ}\rho_{IK}(t) - \frac{1}{2}\delta_{IK}\rho_{KJ}(t)\right) + \sum_{M=2}^{4}\sum_{K=1}^{M-1}\theta(-\sigma_M)\left(-\frac{1}{2}\rho_{IM}(t)\delta_{MJ} - \frac{1}{2}\delta_{IM}\rho_{MJ}(t)\right) \\
&= -\frac{1}{2}\left(\sum_{M=J+1}^{4}\theta(\sigma_M)(1-\delta_{J4})\rho_{IJ}(t) + \sum_{M=I+1}^{4}(1-\delta_{I4})\theta(\sigma_M)\rho_{IJ}(t)\right)
\end{aligned}$$

$$
\begin{aligned}
&-\frac{1}{2}\left((1-\delta_{J1})\sum_{K=1}^{J-1}\theta(-\sigma_J)\rho_{IJ}(t)+(1-\delta_{I1})\sum_{K=1}^{I-1}\theta(-\sigma_I)\rho_{IJ}(t)\right)\\
&=-\frac{1}{2}\Bigg[\left((1-\delta_{J4})\sum_{M=J+1}^{4}\theta(\sigma_M)+(1-\delta_{I4})\sum_{M=I+1}^{4}\theta(\sigma_M)\right)\\
&\quad+\left((1-\delta_{J1})\sum_{K=1}^{J-1}\theta(-\sigma_J)+(1-\delta_{I1})\sum_{K=1}^{I-1}\theta(-\sigma_I)\right)\Bigg]\rho_{IJ}(t)\\
&=-\frac{1}{2}\Bigg[\left((1-\delta_{J4})\sum_{M=J+1}^{4}\theta(\sigma_M)+(1-\delta_{I4})\sum_{M=I+1}^{4}\theta(\sigma_M)\right)\\
&\quad+((1-\delta_{J1})(J-1)\theta(-\sigma_J)+(1-\delta_{I1})(I-1)\theta(-\sigma_I))\Bigg]\rho_{IJ}(t). \qquad \text{(A9)}
\end{aligned}
$$

With that we can write for $I \neq J$,

$$\partial_t \rho_{IJ}(t) = -\frac{\kappa}{2}\Lambda_{IJ}\rho_{IJ}(t), \qquad \text{(A10)}$$

where

$$
\begin{aligned}
\Lambda_{IJ} :=& \left((1-\delta_{J4})\sum_{M=J+1}^{4}\theta(\sigma_M)+(1-\delta_{I,4})\sum_{M=I+1}^{4}\theta(\sigma_M)\right)\\
&+[(1-\delta_{J1})(J-1)\theta(-\sigma_J)\\
&+(1-\delta_{I1})(I-1)\theta(-\sigma_I)]. \qquad \text{(A11)}
\end{aligned}
$$

Equation (A10) can be integrated straightforwardly to

$$\rho_{IJ}(t) = \rho_{IJ}(0)\, e^{-\frac{\kappa}{2}\Lambda_{IJ}t}. \qquad \text{(A12)}$$

For the singlet state we always have $\lambda_{3,4} \leqslant 0$, so Eqs. (A5)–(A8) simplify to

$$\partial_t \rho_{11}(t) = \kappa[\theta(-\sigma_2)\rho_{22}(t)+\rho_{33}(t)+\rho_{44}(t)-\theta(\sigma_2)\rho_{11}(t)], \qquad \text{(A13)}$$

$$\partial_t \rho_{22}(t) = \kappa[\theta(\sigma_2)\rho_{11}(t)-\theta(-\sigma_2)\rho_{22}(t)+\rho_{33}(t)+\rho_{44}(t)], \qquad \text{(A14)}$$

$$\partial_t \rho_{33}(t) = -2\kappa\rho_{33}(t)+\kappa\rho_{44}(t), \qquad \text{(A15)}$$

$$\partial_t \rho_{44}(t) = -3\kappa\rho_{44}(t). \qquad \text{(A16)}$$

We solve these equations starting from the last one:

$$\rho_{44}(t) = \rho_{44}(0)e^{-3\kappa t}. \qquad \text{(A17)}$$

This can be inserted into the third equation and gives

$$\rho_{33}(t) = -\rho_{44}(0)e^{-3\kappa t}+[\rho_{33}(0)+\rho_{44}(0)]e^{-2\kappa t}. \qquad \text{(A18)}$$

