

Leggett-Garg inequalities for quantum fluctuating work

H. J. D. Miller^{1,*} and J. Anders^{1,†}

¹*Department of Physics and Astronomy, University of Exeter, Stocker Road, EX4 4QL, United Kingdom.*

(Dated: February 28, 2018)

The Leggett-Garg inequalities serve to test whether or not quantum correlations in time can be explained within a classical macrorealistic framework. We apply this test to thermodynamics and derive a set of Leggett-Garg inequalities for the statistics of fluctuating work done on a quantum system unitarily driven in time. It is shown that these inequalities can be violated in a driven two-level system, thereby demonstrating that there exists no general macrorealistic description of quantum work. These violations are shown to emerge within the standard Two-Projective-Measurement scheme as well as for alternative definitions of fluctuating work that are based on weak measurement. Our results elucidate the influences of temporal correlations on work extraction in the quantum regime and highlight a key difference between quantum and classical thermodynamics.

I. INTRODUCTION

Much like the celebrated Bell inequalities, which shed light on the deeply non-classical properties of spatial correlations encountered in entangled systems, quantum mechanics possesses a rich temporal structure that distinguishes it from classical physics. In 1985 Leggett and Garg explored this structure by introducing the concept of *macrorealism* [1]. In essence macrorealism can be condensed into two main assumptions about the temporal properties of physical observables within any *classical* description of physics [1–8];

- (i) *Macrorealism per se*: physical observables take on well-defined values at all times independent of the act of observation.
- (ii) *Non-invasive measurability*: in principle it is possible to measure the value of an observable without changing the subsequent evolution of the system.

Assumptions (i) and (ii) can be used to derive mathematical inequalities, the so-called Leggett-Garg inequalities, that serve to test the macrorealism of physical observables. Violations of these inequalities subsequently rule out what would be expected in a classical system, and this quantum behaviour has now been confirmed experimentally in a variety of settings [9–14].

In the quantum regime, thermodynamic quantities such as fluctuating work and heat cannot be represented by hermitian observables, but are conventionally defined via multi-time projective measurements performed on the system [15–17]. For a closed quantum system, one way of defining the fluctuating work done on the system driven out of equilibrium is by the difference in energy eigenvalues observed at the start and end of its evolution. This framework is commonly referred to as the two-projective-measurement scheme, and serves as a route to many of the known fluctuation theorems such as the Jarzynski equality [15, 16] and Tasaki-Crooks relation [18]. Given that these results mirror the corresponding classical fluctuation relations [19, 20], it is often assumed that work

is simply a classical stochastic variable even within the quantum regime. However, the influence of non-classical temporal correlations that arise from two-time quantum measurements on the statistics of fluctuating work have yet to be fully understood. One aspect of this is the fact that work measurements remove coherences in the energy basis and can affect the future evolution of the quantum system, modifying the average work done during the process [21, 22]. Alternative definitions of quantum work related to weak measurement have been proposed in order to circumvent this effect of measurement disturbance on the statistics of work [22–24]. However, it has been shown that the resulting quantum work distributions are not generally positive-definite [25]. The emergence of negative quasi-probabilities is a signature of quantum behaviour, and hints at a link to violations of the Leggett-Garg inequalities [6, 26, 27]. In a similar vein, violations of macrorealism have also been related to the presence of anomalous weak values in quantum systems [8]. Recent work by Blattmann and Mølmer [28] has successfully linked violations of macrorealism to quantum work in the standard TPM approach by utilising the entropic Leggett-Garg inequalities. In their approach one compares the Shannon entropy of the work distribution over different intervals of time. However, the Shannon entropy is not well-defined if the work distribution fails to be positive, and so the entropic Leggett-Garg inequalities cannot be applied to situations in which the work distribution becomes a quasi-probability.

In this paper we will utilise the assumptions of Leggett and Garg to demonstrate that there exists no general macrorealistic description of work for quantum systems driven out of equilibrium. In particular, we show that quantum temporal correlations between energy measurements performed at different times influence the statistical moments of the fluctuating work done on the system during a non-equilibrium process. This result is shown to hold for three different definitions of quantum work: the two-projective measurement (TPM) scheme [15], the full-counting statistics (FCS) [24] and the Margenau-Hill (MH) work distribution [23]. Crucially the inequalities that we derive can be used to test for violations of macrorealism in both strong and weak measurement schemes, regardless of whether or not the work distribution is positive or not.

The paper is organised as follows: we first introduce a set of Leggett-Garg inequalities for the moments of fluctuating

* hm419@exeter.ac.uk

† janet@qipc.org

work, and then consider a driven two-level system and show that the inequalities can be violated. Following that we introduce an alternative set of Leggett-Garg inequalities for the moment-generating function, and apply these inequalities to alternative definitions of quantum work that are based on weak measurement, namely the FCS and MH definitions, subsequently showing that violations of macrorealism can also occur. Finally we conclude with a discussion of our results.

II. INEQUALITIES FOR MOMENTS OF WORK

We first recall the setup for the original Leggett-Garg inequalities [1]. First consider performing three protocols in which the spin $S(t_i) = S_i = \pm 1$ of a qubit is projectively measured at two times within a set of three times $t_0 < t_1 < t_2$. For each of the three protocols one can obtain the temporal correlation function for the values of the spin at times $t = t_i$ and $t = t_j$, denoted $C_{ij} = \langle S_i S_j \rangle$. The macrorealism assumptions (i) and (ii) imply that there exists a three-time probability distribution $\mathcal{P}(S_0, S_1, S_2)$ such that the distributions describing the statistics of each individual protocol can be obtained as marginals of this three-time distribution, eg. $\mathcal{P}(S_0, S_2) = \sum_{S_1} \mathcal{P}(S_0, S_1, S_2)$ and so on [4]. Note that while assumption (i) implies the existence of a three-time probability with the correct marginals, assumption (ii) guarantees that this distribution is the same for all three separate experiments [4]. Finally, using the marginal properties of $\mathcal{P}(S_0, S_1, S_2)$ yields the following Leggett-Garg inequality relating the correlation functions for the three protocols [1];

$$C_{01} + C_{12} - C_{02} \leq 1. \quad (1)$$

This inequality holds for any dichotomic observable. For a simple qubit, the spin at time t_i can be represented by a combination of Pauli matrices; $\hat{S}_i = \mathbf{s}_i \cdot \hat{\boldsymbol{\sigma}}$. By performing successive projective measurements of the spin, the correlation functions can be obtained for each of the three protocols. This leads to a violation the RHS of Eq. (1) which can take a maximum value of $\frac{3}{2}$ [4]. This example illustrates the failure of macrorealism for quantum systems.

