Quantum Fluctuation Theorems, Contextuality, and Work Quasiprobabilities

Matteo Lostaglio
ICFO-Institut de Ciencies Fotoniques, Barcelona Institute of Science and Technology, Castelldefels (Barcelona) 08860, Spain

(Received 2 June 2017; revised manuscript received 7 December 2017; published 24 January 2018)

We discuss the role of contextuality within quantum fluctuation theorems, in the light of a recent no-go result by Perarnau-Llobet *et al.* We show that any fluctuation theorem reproducing the two-point-measurement scheme for classical states either admits a notion of work quasiprobability or fails to describe protocols exhibiting contextuality. Conversely, we describe a protocol that smoothly interpolates between the two-point-measurement work distribution for projective measurements and Allahverdyan's work quasiprobability for weak measurements, and show that the negativity of the latter is a direct signature of contextuality.

DOI: 10.1103/PhysRevLett.120.040602

While quantum thermodynamics thrived in recent years [1,2], we still lack clear evidence that there are any thermodynamically relevant protocols whose results cannot, in a precise sense, be "simulated" classically. In this Letter we show that such protocols do indeed exist; the solution lies in a long-standing debate that surrounded the definition of work in so-called fluctuation theorems (FTs).

FTs are one of the most important set of results in nonequilibrium thermodynamics [3–6]. In their simplest form, a classical system initially prepared in a thermal state at temperature T is driven out of equilibrium. This is achieved by changing the parameters of the Hamiltonian from $H(0)$ to $H(\tau)$, according to a fixed protocol, while keeping the system isolated from the environment. Each repetition requires an amount of work w, corresponding to the realizations of a random variable W. Denote by $Z_{H(t)}$ the partition function of $H(t)$ and by k Boltzmann's constant. The free-energy difference between the equilibrium state with respect to the final Hamiltonian and the initial equilibrium state reads $\Delta F = -kT \log Z_{H(\tau)}/Z_{H(0)}$. Then, the Jarzynski equality characterizes work fluctuations in the protocol above [3],

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \tag{1}$$

where $\langle \cdot \rangle$ denotes averaging and $\beta = (kT)^{-1}$. Through Jensen's inequality, Eq. (1) can be seen as a generalization of the standard thermodynamic inequality $\langle W \rangle \geq \Delta F$. However, it also encodes information about fluctuations, e.g., the probability that $W > \Delta F + x$ is bounded by $e^{-\beta x}$. Equation (1) and related equalities provide refined statements of the second law of thermodynamics beyond averages, are valid at the microscopic scale far from the thermodynamic limit, recover results from linear response theory, and give a way to measure the free energy from nonequilibrium measurements of work [7].

Much effort has been devoted to finding analogous results characterizing work fluctuations of quantum processes and in the presence of quantum coherence [6,8–17]. One of the main challenges is the definition of work for quantum systems (see [18] and references therein). Work in closed systems is defined, in classical physics, as the energy difference between the initial and final phase space point, while in the quantum setting the conventional approach adopts the two-point-measurement (TPM) scheme [19]. The idea is to define work as the energy difference between the outcomes of two projective measurements of energy, performed at the beginning and at the end of the protocol.

Whilst the TPM definition of work has become standard [6], various authors have claimed the approach has some important limitations [8]. First, some have observed that, projective measurements being invasive, the average work according to this definition does *not* in general coincide with the average energy change if no measurement is performed [11,12,17]. Second, others pointed to the fact that quantum coherence and entanglement are destroyed by the TPM scheme at the start of the protocol [12,14,20,21].

Concerning the first of the two issues raised, a recent no-go result has shown that it is a universal feature, proving that within quantum theory no FT can simultaneously reproduce the TPM scheme for classical states and respect the identification of average work with average energy change [16]. However, it is unclear if this identification is indeed a property we should impose on quantum FTs. Regarding the second issue, one can argue that the evolution described by $H(t)$ generates quantum coherence again and, in fact, interference effects do appear in the TPM work distribution when compared to the classical limit [18]. Hence, what are the fundamental limitations, if any, of current FTs?

There is a caveat in the no-go theorem of Ref. [16], in that it can be circumvented at the price of extending the work distributions to quasiprobabilities [11,12,17,20]. The occurrence of negative values in such distributions has been considered a limitation by some authors [11,22], while others claimed they signal quantum effects [12,17], since they can be related to the violation of the Leggett-Garg inequality [23,24]. A second natural question is, then, to what extent quasiprobabilities are a necessary ingredient to capture quantum effects in FTs and what, exactly, we can infer from observing their negativity.

