Superdeterministic hidden-variables models II: arbitrary conspiracy

Indrajit Sen\textsuperscript{1} and Antony Valentini\textsuperscript{1,2}\textsuperscript{*}

\textsuperscript{1}Department of Physics and Astronomy, Clemson University, Kinard Laboratory, Clemson, SC 29634, USA
\textsuperscript{2}Augustus College, 14 Augustus Road, London SW19 6LN UK

(Dated: March 30, 2020)

We prove that superdeterministic models of quantum mechanics are arbitrarily conspiratorial in a mathematically well-defined sense, by further development of the ideas presented in a previous article \textit{A}. We consider a Bell scenario where, in each run and at each wing, the experimenter chooses one of \(N\) devices to determine the local measurement setting. We prove, without assuming any features of quantum statistics, that superdeterministic models of this scenario must have a finely-tuned distribution of hidden variables. Specifically, fine-tuning is required in order for the measurement statistics to not depend on the outputs of the unused devices, or on the experimenter’s choice of which device to use. We quantify this as the overhead fine-tuning \(F\) of the model, and show that \(F \rightarrow 1\) (corresponding to ‘completely fine-tuned’) exponentially fast as \(N \rightarrow \infty\). The notion of fine-tuning assumes that arbitrary (’nonequilibrium’) hidden-variables distributions are possible in principle. We also show how to quantify the conspiracy from a different viewpoint, which requires us to consider only the equilibrium distribution for a given hidden-variables model. We define a formal entropy for superdeterministic models of our scenario, and show that it spontaneously decreases with time. The entropy drop \(\Delta S \rightarrow \infty\) logarithmically fast as \(N \rightarrow \infty\). Both approaches prove that superdeterministic models become arbitrarily conspiratorial as \(N \rightarrow \infty\). We thus quantitatively confirm Bell’s intuition that superdeterministic models are physically implausible.

I. INTRODUCTION

In a previous article \textit{A} \cite{1}, we gave a broad overview of superdeterministic models and criticisms thereof. We also discussed the properties of these models in nonequilibrium. In particular, we showed that nonequilibrium extensions of superdeterministic models have two striking properties. First, the measurement statistics have a general dependence on the mechanism used to determine the measurement settings. This implies that the equilibrium distribution has to be finely-tuned such that these effects disappear in practice. Second, although these models violate marginal-independence (that is, they violate formal no-signalling), there is no information transfer between the wings. Instead, the local outcomes and distant settings are only statistically correlated due to initial conditions. We argued that this mimicking of a signal using statistical correlations is intuitively conspiratorial. Both these properties were discussed qualitatively in \textit{A}. In this paper, we further develop these ideas mathematically and give two separate ways to quantify the conspiratorial character of superdeterministic models.

The paper is structured as follows. In section \textbf{II} we begin with a preliminary discussion which defines the experimental setup on which our arguments are based. This section also defines the constraints that must be imposed on an arbitrary distribution for a superdeterministic model of the setup in order to satisfy some basic empirical observations (unrelated to quantum statistics). We then show in section \textbf{III} that these constraints arbitrarily fine-tune the hidden-variables distribution in a certain limit. In section \textbf{IV} we propose a different approach to quantify the conspiracy which foregoes any appeal to nonequilibrium. This is based on a generalisation of the idea that the mimicking of a physical signal by statistical correlations is intuitively conspiratorial. We show that, even when restricted to equilibrium, a superdeterministic model of our setup suffers from a spontaneous entropy drop at the hidden-variables level, which becomes arbitrarily large in the same limit as in the previous approach. In the concluding section \textbf{V} we discuss our results. The Appendix contains a proof of principle that the fine-tuning argument can be extended to a continuous hidden-variables model by considering a discrete approximation.

\textsuperscript{*}Electronic address: isen@clemson.edu
\textsuperscript{†}Electronic address: antonyv@clemson.edu
II. PRELIMINARY DISCUSSION

Consider a standard Bell scenario \[\mathcal{B}\], where two spin-1/2 particles prepared in the spin-singlet state are each subjected to one of two local spin measurements, say \(M_x\) (corresponding to \(\sigma_z\)) or \(M_z\) (corresponding to \(\sigma_z\)), in a spacelike separated manner. The local measurement settings at both wings are chosen using a *setting mechanism*, which is defined as any physical system that outputs a measurement setting in a stochastic manner. A simple example is a dice whose faces have been painted either \(M_x\) or \(M_z\). An experimenter can roll the dice for each run and use its output to choose the local measurement setting. Other examples of setting mechanisms can include coins, random number generators, humans choosing measurement settings, and so on. The key point is that the output of the setting mechanism (that is, the measurement setting) must be a random variable for practical purposes. It makes no difference to our argument whether the randomness is intrinsic or due to our ignorance of the internal working of the setting mechanism.