Inserting the initial state (34) now reveals that the solution to Eqs. (A17) and (A18) is simply $\rho_{33}(t) = \rho_{44}(t) = 0$.

Then the first two equations are

$$\partial_t \rho_{11}(t) = \kappa\theta(-\sigma_2)\rho_{22}(t)-\kappa\theta(\sigma_2)\rho_{11}(t), \qquad \text{(A19)}$$

$$\partial_t \rho_{22}(t) = \kappa\theta(\sigma_2)\rho_{11}(t)-\kappa\theta(-\sigma_2)\rho_{22}(t), \qquad \text{(A20)}$$

which can be integrated to

$$
\begin{aligned}
\rho_{11}(t) =& \theta(\sigma_2)\rho_{11}(0)e^{-\kappa t}+\theta(-\sigma_2)[\rho_{11}(0)\\
&+\rho_{22}(0)(1-e^{-\kappa t})], \qquad \text{(A21)}
\end{aligned}
$$

$$
\begin{aligned}
\rho_{22}(t) =& \theta(\sigma_2)[\rho_{22}(0)+\rho_{11}(0)(1-e^{-\kappa t})]\\
&+\theta(-\sigma_2)\rho_{22}(0)e^{-\kappa t}. \qquad \text{(A22)}
\end{aligned}
$$

After inserting the initial state $\rho_{11}(0) = q^2$ and $\rho_{22}(0) = 1-q^2$ we get for the diagonal elements

$$\rho_{11}(t) = \theta(\sigma_2)q^2e^{-\kappa t}+\theta(-\sigma_2)[1-(1-q^2)e^{-\kappa t}], \qquad \text{(A23)}$$

$$\rho_{22}(t) = \theta(\sigma_2)(1-q^2e^{-\kappa t})+\theta(-\sigma_2)(1-q^2)e^{-\kappa t}. \qquad \text{(A24)}$$

For the off-diagonal elements of the density matrix, we note that all contributions to $\Lambda_{IJ}$ for $I, J$ equal to 3 or 4 vanish and we only have one nonzero matrix element, which is

$$\Lambda_{12} = \theta(\sigma_2)+\theta(-\sigma_2) = 1, \qquad \text{(A25)}$$

then

$$\rho_{12}(t) = -q\sqrt{1-q^2}\, e^{-\frac{\kappa}{2}t}. \qquad \text{(A26)}$$

Equations (A23), (A24), and (A26) are used in Eq. (36).

[1] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables, I and II, Phys. Rev. **85**, 166 (1952).

[2] G. C. Ghirardi, A. Rimini, and T. Weber, Unified dynamics for microscopic and macroscopic systems, Phys. Rev. D **34**, 470 (1986).

[3] A. Bassi and G. C. Ghirardi, Dynamical reduction models, Phys. Rep. **379**, 257 (2003).

[4] E. Nelson, Review of stochastic mechanics, J. Phys.: Conf. Ser. **361**, 012011 (2012).

[5] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Phys. Phys. Fiz. **1**, 195 (1964).

[6] C. H. Brans, Bell's theorem does not eliminate fully causal hidden variables, Int. J. Theor. Phys. **27**, 219 (1988).

[7] T. N. Palmer, A local deterministic model of quantum spin measurement, Proc. Roy. Soc. A **451**, 585 (1995).