It is also possible to derive a set of Leggett-Garg inequalities reminiscent of Eq. (1) for the moments of fluctuating work in a closed quantum system driven in time. For simplicity we consider a system that can occupy one of two fixed energy states, which we denote by $\frac{\epsilon}{2}$ and $-\frac{\epsilon}{2}$, at three points in time $t = t_0 < t_1 < t_2$ during the driving process. For the statistics of work measured during a particular time interval $t \in [t_i, t_j]$, the k 'th moment of fluctuating work is defined as

$$\langle W^k(t_i, t_j) \rangle = \sum_{\epsilon_i, \epsilon_j} \mathcal{P}(\epsilon_i, \epsilon_j) (\epsilon_j - \epsilon_i)^k. \quad (2)$$

Here the energies occupied by the system at time t_i are denoted by ϵ_i and we assume that the possible work values are given by the energy changes $W_{ij} = \epsilon_j - \epsilon_i$. The probability $\mathcal{P}(\epsilon_i, \epsilon_j)$ governs the statistics of energy at times t_i and t_j , and we make no assumptions about the exact definition of $\mathcal{P}(\epsilon_i, \epsilon_j)$ aside

from assuming it is normalised and non-negative. As with the standard Leggett-Garg experiment described above, the aim is to compare the work statistics observed within different time intervals along the driving process, as shown in Figure 1. Thus in analogy with Eq. (1) we will consider the following quantity;

$$M_k = \langle W^k(t_0, t_1) \rangle + \langle W^k(t_1, t_2) \rangle - \langle W^k(t_0, t_2) \rangle. \quad (3)$$

This quantity can be measured over many runs of the driving process during each of the three time intervals, where the system is prepared in same state at time t_0 for each experiment. For example M_1 is obtained by measuring the average work done on the system sequentially during intervals $t \in [t_0, t_1]$ and $t \in [t_1, t_2]$, and then subtracting the average work done during the total time interval $t \in [t_0, t_2]$. We now seek to bound Eq. (3) through the assumption that the fluctuating work is a macrorealistic variable, as defined by (i) and (ii).

These assumptions imply the existence of a global probability distribution $\mathcal{P}(\epsilon_0, \epsilon_1, \epsilon_2)$ describing the energy statistics at all points in time along the driving process, where each two-time distribution can be obtained as a marginal;

$$\mathcal{P}(\epsilon_i, \epsilon_j) = \sum_{m \neq i, j} \mathcal{P}(\epsilon_0, \epsilon_1, \epsilon_2) \quad \forall i, j. \quad (4)$$

We can show that this condition immediately leads to the following Leggett-Garg inequality for the moments of work (see Appendix A);

$$\begin{aligned} M_k &\geq 0, \quad \text{even } k, \\ M_k &= 0, \quad \text{odd } k. \end{aligned} \quad (5)$$

For $k = 1$ this implies $\langle W(t_0, t_1) \rangle + \langle W(t_1, t_2) \rangle = \langle W(t_0, t_2) \rangle$. This makes intuitive sense; in classical thermodynamics one would not expect to observe any difference between the sum of each intermediate average amount work done and the total average work done between the initial and final points in time. However, we will subsequently show that this does not generally hold for quantum systems, as the bounds in Eq. (5) can be violated for certain driving processes. It should also be noted that while we have assumed a discrete energy spectrum for the time-dependent Hamiltonian, this is not crucial to the derivation of Eq. (5). Indeed, so long as one assumes that the energy moments are always finite then inequalities of the form Eq. (5) can be derived. However, we will restrict our attention to a two-dimensional quantum system throughout the paper for simplicity, as this is sufficient to demonstrate violations of macrorealism in the statistics of fluctuating work.

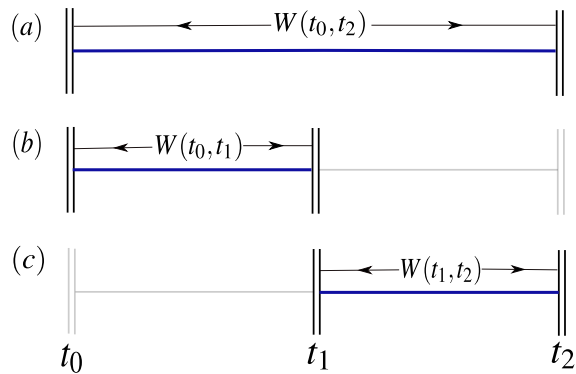


FIG. 1. **Schematic diagram for detecting non-classical work statistics.** An observer performs three separate experiments, (a)-(c), in which the fluctuating work done on the system is measured between the time intervals shown in the diagram. To test the validity of the Leggett-Garg inequality for work, Eq. (5), one compares the statistics of the three experiments, with the same initial state chosen at time t_0 . Note that in experiment (c) no measurement is made at t_0 , thus the system evolves unitarily up to time t_1 .

III. VIOLATIONS OF THE LEGGETT-GARG INEQUALITIES FOR WORK MOMENTS

We will now utilise the inequalities Eq. (5) to show that fluctuating work can not generally be described by a macrorealist theory for quantum systems. The relevant situation that we consider is a standard setup for the thermodynamics of work extraction; an isolated system is initially thermalised and then driven out of equilibrium via changing its Hamiltonian in time, extracting work in the process [15–17]. Suppose that we have a two-level system described by a time-dependent Hamiltonian $\hat{H}_H(t)$ in the Heisenberg picture with initial state $\hat{\rho}$ such that;

$$\hat{H}_H(t_i) = \frac{\varepsilon}{2} \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}; \quad \hat{\rho} = \frac{1}{2}(\hat{I} + \mathbf{r} \cdot \hat{\boldsymbol{\sigma}}), \quad (6)$$

where $|\mathbf{r}| \leq 1$ and $|\mathbf{a}_i| = 1$ are vectors. Note that the time-dependence of the Hamiltonian is attributed only to each vector \mathbf{a}_i . In an experimental setup Eq. (6) describes a spin-1/2 particle coupled to an external classical magnetic field, with the direction of the applied field adjusted by the experimenter in time. Without loss of generality we will set the initial Hamiltonian along the z-axis of the Bloch sphere, i.e. $\mathbf{a}_0 = \{0, 0, 1\}$, and choose an initial thermal state with respect to $\hat{H}(t_0)$ at inverse temperature β ; $\hat{\rho} \propto \exp(-\beta \hat{H}(t_0))$. This in turn implies that $\mathbf{r} = \{0, 0, -\tanh(\beta \varepsilon/2)\}$. To obtain the moments of work in Eq. (3) two projective energy measurements are performed at the start and end of the driving process within the fixed time intervals shown in Figure 1. This method is commonly referred to as the two-projective measurement scheme (TPM) [15, 16, 29]. The joint probability to observe energy ε_i at $t = t_i$ and then ε_j at $t = t_j$ is given by

$$\mathcal{P}\left(\varepsilon_i = \pm \frac{\varepsilon}{2}, \varepsilon_j = \pm \frac{\varepsilon}{2}\right) = \text{Tr}[\hat{\rho} \hat{P}_{a_i}^{\pm}] \cdot \text{Tr}[\hat{P}_{a_j}^{\pm} \hat{P}_{a_i}^{\pm}]. \quad (7)$$