We suppose that the experimenter at each wing has a setup consisting of \(N\) different setting mechanisms, which collectively give \(N\) outputs for each run. The experimenter then chooses one of the outputs and uses it as the actual measurement setting for that run (how he makes this choice has no relevance to our argument, so it can be specified arbitrarily). Nevertheless, the outputs of *all* the \(N\) setting mechanisms are recorded for each run at both the wings.

We label the hidden variables that determine the measurement outcomes at both wings by \(\lambda\). The output of the \(i^{th}\) \((j^{th})\) setting mechanism at wing A (B) is labelled by \(\alpha_i\) \((\beta_j)\), where \(i, j \in \{1, 2, \ldots, N\}\) and \(\alpha_i, \beta_j \in \{M_x, M_z\}\). We label the experimenters’ choice of the outputs by \(\gamma\) in the following manner: if the experimenters choose the \(i^{th}\) setting mechanism at wing A and the \(j^{th}\) setting mechanism at wing B, then we have \(\gamma = \gamma_{ij}\).

Labelling the actual measurement settings at wing A (B) by \(M_a\) \((M_b)\), where \(M_a, M_b \in \{M_x, M_z\}\), the hidden-variables distribution of a superdeterministic model can be expanded as

\[
\rho(\lambda|M_a, M_b) = \sum_{\{\alpha\}, \{\beta\}, \gamma} \rho(\lambda|M_a, \{\alpha\}, \{\beta\}, \gamma)p(\{\alpha\}, \{\beta\}, \gamma|M_a, M_b)
\]  

(1)

where \(\{\alpha\} = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}\) and \(\{\beta\} = \{\beta_1, \beta_2, \ldots, \beta_N\}\). The first term \(\rho(\lambda|M_a, M_b, \{\alpha\}, \{\beta\}, \gamma)\) denotes the hidden-variables distribution given the outputs \(\{\alpha\}\) and \(\{\beta\}\) of all the setting mechanisms, the choice \(\gamma\) made by the experimenters, and the final measurement settings \(M_a\) and \(M_b\). The second term \(p(\{\alpha\}, \{\beta\}, \gamma|M_a, M_b)\) denotes the probability of obtaining a particular set of \((\{\alpha\}, \{\beta\}, \gamma)\) given the final measurement settings \(M_a\) and \(M_b\). This expression can be simplified to

\[
\rho(\lambda|M_a, M_b) = \sum_{\{\alpha\}, \{\beta\}, \gamma} \rho(\lambda|\{\alpha\}, \{\beta\}, \gamma)p(\{\alpha\}, \{\beta\}, \gamma|M_a, M_b)
\]  

(2)

where we have used the fact that \(M_a\) and \(M_b\) are completely determined by \((\{\alpha\}, \{\beta\}, \gamma)\), that is, \(M_a = M_a(\{\alpha\}, \{\beta\}, \gamma)\) and \(M_b = M_b(\{\alpha\}, \{\beta\}, \gamma)\). From \(A\) we know that \(\rho(\lambda|\{\alpha\}, \{\beta\}, \gamma)\) can be a nonequilibrium distribution. Let us consider the most general \(\rho(\lambda|\{\alpha\}, \{\beta\}, \gamma)\) possible. We will derive constraints on this general distribution from some empirical observations.