[8] J. Degorre, S. Laplante, and J. Roland, Simulating quantum correlations as a distributed sampling problem, Phys. Rev. A **72**, 062314 (2005).

[9] T. N. Palmer, The invariant set postulate: A new geometric framework for the foundations of quantum theory and the role played by gravity, Proc. Roy. Soc. London, Ser. A **465**, 3165 (2009).

[10] M. J. W. Hall, Local Deterministic Model of Singlet State Correlations Based on Relaxing Measurement Independence, Phys. Rev. Lett. **105**, 250404 (2010).

[11] G. S. Ciepielewski, E. Okon, and D. Sudarsky, On superdeterministic rejections of settings independence, arXiv:2008.00631.

[12] M. J. W. Hall, The significance of measurement independence for Bell inequalities and locality, in *At the Frontier of Spacetime* (Springer, Switzerland, 2016), pp. 189–204, Ch. 11.

[13] P. A. M. Dirac, On the analogy between classical and quantum mechanics, Rev. Mod. Phys. **17**, 195 (1945).

[14] J. A. Wheeler and R. P. Feynman, Interaction with the absorber as the mechanism of radiation, Rev. Mod. Phys. **17**, 157 (1945).

[15] D. T. Pegg, Objective reality, causality and the aspect experiment, PLA **78**, 233 (1980).

[16] D. T. Pegg, Absorber theory in quantum optics, Phys. Scr. **T12**, 14 (1986).

[17] J. G. Cramer, The transactional interpretation of quantum mechanics, Rev. Mod. Phys. **58**, 647 (1986).

[18] R. E. Kastner, *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility* (Cambridge University Press, Cambridge, England, 2013).

[19] H. Price, Does time-symmetry imply retrocausality? How the quantum world says "Maybe"? Stud. Hist. Philos. Sci. Part B: Stud. Hist. Philos. Mod. Phys. **43**, 75 (2012).

[20] P. D. Drummond and M. D. Reid, Objective quantum fields, retrocausality and ontology, Entropy **23**, 749 (2021).

[21] N. Argaman, Bell's theorem and the causal arrow of time, Am. J. Phys. **78**, 1007 (2010).

[22] L. S. Schulman, *Time's Arrows and Quantum Measurement* (Cambridge University Press, Cambridge, England, 1997).

[23] G. Lindblad, On the generators of quantum dynamical semigroups, Commun. Math. Phys. **119**, 48 (1976).

[24] V. Gorini, A. Kossakowski, and E. C. Sudarsahan, Completely positive semigroups of n-level systems, J. Math. Phys. **17**, 821 (1976).

[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).

[26] S.L. Adler, D. C. Brody, T. A. Brun, and L. P. Hughston, Martingale models for quantum state reduction, J. Phys. A: Math. Gen. **34**, 8795 (2001).

[27] S. Hossenfelder, Born's rule from almost nothing, arXiv:2006.14175 [quant-ph]; A derivation of Born's rule from symmetry, Ann. Phys. **425**, 168394 (2021).

[28] S. Hossenfelder and T. Palmer, Rethinking superdeterminism, Front. Phys. **8**, 139 (2020).

[29] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[30] E. G. Cavalcanti and H. M. Wiseman, Bell nonlocality, signal locality and unpredictability (or what Bohr could have told Einstein at Solvay had he known about Bell experiments), Found. Phys. **42**, 1329 (2012).

[31] H. M. Wiseman, The two Bell's theorems of John Bell, J. Phys. A: Math. Theor. **47**, 424001 (2014).

[32] J. S. Bell, The theory of local beables, Rep. No. CERN-TH-2053, CERN (1975).

[33] H. Price, Locality, independence and the pro-liberty bell, arXiv:quant-ph/9602020.

[34] P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, Prediction, and Search* (MIT Press, Cambridge, MA, 2001).

[35] J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, Cambridge, England, 2009).

[36] C. J. Wood and R. W. Spekkens, The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning, New J. Phys. **17**, 033002 (2015).