Here we have denoted $\hat{P}_{a_i}^{\pm}$ as the projector onto the relevant energy state of the Hamiltonian at time t_i . Substituting Eq. (7) into Eq. (2) gives the moments of work from successive projective energy measurements (see Appendix B):

$$\langle W^k(t_i, t_j) \rangle = \begin{cases} (1 - \mathbf{a}_j \cdot \mathbf{a}_i) \frac{\varepsilon^k}{2}, & \text{even } k, \\ -(\mathbf{r} \cdot \mathbf{a}_i)(1 - \mathbf{a}_j \cdot \mathbf{a}_i) \frac{\varepsilon^k}{2}, & \text{odd } k. \end{cases} \quad (8)$$

To identify the conditions under which fluctuations in work violate Eq. (5), we substitute the above expression into Eq. (3) and parameterise the driving process by introducing $\cos(\theta_{ij}) = \mathbf{a}_j \cdot \mathbf{a}_i$. For even k we have the following quantum bound for M_k after minimising over all normalised vectors $\{\mathbf{a}_i\}$ for $i = 0, 1, 2$ (see Appendix B);

$$\min_{\{\mathbf{a}_i\}} \left[M_k \right] = -\frac{\varepsilon^k}{4}; \quad (\text{even } k). \quad (9)$$

which is saturated by choosing $\theta_{10} = \theta_{21} = \pi/3$. Secondly, for odd k we find that maximising over $\{\mathbf{a}_i\}$ gives

$$\max_{\{\mathbf{a}_i\}} \left[|M_k| \right] = \left| \frac{\varepsilon^k}{2} \tanh(\beta \varepsilon/2) \right|; \quad (\text{odd } k), \quad (10)$$

where the bound is saturated by choosing $\theta_{10} = \theta_{21} = \pi/2$. From Eq. (9) and Eq. (10) it is now apparent that there exists unitary protocols that violate the bounds in Eq. (5). Given that the bounds Eq. (5) necessarily follow from assumptions (i) and (ii), we conclude that quantum fluctuating work generally lacks a macrorealistic description.

Let us note that Eq. (5) cannot be violated for odd k in the high temperature limit. However, Eq. (8) shows that the even work moments are independent of β , and so Eq. (5) can indeed be violated for even k regardless of temperature. This is not surprising, as the standard Leggett-Garg inequality can be violated in the case of an initially maximally mixed qubit [4]. Violations of the work Leggett-Garg inequality, Eq. (5), can still occur at all temperatures since the system may acquire coherences in energy at intermediate times due to the unitary driving.

IV. INEQUALITIES FOR THE CHARACTERISTIC FUNCTION OF WORK

While the inequalities Eq. (5) provide a simple identification of non-classicality for the moments of work, we will show in this section that it is possible to condense this information into two inequalities related to the characteristic function for work rather than the moments themselves. A similar method has previously been applied to investigate the non-classical properties of electron-transport through conductors [30]. The benefit of this approach is two-fold. While the projective energy measurements used to obtain the moments in Eq. (2) may be difficult to implement in practice, measurements of the characteristic function for work can be performed via ancilla-assisted measurement as shown in [31–34]. Secondly, inequalities for the characteristic function allow us to

consider alternative non-invasive measurement schemes such as the full-counting statistics approach to the quantum work distribution proposed in [24], as we later show in the next section.

The characteristic function uniquely defines the probability distribution for work in a unitarily driven system, and is obtained through the Fourier transform of the work distribution;

$$G_\lambda(t_i, t_j) = \sum_{\varepsilon_i, \varepsilon_j} \mathcal{P}(\varepsilon_i, \varepsilon_j) e^{i\lambda(\varepsilon_j - \varepsilon_i)}, \quad (11)$$

with work values $\varepsilon_j - \varepsilon_i$. Assuming the same protocol given by Eq. (6), we can consider a linear combination of characteristic functions for the three intervals of time shown in Figure 1;

$$L_\lambda = G_\lambda(t_0, t_1) + G_\lambda(t_1, t_2) - G_\lambda(t_0, t_2). \quad (12)$$

The assumptions (i) and (ii) for macrorealism imply the following upper bound on the real part of Eq. (12) for the qubit system (see Appendix C);

$$\text{Re}(L_\lambda) \leq 1, \quad (13)$$

whilst the imaginary part of Eq. (12) becomes an *equality*;

$$\text{Im}(L_\lambda) = 0. \quad (14)$$

The bounds Eq. (13) and Eq. (14) now constitute a pair of Leggett-Garg-type inequalities for any classical characteristic function for fluctuating work, assuming that the energy is given by either $\frac{\varepsilon}{2}$ or $-\frac{\varepsilon}{2}$ at all times along the driving protocol.

Within the TPM scheme the characteristic function for work is given by [15, 16];

$$G_\lambda^{\text{TPM}}(t_i, t_j) = \text{Tr}[\hat{\eta}_i e^{i\lambda\hat{H}_H(t_j)} e^{-i\lambda\hat{H}_H(t_i)}], \quad (15)$$

where $\hat{\eta}_i$ is the initial state ρ decohered in the basis of $\hat{H}(t_i)$. The quantum upper bound for the real part of Eq. (12) is as follows (see Appendix D);

$$\max_{\{a_i\}} \left[\text{Re}(L_\lambda) \right] = \frac{5}{4} - \frac{1}{4} \cos(\lambda\varepsilon), \quad (16)$$

Consequently the upper bound exceeds the classical inequality Eq. (13) for all λ . Secondly, the upper bound for the imaginary part of Eq. (12) is

$$\max_{\{a_i\}} \left[\left| \text{Im}(L_\lambda) \right| \right] = \left| \frac{1}{2} \tanh(\beta\varepsilon/2) \sin(\lambda\varepsilon) \right|, \quad (17)$$

with a maximum violation of $\text{Im}(L_\lambda^G) = \frac{1}{2} \tanh(\beta\varepsilon/2)$ for $\lambda = \pi/2\varepsilon$. As one would expect, the quantum bounds Eq. (16) and Eq. (17) are obtained by choosing the same protocol-dependent parameters used to obtain Eq. (9) and Eq. (10) respectively.

To summarise this section we have presented a Leggett-Garg inequality for the characteristic function of fluctuating work, and shown that the non-classicality observed in the moments of work is also exhibited in the characteristic function itself for the TPM scheme.