Here we contribute to these issues by showing that: (1) Any FT that reproduces the two-point-measurement scheme for classical states is either based on work quasiprobabilities or admits a noncontextual description. In other words, restricting to work probabilities prevents us from probing stronger forms of nonclassicality. (2) Conversely, in a generalization of the TPM scheme we show that the appearance of negativity in a work quasiprobability signals the onset of contextuality.

Setting the scene: A no-go result for fluctuation theorems.—A system prepared in a state ρ undergoes a unitary evolution U between time 0 and τ, induced by a time-dependent Hamiltonian $H(t)$. By comparison to the classical case, one wishes to define a distribution $p(w|\mathcal{P})$, giving the probability that the protocol $\mathcal{P}$ requires an amount of work w to be realized. Classically, w can be defined as the internal energy change of the system. Quantum mechanically, however, the mere act of probing the initial state ρ will induce a disturbance $\rho \mapsto \sigma$. We will then consider a general protocol $\mathcal{P}$ schematically described as

$$\mathcal{P} \coloneqq \{(H(0), \rho) \mapsto (H(\tau), U\sigma U^\dagger)\}, \tag{2}$$

where $U = \mathcal{T}\exp\left(-i\int_0^\tau dt H(t)\right)$ and $\mathcal{T}$ is the time-ordering operator. To fix the notation,

$$H(0) = \sum_i E_i |i\rangle\langle i| \coloneqq \sum_i E_i \mathcal{E}_i, \qquad H(\tau) = \sum_i E'_i |i'\rangle\langle i'|.$$

The standard framework extracts the work statistics from the TPM scheme, measuring $H(0)$ at the beginning of the protocol and $H(\tau)$ at the end [19]. Upon observing, respectively, outcomes i and j, one sets $w = E'_j - E_i$. The random variable W defined in this way satisfies Eq. (1) [6].

Note, however, that the first measurement can modify the subsequent statistics. This becomes evident looking at the problem in Heisenberg picture, where one can think of the TPM scheme as the sequential measurement of $H(0)$ followed by $U^\dagger H(\tau) U$. Whenever $[H(0), U^\dagger H(\tau) U] \neq 0$ and $[\rho, H(0)] \neq 0$, the statistics of the second measurement will be disturbed by the first. In fact, $\sigma = \mathcal{D}_{H(0)}(\rho)$, with $\mathcal{D}_{H(0)}$ denoting the operation that fully dephases in the eigenbasis of $H(0)$.

A recent result formalized this clash into a no-go theorem [16]. Consider the definition as follows: *Definition 1.*—A protocol $\mathcal{P}$ is called a FT protocol if, for initial states with no quantum coherence, i.e., ρ satisfying $[\rho, H(0)] = 0$, the results of the TPM scheme are recovered,

$$p(w|\mathcal{P}) = p_{\mathrm{TPM}}(w|\mathcal{P}) \coloneqq \sum_{E'_j - E_i = w} p_i p_{j|i},$$

where $p_i = \langle i|\rho|i\rangle$, $p_{j|i} = |\langle j'|U|i\rangle|^2$.

In other words, FT protocols are those that recover the TPM scheme at least in those situations in which the measurement does not introduce any disturbance to the evolution of the system. This is sufficient to reproduce Eq. (1) for thermal initial states and to match the expected definition of work in the classical limit (in the cases analysed in [18]). Consider now the following assumptions: (1) (Work distribution) $\mathcal{P}$ measures a work probability distribution $p(w|\mathcal{P})$ convex under mixtures of protocols. More precisely, let $q \in [0,1]$ and $\mathcal{P}^i$ be protocols only differing by the initial preparation ρ_i. One requires that if $\rho_0 = q\rho_1 + (1-q)\rho_2$ then

$$p(w|\mathcal{P}^0) = q p(w|\mathcal{P}^1) + (1-q) p(w|\mathcal{P}^2).$$

(2) (Average work) The average measured work should reproduce the average energy change induced by the unitary process on the initial state

$$\langle W \rangle \coloneqq \sum_w p(w|\mathcal{P}) w = \mathrm{Tr}(U\rho U^\dagger H(\tau)) - \mathrm{Tr}(\rho H(0)).$$

Assumption 1 includes the natural demand that, upon conditioning the choice of the protocol on a coin toss, the measured fluctuations are simply the convex combination of those observed in the individual protocols. Assumption 2 is based on the identification of average work with average energy change in closed systems. The main result of Ref. [16] is that no FT protocol can satisfy both assumptions 1 and 2 when the system has no matching gaps, i.e., $E'_{j_1} - E_{i_1} \neq E'_{j_2} - E_{i_2}$ if $(j_1, i_1) \neq (j_2, i_2)$.