For the set of runs where \(\gamma = \gamma_{ij}\), consider two further subsets where for the first subset \(\alpha_k = M_x\) \((k \neq i)\) and \(\beta_l = M_x\) \((l \neq j)\), whereas for the second subset \(\alpha_k = M_z\) and \(\beta_l = M_z\). That is, we are subdividing the set of runs where \(\gamma = \gamma_{ij}\) into subsets according to the outputs of setting mechanisms that have *not* been used to determine the measurement settings. If \(\lambda\) is correlated with \(\alpha_k\) and \(\beta_l\) in this set (for which \(\gamma = \gamma_{ij}\)), then

\[
\rho(\lambda|\{\alpha_i = M_a, \ldots, \alpha_k = M_x\}, \{\beta_j = M_b, \ldots, \beta_l = M_x\}, \gamma_{ij}) \neq \rho(\lambda|\{\alpha_i = M_a, \ldots, \alpha_k = M_z\}, \{\beta_j = M_b, \ldots, \beta_l = M_z\}, \gamma_{ij})
\]  

(3)

Clearly, if equation (3) is true, then the measurement statistics for both these subsets will differ in general, that is, the outputs of the *unused* setting mechanisms will affect the measurement statistics. However, we know from experimental data that there is no such difference in the measurement statistics. We consider the simplest possible explanation for the observed invariance of the measurement statistics with respect to the outputs of the unused setting mechanisms.
in a superdeterministic model: $\lambda$ is correlated only with $\alpha_i$ and $\beta_j$ when $\gamma = \gamma_{ij}$. That is,

$$\rho(\lambda|\{\alpha\}, \{\beta\}, \gamma_{ij}) = \rho(\lambda|\alpha_i, \beta_j, \gamma_{ij}) \tag{4}$$

It follows from equation (4) that there exist pairs of particles which are correlated only with the outputs $\alpha_i$ and $\beta_j$, and further these pairs are measured only when the experimenters choose $\gamma = \gamma_{ij}$. We may say that these particles belong to a sub-ensemble $E_{ij}$ having the hidden-variables distribution $\rho(\lambda|\alpha_i, \beta_j, \gamma_{ij})$. There are $N^2$ such sub-ensembles, each with its own hidden-variables distribution.

Furthermore, the distributions $\rho(\lambda|\alpha_i, \beta_j, \gamma_{ij})$ can all differ from each other, being logically independent. However, as before this implies that the corresponding measurement statistics will also differ in general: using different setting mechanisms will lead to different measurement statistics. We again consider the simplest possible explanation for the empirical fact that the measurement statistics are independent of the experimenters’ choice of $\gamma$: that these distributions are subject to the constraint

$$\rho(\lambda|\alpha_i = M_a, \beta_j = M_b, \gamma_{ij}) = \rho(\lambda|\alpha_k = M_a, \beta_l = M_b, \gamma_{kl}) = \rho(\lambda|M_a, M_b) \tag{5}$$

where $i \neq k$ and $j \neq l$. Thus the usual form of the hidden-variables distribution $\rho(\lambda|M_a, M_b)$ is applicable only after the general distribution $\rho(\lambda|\{\alpha\}, \{\beta\}, \gamma)$ has been constrained by equations (4) and (5).

It is logically possible to conceive a much more complicated superdeterministic model which avoids constraints (4) and (5). For example, say that a model is configured such that in equation (3) the two distributions are different but no such model has been explicitly proposed in the literature. We therefore do not discuss such models in this paper.

Note that $\rho(\lambda|M_a, M_b)$ need not be the equilibrium distribution, since we have not imposed the condition that our model reproduces the predictions of quantum mechanics. Equations (4) and (5) respectively constrain the superdeterministic model only so that the measurement statistics do not depend upon $a)$ the outputs of unused setting mechanisms, and $b)$ the experimenter’s choice of setting mechanism. In the next section we study the effect of these constraints on the hidden-variables distribution.

### III. QUANTIFICATION I: OVERHEAD FINE-TUNING

We begin by giving our definition of overhead fine-tuning for a general hidden-variables model of our setup (including non-superdeterministic models). We first assume for simplicity that $\lambda$ is a discrete variable, but our argument generalises to the continuous case under appropriate limits (see the Appendix). Let the number of possible values of $\lambda$ be $\Lambda$. From the ‘excess-baggage’ theorem [3], we know that $\Lambda$ has to be infinite. Therefore, we prove our results in the limit $\Lambda \to \infty$.

Let $\Omega$ be the total number of independent distributions in the model. In specifying these $\Omega$ distributions, we have a total of $\Omega(\Lambda - 1)$ parameters, after taking into account the normalisation constraints. Suppose $\kappa$ constraints are imposed on these $\Omega$ distributions in order to reproduce the following two empirical observations: 

a) the measurement statistics are independent of the outputs of the unused setting mechanisms.

b) the measurement statistics are independent of the experimenters’ choice of which setting mechanisms to use.