V. GENERALISATION TO WEAK MEASUREMENTS OF WORK

In the previous section we simply recast the original violations of the Leggett-Garg inequality Eq. (5) into the form relevant to the characteristic function for the work statistics of a qubit. While this was applied to the TPM protocol, the inequalities Eq. (13) and Eq. (14) apply to any qubit with a fixed energy spectrum, and this ultimately allows us to investigate alternative measurement schemes and their resulting non-classical violations. Due to the invasive nature of projective measurements, some have argued that the definition of fluctuating work is thermodynamically inconsistent when applied to states with initial coherences in energy [21, 23–25]. In turn this has inspired formulations of non-invasive work statistics that remain consistent with energy conservation in closed systems [22–24]. In particular, we will consider the full-counting statistics for fluctuating work [24, 26, 30, 35–37]. To obtain the full-counting statistics for work, one couples the system's Hamiltonian to the momentum of an external detector, and subsequently measures the phase change acquired by the detector's momentum during the driving process given by the system's time-dependent Hamiltonian $H(t)$ [36]. In turn this allows one to reconstruct a characteristic function of the following form [24, 38];

$$G_\lambda^{\text{FCS}}(t_i, t_j) = \text{Tr}[\hat{\rho} e^{-i\frac{\lambda}{2}\hat{H}_H(t_i)} e^{i\lambda\hat{H}_H(t_j)} e^{-i\frac{\lambda}{2}\hat{H}_H(t_i)}]. \quad (18)$$

For example, as shown in [24] the first moment of Eq. (18) yields $\langle W(t_i, t_j) \rangle = \langle H(t_j) \rangle - \langle H(t_i) \rangle$; in other words the average work done on the system is simply the difference in average energy evaluated in the Heisenberg picture. Consequently one can easily see that the first moment in this measurement scheme still obeys the Leggett-Garg inequality Eq. (5). However, we demonstrate that the same is not true for the higher order moments contained in the characteristic function Eq. (18). Considering again the quantum two-level system described by Eq. (6), it can be shown that the real part of Eq. (18) is equal to the real part of Eq. (15) obtained via the TPM scheme (see Appendix E);

$$\text{Re}\left(G_\lambda^{\text{FCS}}(t_i, t_j)\right) = \text{Re}\left(\text{Tr}[\hat{\eta}_i e^{i\lambda\hat{H}_H(t_j)} e^{-i\lambda\hat{H}_H(t_i)}]\right), \quad (19)$$

This is a surprising result, as it suggests that the same violations of Eq. (13), ie. the upper bound Eq. (16), can be obtained non-invasively. It is only the imaginary part of $G_\lambda^{\text{FCS}}(t_i, t_j)$ that differs from the TPM scheme, in which we find the following upper bound for Eq. (12);

$$\max_{\{a_i\}} \left[\left| \text{Im}(L_\lambda^{\text{FCS}}) \right| \right] = \left| 2\sin^2\left(\frac{\lambda\varepsilon}{4}\right) \sin\left(\frac{\lambda\varepsilon}{2}\right) \tanh\left(\frac{\beta\varepsilon}{2}\right) \right| \quad (20)$$

which is again obtained by choosing $\theta_{10} = \theta_{21} = \pi/2$.

The full-counting statistics are not the only way to characterise non-invasive measurements of work. As proposed by Allahverdyan [23], an alternative characteristic function describing the statistics of work derived from the Margenau-Hill distribution for successive energy measurements [39, 40] is as

follows (see Appendix F):

$$G_{\lambda}^{\text{MH}}(t_i, t_j) = \text{Tr}[\hat{\rho} e^{i\lambda\hat{H}_H(t_j)} \star e^{-i\lambda\hat{H}_H(t_i)}], \quad (21)$$

where $\hat{A} \star \hat{B} = \frac{1}{2}[\hat{A}\hat{B} + \hat{B}\hat{A}]$ denotes the symmetric Jordan product. Notably the corresponding probability distribution can be obtained via sequential weak measurement [41]. While the first and second moments, $\langle W \rangle$ and $\langle W^2 \rangle$, are the same as those obtained from the full-counting statistics, in general higher order moments differ. However, for the isolated driven qubit we again find precisely the same violations of Eq. (13) because the real part of $G_{\lambda}^{\text{MH}}(t_i, t_j)$ is also equivalent to the real part of $G_{\lambda}^{\text{FCS}}(t_i, t_j)$. On the other hand the inequality for the imaginary term, Eq. (14), cannot be violated in our setup (Appendix F).

VI. DISCUSSION

In the paper we have demonstrated a violation of macro-realism in the statistics of fluctuating work for a quantum system unitarily driven out of thermal equilibrium for three different characterisations of the work statistics. As with the original Leggett-Garg inequalities, these violations stem from the absence of a global three-time probability distribution of the form Eq. (4) for both strong and weak measurement

schemes. This emphasises the fact that quantum fluctuations in work are manifestly different from the stochastic fluctuations encountered in classical non-equilibrium thermodynamics due to the influence of temporal correlations on the work moments. These findings compliment recent results showing that fluctuating work cannot always be assigned a well-defined probability distribution when quantum coherence is taken into account [22–25, 42, 43]. Ultimately our analysis shows that the Leggett-Garg inequalities provide a useful tool for understanding the difference between quantum and classical thermodynamics, and the inequalities Eq. (5), Eq. (13) and Eq. (14) may find an application in identifying quantum behaviour in thermal machines [44].

ACKNOWLEDGMENTS

We would like to thank Paolo Solinas for suggesting a link between quantum work and the Leggett-Garg inequality, and Clive Emary and Alexander Friedenberger for insightful discussions. HM is supported by EPSRC through a Doctoral Training Grant. J.A. acknowledges support from EPSRC, grant EP/M009165/1, and the Royal Society. This research was supported by the COST network MP1209 ‘‘Thermodynamics in the quantum regime’’.