Genuinely nonclassical effects.—While the result of Ref. [16] was phrased mainly as an incompatibility between the requirement that $\mathcal{P}$ is a FT protocol and assumption 2, we note that assumption 1 contains the often implicit condition that a work distribution exists. Since work involves two generally noncommuting observables $H(0)$ and $U^\dagger H(\tau) U$, restrictions arise from the lack of a joint probability distribution for them.

The above point can be sharpened as follows. Broadly speaking, we want to make a distinction between a phenomenon that is essentially classical in nature from one that is irreducibly quantum mechanical [25]. We may think in terms of a challenge involving two parties, Alice and Bob. Alice sets up a quantum experiment, specifying the quantum systems involved, their interactions and the measurements performed, and tells Bob about this. Bob then prepares a box, in which he devises some classical mechanism that tries to reproduce the same statistics of Alice's experiment. If—despite Bob's best efforts—Alice can always find discrepancies between the statistics produced by Bob and her quantum experiment, we call the quantum phenomenon involved genuinely nonclassical.

The idea of classical mechanism is encapsulated in the notion of hidden variable model or ontological model. At the operational level, consider a set of instructions defining preparations procedures P and measurement procedures M with outcomes k; Alice observes k with probability $p(k|P,M)$. Bob's classical mechanism reproduces this statistics using a set of states λ that are, in general, randomly sampled from a set Λ according to a probability distribution $p(\lambda|P)$ every time the preparation P is performed. For example, Λ may be the phase space of the classical mechanism, but we are not limited to that. Moreover, Bob's mechanism can include a measurement device M that takes in the physical state λ and outputs an outcome k with probability $p(k|\lambda,M)$. Bob wins the challenge if he can reproduce the statistics $p(k|P,M)$ as an average over the unobserved states of the classical mechanism [26,27],

$$p(k|P,M) = \int_\Lambda d\lambda p(\lambda|P)p(k|\lambda,M). \qquad (3)$$

(For further details, see Sec. A of the Supplemental Material [28], which includes Refs. [29–43]).

For a mechanism to be classical it should operate in a noncontextual way. Specifically, the mechanism is called preparation noncontextual if $p(\lambda|P)$ is a function of the quantum state alone, i.e., $p(\lambda|P) \equiv p(\lambda|\rho)$; for example, Bob's mechanism cannot distinguish different ensembles associated to the same ρ. Furthermore, the mechanism is called measurement noncontextual if $p(k|\lambda,M)$ depends only on the positive-operator-valued measure (POVM) element M_k associated to the corresponding outcome of the measurement M, i.e., $p(k|\lambda,M) \equiv p(k|\lambda,M_k)$ [26,31]. If a mechanism is both preparation and measurement noncontextual, it is called universally noncontextual. See Sec. A of the Supplemental Material [28] for the relation to Kochen-Specker contextuality [44].

It is a question of fundamental as well as practical importance to know if FT protocols can uncover genuinely nonclassical phenomena. Here we make this precise by showing the following:

Theorem 1.—Assume the FT protocol $\mathcal{P}$ satisfies assumption 1. Then, there exists a universally noncontextual ontological model for every preparation ρ and measurement of $p(w|\mathcal{P})$.

Theorem 1 says that in any FT protocol we either (i) lift the assumption that the work distribution should be a probability (assumption 1), or (ii) only probe quantum effects admitting a classical noncontextual model.

Existence of a work distribution forces noncontextuality.—Let us prove Theorem 1. From now on, we focus on the case in which for every w there exists a unique couple of indexes i, j such that $E'_j - E_i = w$. Let P and M denote the preparation and measurement procedures involved in the FT protocol $\mathcal{P}$. As shown in Ref. [16], assumption 1 can be reformulated as follows: there exists a POVM $M(\mathcal{P}) = \{M_w(\mathcal{P})\}$, such that

$$p(w|\mathcal{P}) = \mathrm{Tr}(M_w(\mathcal{P})\rho), \qquad (4)$$

with $M_w(\mathcal{P})$ being a function of $H(0)$, U, and $H(\tau)$, but *not* of ρ.