[37] S. Hossenfelder, Superdeterminism: A guide for the perplexed, arXiv:2010.01324.

[38] D. Schmid, J. H. Selby, and R. W. Spekkens, Unscrambling the omelette of causation and inference: The framework of causal-inferential theories, arXiv:2009.03297.

[39] K. B. Wharton and N. Argaman, Bell's theorem and spacetime-based reformulations of quantum mechanics, Rev. Mod. Phys. **92**, 021002 (2020).

[40] E. Adlam, Spooky action at a temporal distance, Entropy **20**, 41 (2018).

[41] E. Adlam, Contextuality, Fine-Tuning, and Teleological Explanation, arXiv:2110.15898.

[42] J. R. Hance and S. Hossenfelder, The wave-function as a true ensemble, Proc. R. Soc. London, Ser. A, **478**, 20210705 (2022).

[43] A. Shimony, M. A. Horne, and J. F. Clauser, Comment on "The theory of local beables," Epistemological Lett. **13**, 1 (1976).

[44] J. S. Bell, Free variables and local causality, Epistemological Lett. **15**, (1977).

[45] C. Leung, A. Brown, H. Nguyen, A. S. Friedman, D. I. Kaiser, and J. Gallicchio, Astronomical random numbers for quantum foundations experiments, Phys. Rev. A **97**, 042120 (2018).

[46] The BIG Bell Test Collaboration, Challenging local realism with human choices, Nature (London) **557**, 212 (2018).

[47] A. S. Friedman, A. H. Guth, M. J. W. Hall, D. I. Kaiser, and J. Gallicchio, Relaxed Bell inequalities with arbitrary measurement dependence for each observer, Phys. Rev. A **99**, 012121 (2019).

[48] S. Hossenfelder, Screams for explanation: Finetuning and naturalness in the foundations of physics, Synthese **198** (Suppl 16), 3727 (2021).

[49] S. Hossenfelder, *Lost in Math* (Basic Books, New York, 2018).

[50] I. Sen and A. Valentini, Superdeterministic hidden-variables models I: Nonequilibrium and signalling, Proc. R. Soc. London, Ser. A **476**, 20200212 (2020).

[51] I. Sen and A. Valentini, Superdeterministic hidden-variables models II: Conspiracy, Proc. R. Soc. London, Ser. A **476**, 20200214 (2020).
[52] E. K. Chen, *The Routledge Companion to Philosophy of Physics* (Routledge, 2021), pp. 184–199.
[53] N. Gisin, Stochastic quantum dynamics and relativity, Helv. Phys. Acta **62**, 363 (1989).
[54] A. Bassi and K. Hejazi, No-faster-than-light-signaling implies linear evolution. A re-derivation, Eur. J. Phys. **36**, 055027 (2015).
[55] A. Valentini, Signal-locality, uncertainty, and the subquantum H-theorem. I, Phys. Lett. A **156**, 5 (1991).
[56] A. Valentini, Signal-locality, uncertainty, and the subquantum H-theorem. II, Phys. Lett. A **158**, 1 (1991).
[57] A. Valentini, Hidden variables, statistical mechanics and the early universe, in *Chance in Physics* (Springer, New York, 2001), pp. 165–181.
[58] A. Valentini, Astrophysical and cosmological tests of quantum theory, J. Phys. A: Math. Theor. **40**, 3285 (2007).
[59] N. G. Underwood and A. Valentini, Anomalous spectral lines and relic quantum nonequilibrium, Phys. Rev. D **101**, 043004 (2020).
[60] T. Norsen, John S. Bell's concept of local causality, Am. J. Phys. **79**, 1261 (2011).
[61] W. H. Zurek, Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? Phys. Rev. D **24**, 1516 (1981).
[62] S. Hossenfelder, Testing super-deterministic hidden variables theories, Found. Phys. **41**, 1521 (2011).