We then define the ratio

$$F = \frac{\kappa}{\Omega(\Lambda - 1)} \tag{6}$$

as the overhead fine-tuning of the hidden-variables model. If $F = 0$ we call the model completely general, as the model does not need to impose any constraint to reproduce the observations $a)$ and $b)$. On the other hand if $F = 1$ we call the model completely fine-tuned, as the model has no parameters left after imposing the constraints. Any value of $F$ outside $[0, 1]$ reflects inconsistency in the model. We now prove that for the superdeterministic model discussed in the previous section, $F \to 1$ exponentially fast as $N \to \infty$ (where $N$ is the number of setting mechanisms).
First we note that, since the variables \( \{ \alpha \} \) and \( \{ \beta \} \) can each take \( 2^N \) values whereas \( \gamma \) can take \( N^2 \) values, there are \( \Omega = 2^N \times 2^N \times N^2 \) independent distributions \( \rho(\lambda|\{ \alpha \}, \{ \beta \}, \gamma) \). Let us now count the number of constraints imposed by equations (4) and (5). Replacing the density \( \rho(\lambda|\{ \alpha \}, \{ \beta \}, \gamma) \) by the discrete distribution \( p(\lambda|\{ \alpha \}, \{ \beta \}, \gamma) \) in equation (4), \( \lambda \) can take \( \Lambda \) values (where it is understood that \( \Lambda \to \infty \)). Meanwhile, \( \alpha_i \) and \( \beta_j \) can each take 2 values, and \( \gamma \) can take \( N^2 \) values. Therefore, in equation (4), \( 2^N \times 2^N \times N^2(\Lambda - 1) \) parameters on the left-hand side are substituted by \( 2 \times 2 \times N^2(\Lambda - 1) \) parameters on the right-hand side. The number of constraints imposed is \( N^2(2^N - 4)(\Lambda - 1) \). Similarly, equation (5) replaces \( 2 \times 2 \times N^2(\Lambda - 1) \) parameters on the left-hand side by \( 2 \times 2 \times (\Lambda - 1) \) parameters on the right-hand side, and so the number of constraints imposed is equal to \( 4(N^2 - 1)(\Lambda - 1) \). Thus the total number of constraints is given by

\[
\kappa = N^2(2^N - 4)(\Lambda - 1) + 4(N^2 - 1)(\Lambda - 1)
\]

From our definition (6), the degree of fine-tuning is then

\[
F = 1 - \frac{1}{N^2 \times 2^{2(N-1)}}
\]

It is clear from equation (9) that \( F \to 1 \) exponentially fast as \( N \to \infty \). This arbitrarily high fine-tuning is required because the measurement statistics depend crucially on the correlations between \( \lambda \) and the variables \( \{ \{ \alpha \}, \{ \beta \}, \gamma \} \) which determine the measurement settings. In the Appendix, we prove that if \( \lambda \) is a continuous variable we can retrieve the above result for arbitrarily close discrete approximations to \( \lambda \). Note that in deriving equation (4) we did not impose any condition on the distribution \( p(\lambda|M_\alpha, M_\beta) \). That is, we did not impose features of quantum statistics – such as the Born rule, marginal-independence (formal no-signalling), or non-contextuality – to argue that superdeterministic models are arbitrarily conspiratorial. Thus, even nonequilibrium distributions which do not reproduce quantum statistics, if they reproduce the observations \( a \) and \( b \), must be arbitrarily fine-tuned according to our definition.

IV. QUANTIFICATION II: SPONTANEOUS ENTROPY DROP

In the previous section we proved that superdeterministic models of our setup are arbitrarily fine-tuned: \( F \to 1 \) in the limit \( N \to \infty \). However, the definition of \( F \) rests on the concept of quantum nonequilibrium as the total number of parameters needed to specify all of the independent distributions, \( \Omega(\Delta - 1) \), is defined by including contributions from all possible distributions – including those which do not reproduce quantum statistics. We now show how to quantify the conspiracy of superdeterministic models without having to invoke the possibility (in principle) of quantum nonequilibrium.\(^2\) For this purpose, we first remind ourselves of the conspiratorial nature of apparent-signalling in superdeterminism, as discussed in \( A \). The violation of marginal-independence (or formal no-signalling) is not equivalent to a transfer of information between the wings in a superdeterministic context. Instead, the apparent-signal consists entirely of a statistical correlation, set up by initial conditions, between the local marginals and the distant settings. Thus a physical signal can be mimicked by superdeterministic models by invoking coincidences set up by initial conditions. We will now argue that such coincidences are not restricted to nonequilibrium or signalling, but present in superdeterministic models generally.