We want to derive the existence of a noncontextual model reproducing the observed work statistics $p(w|\mathcal{P})$; i.e., from Eq. (3),

$$\mathrm{Tr}(\rho M_w(\mathcal{P})) = \int_\Lambda d\lambda p(\lambda|\rho)p(w|\lambda, M_w(\mathcal{P})). \qquad (5)$$

This, in fact, arises as a simple consequence of the main results of Ref. [16]. There it is shown that assumption 1 enforces on the FT protocol $M_w(\mathcal{P}) = M_w^{\mathrm{TPM}}(\mathcal{P})$, where $M_w^{\mathrm{TPM}}(\mathcal{P})$ is the two-point-measurement POVM [45], $M_w^{\mathrm{TPM}}(\mathcal{P}) = |\langle j'|U|i\rangle|^2|i\rangle\langle i|$. Λ can then be taken to be a space labeling energies and for all $H(0)$, U, and $H(\tau)$,

$$p(\lambda|\rho) = \langle\lambda|\rho|\lambda\rangle,$$
$$p(w|\lambda, M_w(\mathcal{P})) = \mathrm{Tr}(M_w^{\mathrm{TPM}}(\mathcal{P})|\lambda\rangle\langle\lambda|). \qquad (6)$$

Substituting Eqs. (6) in Eq. (5), and using Eq. (4) we get the claimed result.

As is known, there are prepare-and-measure experiments that *cannot* be reproduced by a noncontextual mechanism [26,30]. While Bob can in principle emulate core aspects of many quantum phenomena with a classical mechanism [46–48], contextuality is beyond his reach. Theorem 1 says that any FT protocol satisfying assumption 1 will not allow the correspondent experiments to manifest genuine nonclassicality; i.e., it will restrict us to probing a "fragment" of quantum theory [25,32] that admits a classical representation [30].

Some clarifications are now in order. First, when there is no POVM satisfying Eq. (4) (i.e., assumption 1 is lifted), we study if noncontextual ontological models exist for the scheme collecting the statistics through which $p(w|\mathcal{P})$ is reconstructed. Lifting assumption 1 expands the set of FT protocols $\mathcal{P}$ beyond the TPM POVM, but this is distinct from proving that any of the reconstruction protocols lacks a noncontextual mechanism.

Second, Ref. [43] highlighted that the absence of nonnegative quasiprobability representations for a given protocol and contextuality are the same concept. However, one should be careful not to identify the negativity of $p(w|\mathcal{P})$, reconstructed from a set of preparations and measurements in $\mathcal{P}$, with the negativity of *every* quasiprobability representations of such preparations and measurements [$p(w|\mathcal{P})$ is not a representation, see Sec. C of the Supplemental Material [28]]. It is a nontrivial task to define FT protocols in which negativity of a work quasiprobability can be provably associated to contextuality. This is what we will do in the rest of this Letter.

Negativity of the work distribution implies contextuality.—Since we are interested in FT protocols able to witness genuinely nonclassical features, because of Theorem 1 we investigate here the possibility of lifting assumption 1. This means that $p(w|\mathcal{P})$ is a quasiprobability, exhibiting negativity (as in Ref. [11]) or lacking convexity (as in Ref. [49]). Here we investigate the former possibility.

First, let us describe a family of protocols that smoothly interpolates between the TPM protocol, estimating $p_{\mathrm{TPM}}(w|\mathcal{P})$, and a protocol probing the recent work quasiprobability introduced by Allahverdyan in Ref. [11], denoted by $p_{\mathrm{weak}}(w|\mathcal{P})$. While the former is obtained in the strong (projective) measurement limit, the latter is achieved through a weak measurement [36,37]. Second, we will show that the negativity of $p_{\mathrm{weak}}(w|\mathcal{P})$ directly signals contextuality of the weak measurement protocol.