-
- [1] A. Leggett and A. Garg, *Phys. Rev. Lett.* **54**, 857 (1985).
 - [2] J. Kofler and Č. Brukner, *Phys. Rev. Lett.* **101**, 090403 (2008).
 - [3] M. M. Wilde and A. Mizel, *Found. Phys.* **42**, 256 (2012).
 - [4] C. Emary, N. Lambert, and F. Nori, *Rep. Prog. Phys.* **77**, 016001 (2014).
 - [5] J. Kofler and Č. Brukner, *Phys. Rev. A* **87**, 052115 (2013).
 - [6] J. J. Halliwell, *Phys. Rev. A* **93**, 022123 (2016).
 - [7] L. Clemente and J. Kofler, *Phys. Rev. A* **91**, 062103 (2015).
 - [8] N. S. Williams and A. N. Jordan, *Phys. Rev. Lett.* **100**, 026804 (2008).
 - [9] A. Palacios-Laloy, F. Mallet, F. Nguyen, P. Bertet, D. Vion, D. Esteve, and A. N. Korotkov, *Nat. Phys.* **6**, 442 (2010).
 - [10] J. Dressel, C. J. Broadbent, J. C. Howell, and A. N. Jordan, *Phys. Rev. Lett.* **106**, 040402 (2011).
 - [11] M. E. Goggin, M. P. Almeida, M. Barbieri, B. P. Lanyon, J. L. O’Brien, A. G. White, and G. J. Pryde, *Proc. Natl. Acad. Sci. USA* **108**, 1256 (2011).
 - [12] G. C. Knee, S. Simmons, E. M. Gauger, J. J. L. Morton, H. Riemann, N. V. Abrosimov, P. Becker, H.-J. Pohl, K. M. Itoh, M. L. W. Thewalt, G. A. D. Briggs, and S. C. Benjamin, *Nat. Comm.* **3**, 606 (2012).
 - [13] Z. Q. Zhou, S. F. Huelga, C. F. Li, and G. C. Guo, *Phys. Rev. Lett.* **115**, 113002 (2015).
 - [14] C. Robens, W. Alt, D. Meschede, C. Emary, and A. Alberti, *Phys. Rev. X* **5**, 011003 (2015).
 - [15] P. Talkner, E. Lutz, and P. Hänggi, *Phys. Rev. E* **75**, 050102 (2007).
 - [16] M. Esposito, U. Harbola, and S. Mukamel, *Rev. Mod. Phys.* **81**, 1665 (2009).
 - [17] P. Hänggi and P. Talkner, *Nat. Phys.* **11**, 108 (2015).
 - [18] P. Talkner and P. Hänggi, *J. Phys. A* **40**, F569 (2007).
 - [19] C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
 - [20] G. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
 - [21] P. Kammerlander and J. Anders, *Sci. Rep.* **6**, 22174 (2016).
 - [22] H. J. D. Miller and J. Anders, *N. J. Phys.* **19**, 062001 (2017).
 - [23] A. E. Allahverdyan, *Phys. Rev. E* **90**, 032137 (2014).
 - [24] P. Solinas and S. Gasparinetti, *Phys. Rev. E* **92**, 042150 (2015).
 - [25] M. Perarnau-Llobet, E. Baumer, K. V. Hovhannissyan, M. Huber, and A. Acin, *Phys. Rev. Lett.* **118**, 070601 (2017).
 - [26] A. Bednorz, W. Belzig, and A. Nitzan, *N. J. Phys.* **14**, 013009 (2012).
 - [27] P. P. Hofer, *Quantum* **1**, 32 (2017).
 - [28] R. Blattmann and K. Mølmer, *Phys. Rev. A* **96**, 012115 (2017).
 - [29] C. Jarzynski, H. T. Quan, and S. Rahav, *Phys. Rev. X* **5**, 031038 (2015).
 - [30] C. Emary, *Phys. Rev. B* **86**, 085418 (2012).
 - [31] R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, and V. Vedral, *Phys. Rev. Lett.* **110**, 230601 (2013).
 - [32] L. Mazzola, G. De Chiara, and M. Paternostro, *Phys. Rev. Lett.* **110**, 230602 (2013).
 - [33] A. J. Roncaglia, F. Cerisola, and J. P. Paz, *Phys. Rev. Lett.* **113**, 250601 (2014).
 - [34] T. B. Batalhão, A. M. Souza, L. Mazzola, R. Auccaise, R. S. Sarthour, I. S. Oliveira, J. Goold, G. De Chiara, M. Paternostro, and R. M. Serra, *Phys. Rev. Lett.* **113**, 140601 (2014).
 - [35] Y. V. Nazarov and M. Kindermann, *Euro. Phys. J. B* **35**, 413 (2003).
 - [36] A. A. Clerk, *Phys. Rev. A* **84**, 043824 (2011).

- [37] P. P. Hofer and A. A. Clerk, *Phys. Rev. Lett.* **116**, 013603 (2016).
 [38] P. Solinas and S. Gasparinetti, *Phys. Rev. A* **94**, 052103 (2016).
 [39] H. Margenau and R. N. Hill, *Prog. Theo. Phys.* **26**, 722 (1961).
 [40] L. M. Johansen, *Phys. Rev. A* **76**, 012119 (2007).
 [41] J. S. Lundeen and C. Bamber, *Phys. Rev. Lett.* **108**, 070402 (2012).
 [42] P. Solinas, H. J. D. Miller, and J. Anders, *Phys. Rev. A* **96**, 052115 (2017).
 [43] M. Lostaglio, [arXiv:1705.05397](https://arxiv.org/abs/1705.05397) (2017).
 [44] A. Friedenberger and E. Lutz, [arXiv:1508.04128](https://arxiv.org/abs/1508.04128) (2015).

Appendix A: Derivation of Eq. (5)

By using the condition Eq. (4) for the marginal probabilities describing the statistics of the three experiments displayed in Figure (1), one can express the following linear combination of work, Eq. (3) as;

$$\begin{aligned} M_k &= \sum_{\varepsilon_0, \varepsilon_1, \varepsilon_2} \mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \left[(\varepsilon_2 - \varepsilon_1)^k + (\varepsilon_1 - \varepsilon_0)^k - (\varepsilon_2 - \varepsilon_0)^k \right] \\ &= \left(\varepsilon^k + (-1)^k \varepsilon^k \right) \cdot \left[\mathcal{P}(+, -, +) + \mathcal{P}(-, +, -) \right] \end{aligned} \quad (\text{A1})$$

Note here that $\mathcal{P}(+, -, +)$ refers to the sequence $\{\varepsilon_0 = +\varepsilon/2, \varepsilon_1 = -\varepsilon/2, \varepsilon_2 = +\varepsilon/2\}$ and similarly for $\mathcal{P}(-, +, -)$. The last line in Eq. (A1) is reached by using the fact $\varepsilon_i = \pm\varepsilon/2, \forall i$ and that all other sequences drop out of the above summation regardless of the probability. Given that one is free to choose any arbitrary distribution $\mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2)$, we arrive at a Leggett-Garg inequality for fluctuating work;

$$0 \leq M_k \leq \varepsilon^k + (-1)^k \varepsilon^k, \quad \forall k. \quad (\text{A2})$$

where the lower bound is achieved by choosing $\mathcal{P}(+, -, +) = \mathcal{P}(-, +, -) = 0$ and the upper bound by setting $\mathcal{P}(+, -, +) = \mathcal{P}(-, +, -) = \frac{1}{2}$.