Suppose the results of a pseudo-random number generator are used to determine the measurement settings in a Bell experiment. Then, in a superdeterministic model the numbers generated must be statistically correlated with \( \lambda \) due to initial conditions. After a certain number of runs, let a radioactive device be used to choose the measurement settings. Then, for the superdeterministic model to still hold, the radioactive decay occurring in that device must become correlated with \( \lambda \). Next, perhaps a coin toss can be used to choose the measurement settings. Then, the toss results must become correlated with \( \lambda \). One can go on, and keep changing the setting mechanism indefinitely, and \( \lambda \) must promptly become correlated with it. This strikes one as conspiratorial intuitively. Let us make this intuition

\(^2\) Although there are compelling theoretical reasons \(^4\) that motivate nonequilibrium extensions of hidden-variables models, a debate about nonequilibrium, for the purposes of this article, is secondary to proving that superdeterministic models are unequivocally conspiratorial.
precise by considering a superdeterministic model of our setup (restricting ourselves to quantum equilibrium). The hidden-variables distribution must satisfy the constraints represented by equations (4) and (5). Consider the pairs of particles that belong to the sub-ensemble $E_{ij}$. The particles in this sub-ensemble are correlated only to the outputs $\alpha_i$ and $\beta_j$. Further, these particles are measured only when the experimenters choose $\gamma = \gamma_{ij}$. However, while choosing $\gamma$ the experimenters themselves do not know which pair of particles belongs to which sub-ensemble. Neither can the later event ‘experimenters choose $\gamma$’ causally affect the hidden-variables state at the earlier preparation event (without retrocausality). Therefore, one can ask if it is possible for the experimenters to choose the wrong setting mechanisms, say the $k^{th}$ and $l^{th}$ setting mechanisms ($k \neq i, l \neq j$), to determine the measurement settings for particles belonging to the sub-ensemble $E_{ij}$. If this happens, the correlation between $\lambda$ and the measurement settings will disappear, since $\alpha_k$ and $\beta_l$ will determine the measurement settings $M_a$ and $M_b$ respectively, whereas the particles are correlated only to $\alpha_i$ and $\beta_j$. The Bell correlations will not be reproduced in this case. Such a mismatch is precluded, however, by the constraint represented by equation (4). This means that, for a superdeterministic model whose hidden-variables distribution satisfies equation (4), the experimenters’ choices are determined from the past in such a manner that they coincidentally always pick the right combinations of particles and setting mechanisms. We now show that this constitutes a spontaneous decrease of an appropriately-defined formal ‘entropy’ at the hidden-variables level.

Let there be $N_0$ number of spin-singlet pairs, and therefore $N_0$ runs in the experiment. Let the two experimenters determine the sequence in which they will use the various setting mechanisms, that is, they decide the value of $\gamma$ for each run beforehand. Now they need to determine the sequence in which they will use the various spin-singlet pairs. We know that, in a superdeterministic model of our setup, each pair belongs to a particular sub-ensemble $E_{ij}$. To reproduce the Bell correlations, the experimenters have to ensure that there are no ‘mismatches’. To do this, they need to collect together all the pairs belonging to each sub-ensemble $E_{ij}$, and then arrange them so that they are used only for those runs where $\gamma = \gamma_{ij}$.