Consider the following one-parameter family of protocols, parametrized by $s \in \mathbb{R}$, involving steps as follows (for related schemes, see Refs. [20,22,50] and references therein): (1) A measurement device or "pointer," represented by a one-dimensional quantum system with canonical observables X and P, is prepared in a Gaussian state with spread s,

$$|\Psi\rangle = (\pi s^2)^{-1/4} \int dx \exp\left(-\frac{x^2}{2s^2}\right)|x\rangle. \tag{7}$$

The system is prepared in state ρ, initially uncorrelated from the device. (2) The device is coupled to the system through the interaction Hamiltonian $H_{\mathrm{int}} = g(t)\mathcal{E}_i \otimes P$ over a time interval $[-t_M, 0]$ (recall $\mathcal{E}_i = |i\rangle\langle i|$). We can choose units such that $g = \int_{-t_M}^{0} dt g(t) = 1$. After the interaction, a projective measurement $\{|x\rangle\langle x|\}$ on the device induces a corresponding POVM $\{M_x^s\}_{x\in\mathbb{R}}$ on the system. In the limit $s \to \infty$, this is called a weak measurement of $\mathcal{E}_i$. (3) At $t = 0$, the system is evolved according to U, the driving unitary of Eq. (2) (we neglect the free evolution of the system during the measurement). (4) Finally, at $t = \tau$, a projective measurement of $H(\tau)$ is performed on the system and outcome j is postselected.

Denote by q_j the probability of observing outcome j in the final measurement. Moreover, let $\langle X\rangle_j$ be the expectation value of the shift in the pointer upon postselecting outcome j. There are two important limits (see Sec. B of the Supplemental Material [28]): (1) In the strong measurement limit, $s \to 0$, $\{M_x^s\}$ approaches a projective measurement and

$$q_j\langle X\rangle_j \to p_{\mathrm{TPM}}(w|\mathcal{P}) = p_i p_{i|j}.$$

(2) In the weak measurement limit, $s \to \infty$,

$$q_j\langle X\rangle_j \to p_{\mathrm{weak}}(w|\mathcal{P}) = \mathrm{ReTr}(\rho\mathcal{E}_i\Pi_j).$$

Here we defined $\Pi_j = U^\dagger|j'\rangle\langle j'|U$ and $p_{\mathrm{weak}}(w|\mathcal{P})$ corresponds to the work quasidistribution recently introduced by Allahverdyan [11],

$$p_{\mathrm{weak}}(w|\mathcal{P}) = \mathrm{ReTr}(\rho\mathcal{E}_i\Pi_j), \qquad w = E'_j - E_i. \tag{8}$$

We see that both $p_{\mathrm{TPM}}(w|\mathcal{P})$ and $p_{\mathrm{weak}}(w|\mathcal{P})$ are obtained within the same general class of protocols. p_{weak} is known as the Margenau-Hill distribution [51]. The weak measurement protocol so defined is a FT protocol and we notice in passing that it satisfies assumption 2. Furthermore, it gives rise to a FT [11],

$$\langle e^{-\beta W}\rangle = e^{-\beta\Delta F}\Upsilon, \tag{9}$$

where $\Upsilon = \mathrm{ReTr}\big(U^\dagger\gamma(\tau)U\gamma(0)^{-1}\rho\big)$ and $\gamma(t) = e^{-\beta H(t)}/\mathrm{Tr}(e^{-\beta H(t)})$. If $\rho = \gamma(0)$, we have $\Upsilon = 1$ and the equality of Eq. (1) is recovered [for a full interpretation of Eq. (9), see Ref. [11]].

While p_{weak} can attain negative values, contrary to Ref. [11] we suggest this is not a limitation. Quite the opposite, due to Theorem 1 negativity is necessary to probe contextuality (p_{weak} is convex in ρ). Remarkably, it is also *sufficient*. Recall that a classical mechanism is called outcome deterministic if projective measurements M give a deterministic response, i.e., $p(k|\lambda, M) \in \{0, 1\}$. With reference to the protocols provided, we have the following.

Theorem 2.—Let ρ be a quantum state, $\mathcal{E}_i$ and Π_j projectors. If $\mathrm{ReTr}(\rho\mathcal{E}_i\Pi_j) < 0$, for s large enough there is no measurement non-contextual ontological model for preparation ρ, measurement $\{M_x^s\}_{x\in\mathbb{R}}$, and postselection Π_j that satisfies outcome determinism.

The complete proof is given in Sec. D of the Supplemental Material [28]. The quantity $\mathrm{ReTr}(\rho\mathcal{E}_i\Pi_j)/\mathrm{Tr}(\rho\Pi_j)$ is known as generalized weak value [52] and its negative values are called anomalous. Hence, the above theorem is an extension of the main result of Ref. [40] to mixed states. [One may define generalized weak values as $\mathrm{Tr}(\rho\mathcal{E}_i\Pi_j)/\mathrm{Tr}(\rho\Pi_j)$, since this is a direct generalization of the initial proposal [38,53]. However, complex weak values can be achieved in Gaussian quantum mechanics, which admits a noncontextual model [42,47]. Hence, following Ref. [40], we focused on the real part.] The result implies that the observation of a generalized anomalous weak value provides a proof of contextuality of the FT protocol introduced above (note that U is included in Π_j).