Appendix B: Derivation of Eq. (8)

We begin with the standard formula for the joint probability to observe energy $\varepsilon_i = \pm\frac{\varepsilon}{2}$ at time t_i and then energy $\varepsilon_j = \pm\frac{\varepsilon}{2}$ at time t_j from projective measurements of the Hamiltonian;

$$\mathcal{P}\left(\varepsilon_i = \pm\frac{\varepsilon}{2}, \varepsilon_j = \pm\frac{\varepsilon}{2}\right) = \text{Tr}[\hat{\rho} \hat{P}_{a_i}^{\pm}] \cdot \text{Tr}[\hat{P}_{a_j}^{\pm} \hat{P}_{a_i}^{\pm}]. \quad (\text{B1})$$

The above formula can be simplified by using the fact that each projector can be written as $\hat{P}_{a_i}^{\pm} = \frac{1}{2}(\hat{I} \pm \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})$, and then applying the identity $\text{Tr}[(\mathbf{a} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{b} \cdot \hat{\boldsymbol{\sigma}})] = 2\mathbf{a} \cdot \mathbf{b}$. This leads to

$$\begin{aligned} \mathcal{P}\left(\varepsilon_i = -\frac{\varepsilon}{2}, \varepsilon_j = +\frac{\varepsilon}{2}\right) &= \frac{1}{16} \text{Tr} \left[(\hat{I} + \mathbf{r} \cdot \hat{\boldsymbol{\sigma}}) \cdot (\hat{I} - \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}) \right] \cdot \text{Tr} \left[(\hat{I} + \mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) \cdot (\hat{I} - \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}) \right] \\ &= \frac{1}{4} \left[(1 - \mathbf{r} \cdot \mathbf{a}_i) \cdot (1 - \mathbf{a}_j \cdot \mathbf{a}_i) \right], \end{aligned} \quad (\text{B2})$$

and similarly

$$\mathcal{P}\left(\varepsilon_i = +\frac{\varepsilon}{2}, \varepsilon_j = -\frac{\varepsilon}{2}\right) = \frac{1}{4} \left[(1 + \mathbf{r} \cdot \mathbf{a}_i) \cdot (1 - \mathbf{a}_j \cdot \mathbf{a}_i) \right] \quad (\text{B3})$$

Finally from the definition of the moments for work, Eq. (2), we arrive at Eq. (8);

$$\begin{aligned} \langle W^k(t_i, t_j) \rangle &= \sum_{\varepsilon_i, \varepsilon_j} \mathcal{P}(\varepsilon_i, \varepsilon_j) (\varepsilon_j - \varepsilon_i)^k \\ &= \varepsilon^k \left[\mathcal{P}(-, +) + (-1)^k \mathcal{P}(+, -) \right] \\ &= \frac{(1 - \mathbf{a}_j \cdot \mathbf{a}_i)}{4} \varepsilon^k \left[1 - \mathbf{r} \cdot \mathbf{a}_i + (-1)^k (1 + \mathbf{r} \cdot \mathbf{a}_i) \right]. \end{aligned} \quad (\text{B4})$$

TABLE I. Displays the various combinations of θ_{01} and θ_{12} in Eq. (C6)

θ_{01}	θ_{12}	$\cos(\theta_{12}) + \cos(\theta_{01}) - \cos(\theta_{12} + \theta_{01})$	$\sin(\theta_{12}) + \sin(\theta_{01}) - \sin(\theta_{12} + \theta_{01})$
0	0	1	0
0	$-\lambda\varepsilon$	1	0
0	$\lambda\varepsilon$	1	0
$-\lambda\varepsilon$	$+\lambda\varepsilon$	$2\cos(\lambda\varepsilon) - 1$	0
$\lambda\varepsilon$	$-\lambda\varepsilon$	$2\cos(\lambda\varepsilon) - 1$	0
$-\lambda\varepsilon$	0	1	0
$\lambda\varepsilon$	0	1	0

Now the moments can be substituted into the Leggett-Garg equation Eq. (3). For even k the inequality is state independent and can be rewritten in terms of the angles between each Bloch vector, denoted $\mathbf{a}_i \cdot \mathbf{a}_j = \cos(\theta_{ij})$;

$$M_k = \frac{\varepsilon^k}{2} \left[1 + \cos(\theta_{10} + \theta_{21}) - \cos(\theta_{21}) - \cos(\theta_{10}) \right] \quad (\text{B5})$$

where we have used the trigonometric relation $\theta_{10} + \theta_{21} = \theta_{20}$. The lower bound, Eq. (9) is obtained by setting the angles to $\theta_{10} = \theta_{21} = \pi/3$. To bound Eq. (3) for odd k we note the following relation; $\mathbf{r} \cdot \mathbf{a}_1 = \tanh(\beta\varepsilon/2)\cos(\theta_{01})$. In this case Eq. (3) can be written as follows;

$$M_k = \frac{\varepsilon^k}{2} \tanh(\beta\varepsilon/2) \left[\cos(\theta_{01})\cos(\theta_{12}) - \cos(\theta_{01} + \theta_{12}) \right] \quad (\text{B6})$$

The bound Eq. (10) is obtained by setting $\theta_{10} = \theta_{21} = \pi/2$.

Appendix C: Derivation of Eq. (13) and Eq. (14)

The characteristic function for work done by a system driven unitarily in time between $t = t_i$ and $t = t_j$ is given by the following;

$$G(\lambda, t_i, t_j) = \langle e^{i\lambda W(t_i, t_j)} \rangle, \quad (\text{C1})$$

$$= \sum_{\varepsilon_i, \varepsilon_j} \mathcal{P}(\varepsilon_i, \varepsilon_j) e^{i\lambda(\varepsilon_j - \varepsilon_i)}, \quad (\text{C2})$$

where $\mathcal{P}(\varepsilon_i, \varepsilon_j)$ is some arbitrary joint probability governing the statistics of energy at two separate times. Under the assumptions (i) and (ii) for macrorealism this again implies the existence of a three-time distribution of the form Eq. (4), meaning that Eq. (12) can be rewritten as

$$L_\lambda = G(\lambda, t_1, t_2) + G(\lambda, t_0, t_1) - G(\lambda, t_0, t_2) \quad (\text{C3})$$

$$= \sum_{\varepsilon_1, \varepsilon_2} \mathcal{P}(\varepsilon_1, \varepsilon_2) e^{i\lambda(\varepsilon_2 - \varepsilon_1)} + \sum_{\varepsilon_0, \varepsilon_1} \mathcal{P}(\varepsilon_0, \varepsilon_1) e^{i\lambda(\varepsilon_1 - \varepsilon_0)} - \sum_{\varepsilon_0, \varepsilon_2} \mathcal{P}(\varepsilon_0, \varepsilon_2) e^{i\lambda(\varepsilon_2 - \varepsilon_0)} \quad (\text{C4})$$

$$= \sum_{\varepsilon_0, \varepsilon_1, \varepsilon_2} \mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \left[e^{i\lambda(\varepsilon_2 - \varepsilon_1)} + e^{i\lambda(\varepsilon_1 - \varepsilon_0)} - e^{i\lambda(\varepsilon_2 - \varepsilon_0)} \right]. \quad (\text{C5})$$