However, the experimenters do not have access to the information ‘which pair belongs to which sub-ensemble’. They have access to only macroscopic information, in particular the order in which the setting mechanisms are to be used (the value of $\gamma$ for each run), and the measurement statistics after the experiment is over. If the measurement statistics satisfy the Bell correlations, the experimenters will conclude that the matching was successful; otherwise the experimenters will conclude that the matching was unsuccessful. Initially, they cannot know if the pairs are already arranged in a correct or incorrect manner. For example, if it were known that the initial unsorted sequence of spin-singlet pairs is correct, then the experimenters would know exactly which pair belongs to which sub-ensemble. They could then scramble the sequence to yield an incorrect sequence. The incorrect sequence would not reproduce the Bell correlations. Similarly, if the initial unsorted sequence were known to be incorrect, then this would again yield statistics violating the Bell correlations. But for a superdeterministic model restricted to equilibrium, the measurement statistics cannot violate the Bell correlations. Thus only the final sorted sequence can be guaranteed to agree with observation. The experimenters do know, however, that if there are $x_{ij}$ runs for which they have chosen $\gamma = \gamma_{ij}$ (where $\sum_{i,j} x_{ij} = N_0$), then there have to be $x_{ij}$ pairs belonging to the sub-ensemble $E_{ij}$, for otherwise it would not be possible to reproduce the Bell correlations. Therefore, the number of possible initial sequences of singlet-state pairs is

$$W = \frac{N_0!}{\prod_{i,j} x_{ij}!}.$$  \hfill (10)

Since all the sequences are equally likely to the experimenters, the number $W$ may be associated with a formal mathematical entropy of

$$S = \ln\left(\frac{N_0!}{\prod_{i,j} x_{ij}!}\right).$$  \hfill (11)

This is nothing but the Shannon entropy of the distribution of singlet-state pairs (each corresponding to a specific sub-ensemble) over the $N_0$ runs. Intuitively, it may be thought as a measure of the variability of this distribution.

---

3 Recall that, in the correct sequence, if for a certain run the value of $\gamma$ is $\gamma_{ij}$, then the singlet pair for that run belongs to the sub-ensemble $E_{ij}$. 
FIG. 1: Illustration of spontaneous entropy drop at the hidden-variables level in superdeterministic models. Each solid ball represents a spin-singlet pair and its colour represents the sub-ensemble it belongs to. Each socket represents an experimental run and its colour represents the value of $\gamma$ for that run. Initially, as shown in the middle figure, there are a number of balls kept in a bag, arranged in no particular order, along with an equal number of sockets. The experimenters have to ensure that a spin-singlet pair belonging to a particular sub-ensemble, say $E_{ij}$, is measured only in a run where $\gamma = \gamma_{ij}$. Pictorially, this corresponds to matching the colours of both the balls and the sockets. However, the information ‘which spin-singlet pair belongs to which sub-ensemble’ is at the hidden-variables level, which the experimenters do not have access to. That is, the experimenters know the colour of the sockets but not the colour of the balls. Therefore, they perform the task by pure guesswork. One would expect that a situation like that shown in the left-hand figure would therefore arise. However they always find out that they have successfully matched each and every ball with its appropriate socket, as shown in the right-hand figure, because of some conspiratorial mechanism. Though logically possible, the mechanism causes a spontaneous entropy drop at the hidden-variables level. The entropy drop $\Delta S \to \infty$ linearly with the number of balls (spin-singlet pairs) and logarithmically with the number of colours (sub-ensembles).

After the two experimenters have arranged the pairs for performing the measurements, they can be confident that the matching has been successful (assuming quantum equilibrium). Since they know $\gamma$ for each run, and they also know that the matching has been successful, they now also know the value of $E_{ij}$ for each run. The number of possible sequences is now just $W = 1$, and so the associated entropy $S = \ln W = 0$. Thus the two experimenters have somehow caused the entropy $S$ to decrease by

$$\Delta S = \ln\left(\frac{N_0!}{\prod_{i,j} x_{ij}!}\right)$$

(12)

However, according to superdeterminism such a decrease always occurs (in equilibrium), regardless of whether or not the two experimenters deliberately attempt to produce it. Therefore, it is fair to say that this entropy decrease is spontaneous (see Fig. 1). Note the importance of assuming quantum equilibrium for this argument. In nonequilibrium, the Bell correlations will not be reproduced in general. The experimenters will interpret it as the matching being unsuccessful. The macroscopic information ‘matching was unsuccessful’ will not select any one sequence of pairs.

We now evaluate the net spontaneous entropy decrease for the simplest case that $x_{ij} = N_0/N^2 \forall i,j$ ($N_0/N^2$ is assumed to be a natural number). Assuming $N_0/N^2 \gg 1$ and using Stirling’s approximation$^4$ this is found to be

$$\Delta S = \ln(N_0!) - N^2 \ln[(N_0/N^2)!] = 2N_0 \ln(N)$$

(13)

Thus the entropy drop $\Delta S$ increases linearly with the number $N_0$ of spin-singlet pairs in the ensemble, and logarithmically with the number $N$ of setting mechanisms available at each wing of the experiment. In the limit $N \to \infty$ (with the constraint that $N_0/N^2 = x_{ij}$ is a natural number), we have $\Delta S \to \infty$ and in this sense we may say that the superdeterministic model becomes arbitrarily conspiratorial.

$^4$ For large $n$, $\ln(n!) = n \ln n - n$. 
This way of quantifying the conspiracy does not assume the possibility of nonequilibrium distributions, or that \( \lambda \) is discrete. It only quantifies the counter-intuitiveness of making the correct choice spontaneously at all times.

V. CONCLUSION

We have provided two different ways to quantify the often-alleged conspiratorial character of superdeterministic models, using intuitive suggestions made from a previous article \( \mathcal{A} \). The first approach uses the idea that different setting mechanisms lead to different measurement statistics for a nonequilibrium superdeterministic model. We defined a parameter called the overhead fine-tuning \( F \), which quantifies the amount by which an arbitrary distribution of the model must be fine-tuned in order to reproduce two basic empirical observations: the measurement statistics do not depend on \( a \) the outputs of the unused setting mechanisms, and \( b \) the experimenters’ choice of which setting mechanisms to use. We then proved that, for a superdeterministic model of our setup, \( F \to 1 \) exponentially fast (corresponding to a ‘completely fine-tuned model’) as the number of setting mechanisms \( N \to \infty \). No features of quantum statistics, such as the Born rule, marginal-independence (formal no-signalling), or non-contextuality were assumed in deriving this result.

Intuitively, \( F \) quantifies how special the hidden-variables distribution has to be in order to reproduce the properties \( a \) and \( b \). Clearly, the notion of ‘special’ is meaningful only when there are multiple possible distributions, that is, if nonequilibrium is allowed (at least in principle). Those who consider nonequilibrium distributions to be unmotivated may find this approach to be artificial. We have, therefore, provided a second way to quantify the conspiracy, which does not employ the concept of nonequilibrium.

The second approach uses another idea from \( \mathcal{A} \) that a superdeterministic model in nonequilibrium can mimic a physical signal using a series of coincidences set-up by initial conditions. We argued that such coincidences are in fact present even in equilibrium for a superdeterministic model, but at the hidden-variables level. We started off with a superdeterministic model of our setup restricted to equilibrium, and then argued that, for the model to be consistent, the experimenters’ choices have to be determined from the past in such a manner that they coincidentally always pick the ‘right combinations’ of singlet-state pairs and setting mechanisms. We have shown that this leads to a spontaneous drop \( \Delta S \) of a formal entropy \( S \) at the hidden-variables level, and that \( \Delta S \to \infty \) logarithmically fast with \( N \). By quantifying the conspiracy in this way, as a spontaneous drop of entropy, we may say that the equilibrium superdeterministic model becomes arbitrarily conspiratorial as \( N \to \infty \).

Are other hidden-variables models of our setup conspiratorial according to our two approaches? For nonlocal models, the overhead fine-tuning \( F = 0 \) by construction, since the hidden-variables distribution does not depend upon the variables \((\{\alpha\}, \{\beta\}, \gamma)\). There are also no sub-ensembles associated with different setting mechanisms, and so the argument leading to a spontaneous entropy drop does not apply. Therefore nonlocal models of our setup are not conspiratorial. On the other hand, the situation is not immediately clear for retrocausal models. This is primarily because, up till now, no clear retrocausal mechanism has been proposed which explains exactly how the information about the future measurement settings is made available to the preparation event in the past. These issues shall be discussed in a forthcoming paper \( \mathcal{B} \).

We conclude that superdeterministic models of Bell correlations are arbitrarily conspiratorial. On this basis it seems fair to conclude that superdeterminism is scientifically unattractive.

VI. APPENDIX

We provide here a proof of principle that our result for arbitrarily large fine-tuning holds for a discrete approximation to a continuous distribution, where the approximation can be made as close as desired to the continuous case.