Defining $\theta_{ij} = \lambda(\varepsilon_j - \varepsilon_i)$, the real and imaginary parts of Eq. (12) are then

$$\begin{aligned} \text{Re}(L_\lambda) &= \sum_{\varepsilon_0, \varepsilon_1, \varepsilon_2} \mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \left[\cos(\theta_{12}) + \cos(\theta_{01}) - \cos(\theta_{01} + \theta_{12}) \right] \\ \text{Im}(L_\lambda) &= \sum_{\varepsilon_0, \varepsilon_1, \varepsilon_2} \mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2) \left[\sin(\theta_{12}) + \sin(\theta_{01}) - \sin(\theta_{01} + \theta_{12}) \right] \end{aligned} \quad (\text{C6})$$

To obtain bounds on the real and imaginary terms in Eq. (C6) we use the fact that $\varepsilon_i = \pm \frac{\varepsilon}{2}$ at all times and consider the various combinations of θ_{01} and θ_{12} shown in Table I. Given that we are free to choose any $\mathcal{P}(\varepsilon_0, \varepsilon_1, \varepsilon_2)$, the bounds in Eq. (13) and Eq. (14) follow immediately.

Appendix D: Maximal violations for the characteristic function

Here we derive the quantum bounds for Eq. (12). The moments of work obtained by successive projective measurements are given by Eq. (8), and can be substituted into Eq. (15) by using

$$G_\lambda(t_i, t_j) = \langle e^{i\lambda W(t_i, t_j)} \rangle = \sum_{k=0}^{\infty} \frac{(i\lambda)^k}{k!} \langle W^k(t_i, t_j) \rangle. \quad (\text{D1})$$

where we have used the series expansion for the complex exponential function in terms of the work moments. This leads to an expression for the real part of the characteristic function;

$$\begin{aligned} \Re(G_\lambda(t_i, t_j)) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \lambda^{2k} \langle W^{2k}(t_i, t_j) \rangle, \\ &= 1 + \frac{(1 - \mathbf{a}_j \cdot \mathbf{a}_i)}{2} \sum_{k=1}^{\infty} \frac{(-1)^k}{(2k)!} (\lambda \varepsilon)^{2k} \\ &= 1 + \frac{(1 - \mathbf{a}_j \cdot \mathbf{a}_i)}{2} [\cos(\lambda \varepsilon) - 1], \end{aligned} \quad (\text{D2})$$

Substituting Eq. (D2) into Eq. (12) yields the following;

$$\begin{aligned} \Re(L_\lambda) &= \Re \left(G_\lambda(t_1, t_2) + G_\lambda(t_0, t_1) - G_\lambda(t_0, t_2) \right) \\ &= 1 + \frac{(\cos(\lambda \varepsilon) - 1)}{2} \left[1 + \mathbf{a}_2 \cdot \mathbf{a}_0 - \mathbf{a}_2 \cdot \mathbf{a}_1 - \mathbf{a}_1 \cdot \mathbf{a}_0 \right], \end{aligned} \quad (\text{D3})$$

To bound Eq. (D3) we denote $\mathbf{a}_{i+1} \cdot \mathbf{a}_i = \cos(\phi_i)$ and use the trigonometric relation $\mathbf{a}_2 \cdot \mathbf{a}_0 = \cos(\phi_0 + \phi_1)$. This leads to the following inequality;

$$\Re(L_\lambda) \leq \frac{5}{4} - \frac{1}{4} \cos(\lambda \varepsilon), \quad (\text{D4})$$

where the upper bound is obtained by choosing $\theta_{10} = \theta_{21} = \pi/3$, as with the Leggett-Garg inequality for even moments of work in Eq. (9).

Turning to the imaginary part of Eq. (15) and using the expression Eq. (8) for the work moments we find the following;

$$\begin{aligned} \Im(G_\lambda(t_i, t_j)) &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \lambda^{2k+1} \langle W^{2k+1}(t_i, t_j) \rangle \\ &= \frac{(\mathbf{r} \cdot \mathbf{a}_i)(\mathbf{a}_j \cdot \mathbf{a}_i - 1)}{2} \sin(\lambda \varepsilon). \end{aligned} \quad (\text{D5})$$

For the bounds on the imaginary term, we substitute Eq. (D5) into Eq. (12) and choose the same set of parameters that were applied to Eq. (10), namely $\theta_{10} = \theta_{21} = \pi/2$, ending with Eq. (17).

Appendix E: Maximum violations for full-counting statistics

In this section we provide a derivation of Eq. (19) and Eq. (20). We first begin with the characteristic function Eq. (18) for a two-level system described by the Hamiltonian Eq. (6);

$$G_\lambda^{\text{FCS}}(t_i, t_j) = \text{Tr}[\hat{\rho} e^{-i\frac{\lambda}{4} \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} e^{i\frac{\lambda}{2} \mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}} e^{-i\frac{\lambda}{4} \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}}] \quad (\text{E1})$$

We now expand the exponential operators using $e^{-i\lambda \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} = \cos(\lambda) \hat{I} - i \sin(\lambda) \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}$;

$$\begin{aligned} e^{-i\frac{\lambda}{4} \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} e^{i\frac{\lambda}{2} \mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}} e^{-i\frac{\lambda}{4} \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} &= \cos^2(\lambda \varepsilon/2) \hat{I} - \frac{i}{2} \sin(\lambda \varepsilon) (\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}) + i \sin(\lambda \varepsilon/2) \cos^2(\lambda \varepsilon/4) (\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) \\ &\quad - \frac{1}{2} \sin^2(\lambda \varepsilon/2) \{ (\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}), (\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) \} - i \sin^2(\lambda \varepsilon/4) \sin(\lambda \varepsilon/2) (\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}) (\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) (\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}) \end{aligned} \quad (\text{E2})$$

Now we substitute this expansion back into $G_\lambda^{\text{FCS}}(t_i, t_j)$ above, along with $\hat{\rho} = \frac{1}{2}(\hat{I} + \mathbf{r} \cdot \hat{\boldsymbol{\sigma}})$, and apply the identity $\text{Tr}[(\mathbf{a} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{b} \cdot \hat{\boldsymbol{\sigma}})] = 2\mathbf{a} \cdot \mathbf{b}$;

$$\begin{aligned} G_\lambda^{\text{FCS}}(t_i, t_j) &= \cos^2(\lambda\varepsilon/2) - \frac{i}{2}\sin(\lambda\varepsilon)(\mathbf{r} \cdot \mathbf{a}_i) + i\sin(\lambda\varepsilon/2)\cos^2(\lambda\varepsilon/4)(\mathbf{r} \cdot \mathbf{a}_j) + \sin^2(\lambda\varepsilon/2)(\mathbf{a}_i \cdot \mathbf{a}_j) \\ &\quad + \frac{1}{4}\sin^2(\lambda\varepsilon/2)\text{Tr}\left[(\mathbf{r} \cdot \hat{\boldsymbol{\sigma}})\{(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})\}\right] \\ &\quad - \frac{i}{2}\sin^2(\lambda\varepsilon/4)\sin(\lambda\varepsilon/2)\left(\text{Tr}\left[(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})\right] + \text{Tr}\left[(\mathbf{r} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})\right]\right) \end{aligned} \quad (\text{E3})$$