Let there be a superdeterministic model with a continuous distribution \( \rho(\lambda|\{\alpha\}, \{\beta\}, \gamma) \) possessing a first derivative at all points, where \( \lambda \in (\infty, +\infty) \). In a discrete approximation of this model, we replace \( \rho(\lambda|\{\alpha\}, \{\beta\}, \gamma) \) by \( p(\lambda|\{\alpha\}, \{\beta\}, \gamma) \), where \( \lambda \) is restricted to points in the set \( S = \{-\lambda_o, -\lambda_o + \epsilon, -\lambda_o + 2\epsilon, \ldots, -\lambda_o + (\Lambda - 1)\epsilon\} \). Here \( \lambda_o \)
can be made as large as desired. In the limit \( \epsilon \to 0 \) and \( \Lambda \to \infty \), we impose the condition
\[
\lim_{\epsilon \to 0} \epsilon \Lambda = 2\lambda_o
\]  
(14)

We define \( p(-\lambda_o|\{\alpha\},\{\beta\},\gamma) = \rho(-\lambda_o|\{\alpha\},\{\beta\},\gamma)\epsilon \). For the points contained in the set \( S \setminus \{-\lambda_o\} \), we define
\[
p(-\lambda_o + (l+1)\epsilon|\{\alpha\},\{\beta\},\gamma) = p(-\lambda_o + l\epsilon|\{\alpha\},\{\beta\},\gamma) + \rho'(-\lambda_o + l\epsilon|\{\alpha\},\{\beta\},\gamma)\epsilon^2
\]  
(15)

where \( l \in \{0,1,2,3,...,(\Lambda-2)\} \). Therefore, the discrete approximation to the continuous distribution \( \rho(\lambda|\{\alpha\},\{\beta\},\gamma) \) has \( \Lambda \) independent parameters corresponding to \( \rho(-\lambda_o|\{\alpha\},\{\beta\},\gamma), \rho'(-\lambda_o|\{\alpha\},\{\beta\},\gamma), \rho'(-\lambda_o + \epsilon|\{\alpha\},\{\beta\},\gamma), \rho'(-\lambda_o + 2\epsilon|\{\alpha\},\{\beta\},\gamma),..., \rho'(-\lambda_o + (\Lambda-2)\epsilon|\{\alpha\},\{\beta\},\gamma) \). The normalisation constraint on the discrete distribution
\[
\sum_{l=0}^{\Lambda-1} p(-\lambda_o + l\epsilon|\{\alpha\},\{\beta\},\gamma) = 1
\]  
(16)

reduces the parameters by 1, that is, to \( \Lambda - 1 \), which is the exact same number of parameters for an arbitrary discrete distribution considered in the main text. Since each of \( \{\alpha\}, \{\beta\} \) and \( \gamma \) is already discrete, we have the same number of distributions as in the main text. Lastly, as there are \( \Lambda \) points in the set \( S \), the total number of constraints \( \kappa \) implied by equations (1) and (3) are also identical to that in the main text. Therefore, applying the definition of fine-tuning given by equation (6) to the discrete approximation to the continuous case leads us to the same result, that \( F \to 1 \) exponentially fast as \( N \to \infty \). The proof can be easily extended to piecewise continuously differentiable distributions.

[1] I. Sen and A. Valentini. Superdeterministic hidden-variables models I: nonequilibrium and signalling. arXiv preprint arXiv: 2003.11989.
[2] J. S. Bell. Speakable and unspeakable in quantum mechanics: Collected papers on quantum philosophy. Cambridge Univ. Press, 2004.
[3] L. Hardy. Quantum ontological excess baggage. Stud. Hist. Phil. Sci. B, 35, 2004.
[4] A. Valentini. Signal-locality in hidden-variables theories. Phys. Lett. A, 297(5-6), 2002.
[5] A. Valentini. Universal signature of non-quantum systems. Phys. Lett. A, 332(3), 2004.
[6] A. Valentini. Astrophysical and cosmological tests of quantum theory. J. Phys. A, 40(12), 2007.
[7] G. Carcassi, C. A. Aidala, and J. Barbour. Variability as a better characterization of Shannon entropy. arXiv preprint arXiv:1912.02012, 2019.
[8] I. Sen. The physical interpretation of retrocausal hidden-variable models. 2003.11989.