To simplify Eq. (E3) we make use of the identity $(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) = (\mathbf{a}_i \cdot \mathbf{a}_j)\hat{I} + i(\mathbf{a}_i \times \mathbf{a}_j) \cdot \hat{\boldsymbol{\sigma}}$. This leads to the following set of relations;

$$\text{Tr}\left[(\mathbf{r} \cdot \hat{\boldsymbol{\sigma}})\{(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})\}\right] = 2i[(\mathbf{a}_i \times \mathbf{a}_j) \cdot \mathbf{r} + (\mathbf{a}_j \times \mathbf{a}_i) \cdot \mathbf{r}] = 0, \quad (\text{E4})$$

$$\text{Tr}\left[(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})\right] = 2i(\mathbf{a}_j \times \mathbf{a}_i) \cdot \mathbf{a}_i = 0, \quad (\text{E5})$$

$$\text{Tr}\left[(\mathbf{r} \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}})(\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}})\right] = 2(\mathbf{r} \cdot \mathbf{a}_i)(\mathbf{a}_j \cdot \mathbf{a}_i) - 2(\mathbf{r} \times \mathbf{a}_i) \cdot (\mathbf{a}_j \times \mathbf{a}_i). \quad (\text{E6})$$

Substituting these into Eq. (E3), collecting the real and imaginary terms and simplifying leads to Eq. (19);

$$\Re\left(G_\lambda^{\text{FCS}}(t_i, t_j)\right) = 1 + \frac{(1 - \mathbf{a}_j \cdot \mathbf{a}_i)}{2} [\cos(\lambda\varepsilon) - 1] \quad (\text{E7})$$

and similarly

$$\begin{aligned} \text{Im}\left(G_\lambda^{\text{FCS}}(t_i, t_j)\right) &= \sin(\lambda\varepsilon/2)\cos^2(\lambda\varepsilon/4)(\mathbf{r} \cdot \mathbf{a}_j) - \frac{1}{2}\sin(\lambda\varepsilon)(\mathbf{r} \cdot \mathbf{a}_i) \\ &\quad - \sin(\lambda\varepsilon/2)\sin^2(\lambda\varepsilon/4)\left[(\mathbf{r} \cdot \mathbf{a}_i)(\mathbf{a}_j \cdot \mathbf{a}_i) - (\mathbf{r} \times \mathbf{a}_i) \cdot (\mathbf{a}_j \times \mathbf{a}_i)\right] \end{aligned} \quad (\text{E8})$$

Noting that one can rewrite $(\mathbf{r} \times \mathbf{a}_1) \cdot (\mathbf{a}_2 \times \mathbf{a}_1) = \tanh(\beta\varepsilon/2) [\cos(\theta_{01} + \theta_{12}) - \cos(\theta_{01})\cos(\theta_{12})]$, substituting Eq. (E8) into the Leggett-Garg equation Eq. (12) leads to the following simple expression;

$$\text{Im}(L_\lambda^{\text{FCS}}) = 2\sin^2(\lambda\varepsilon/4)\sin(\lambda\varepsilon/2)\tanh(\beta\varepsilon/2)\left[\cos(\theta_{01} + \theta_{12}) - \cos(\theta_{01})\cos(\theta_{12})\right] \quad (\text{E9})$$

with the upper bound Eq. (20) given by setting $\theta_{10} = \theta_{21} = \pi/2$.

Appendix F: Maximum violations for the Margenau-Hill distribution

In this section we provide details of Leggett-Garg violations found from the characteristic function Eq. (21). As noted in the main text, a work distribution based on the Margenau-Hill distribution was proposed in [23] to extend the definition of fluctuating to states with initial energy coherences. Firstly, note that the Margenau-Hill distribution can be used to define a joint probability for the successive energy outcomes $\varepsilon_i, \varepsilon_j$ at times t_i and t_j respectively [40];

$$\mathcal{P}\left(\varepsilon_i = \pm \frac{\varepsilon}{2}, \varepsilon_j = \pm \frac{\varepsilon}{2}\right) = \text{Tr}[\hat{\rho} \hat{P}_{a_i}^\pm \star \hat{P}_{a_j}^\pm]. \quad (\text{F1})$$

The corresponding characteristic function of work is then as follows;

$$\begin{aligned} G_\lambda^{\text{MH}}(t_i, t_j) &= \langle e^{i\lambda W(t_i, t_j)} \rangle \\ &= \sum_{\varepsilon_i, \varepsilon_j} \mathcal{P}(\varepsilon_i, \varepsilon_j) e^{i\lambda(\varepsilon_j - \varepsilon_i)} \\ &= \text{Tr}[\hat{\rho} e^{i\lambda \mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}} \star e^{-i\lambda \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}}. \end{aligned} \quad (\text{F2})$$

Taking the expansion $e^{-i\lambda \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} = \cos(\lambda)\hat{I} - i\sin(\lambda)\mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}$, we obtain the following operator expression;

$$e^{i\lambda \mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}} \star e^{-i\lambda \mathbf{a}_i \cdot \hat{\boldsymbol{\sigma}}} = \cos^2(\lambda\varepsilon/2)\hat{I} + i\sin(\lambda\varepsilon)(\mathbf{a}_j - \mathbf{a}_i) \cdot \hat{\boldsymbol{\sigma}} + \sin^2(\lambda\varepsilon/2)(\mathbf{a}_j \cdot \hat{\boldsymbol{\sigma}}) \quad (\text{F3})$$

Substituting this into Eq. (F2) and applying the same identities used throughout Appendix E gives an expression for the characteristic function;

$$G_{\lambda}^{\text{MH}}(t_i, t_j) = \cos^2(\lambda \varepsilon / 2) + \sin^2(\lambda \varepsilon / 2)(\mathbf{a}_j \cdot \mathbf{a}_i) + i \sin(\lambda \varepsilon)(\mathbf{a}_j - \mathbf{a}_i) \cdot \mathbf{r} \quad (\text{F4})$$

We can immediately see that the real part of $G_{\lambda}^{\text{MH}}(t_i, t_j)$ is again equivalent to Eq. (19). As for the imaginary term, a straightforward substitution of Eq. (F4) into Eq. (12) reveals that

$$\Im(L_{\lambda}^{\text{MH}}) = 0, \quad (\text{F5})$$

for all choices of θ_{ij} .