We need to be more precise here, since the theorem involves the condition of outcome determinism. First, note that this condition is indeed necessary to get any result: a measurement-noncontextual (but not outcome-deterministic) model exists for full quantum mechanics [26], and hence for any FT protocol. Second, note that many authors include outcome determinism in the definition of contextuality. This is the case of the Kochen and Specker theorem [44]. Third, the following corollary of Theorem 1 holds:

Corollary 3.—Assume quantum mechanics holds. If $\mathrm{ReTr}(\rho\mathcal{E}_i\Pi_j) < 0$, for s large enough there is no universally noncontextual model for preparation ρ, measurement $\{M_x^s\}_{x\in\mathbb{R}}$, and postselection Π_j.

This corollary holds because some elementary operational conditions (obviously satisfied by the operational theory associated to quantum mechanics) are sufficient to prove outcome determinism from preparation noncontextuality, as detailed in Refs. [26,31].

A consequence of Theorem 2 is that in the presence of noncommutativity there is always a state able to witness contextuality in the FT protocol given above: for any i, j, with $[\mathcal{E}_i, \Pi_j] \neq 0$, there are quantum states ρ such that $p_{\text{weak}}(w|\mathcal{P}) < 0$ [11]. Conversely, necessarily one must have $[\rho, \mathcal{E}_i] \neq 0$ and $[\rho, \Pi_j] \neq 0$ for some i, j to observe negative values of the work distribution.

Conclusions.—In this Letter we presented the first example of contextuality in a thermodynamic framework. The no-go result of Theorem 1 shows that FT protocols are unable to access contextuality, *unless* the notion of work distribution is extended to a work quasiprobability, lacking non-negativity or convexity. Conversely, from Theorem 2 we have seen that the negative values of a work quasiprobability (accessible through weak measurements) imply contextuality of a FT protocol that naturally generalizes the TPM scheme.

Importantly, since noncontextual models exist reproducing phenomena such as quantum interference, complementarity, and noncommutativity—among others [46,47]—contextuality cannot be understood as a consequence of measurement disturbance and lack of knowledge of an underlying classical variable. Hence, the negativity of p_{weak} is inherently nonclassical (differentiating these protocols from TPM schemes). While a FT exists for p_{weak} [Eq. (9)], it will be important to derive a *direct* thermodynamic interpretation of negativity.

This work also paves the way to an experimental verification of contextuality in a FT protocol, but more work needs to be done to make the present proposal robust to experimental imperfections [27,54].

Analogous questions are being investigated in the context of quantum computing [55–59], and this work suggests a path to finding quantum advantages in thermodynamics. Eventually we hope this will lead to the design of thermodynamic machines, such as engines, whose performance provably outperforms classical counterparts, independently of the specific assumptions on the underlying model.

I would like to thank A. Acin, J. Bowles, D. Jennings, K. Korzekwa, R. Kunjwal, M. Oszmaniec, and M. Perarnau-Llobet for insightful discussions and comments on a previous draft. I acknowledge financial support from the Spanish MINECO (Severo Ochoa Grant No. SEV-2015-0522 and Project No. QIBEQI FIS2016-80773-P), Fundacio Cellex, Generalitat de Catalunya (CERCA Programme and SGR 875), and COST Action MP1209.

[1] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, J. Phys. A **49**, 143001 (2016).
[2] S. Vinjanampathy and J. Anders, Contemp. Phys. **57**, 545 (2016).
[3] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[4] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).
[5] E. M. Sevick, R. Prabhakar, S. R. Williams, and D. J. Searles, Annu. Rev. Phys. Chem. **59**, 603 (2008).
[6] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).
[7] C. Jarzynski, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).
[8] A. E. Allahverdyan and T. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).
[9] T. Albash, D. A. Lidar, M. Marvian, and P. Zanardi, Phys. Rev. E **88**, 032146 (2013).
[10] A. E. Rastegin and K. Życzkowski, Phys. Rev. E **89**, 012127 (2014).
[11] A. E. Allahverdyan, Phys. Rev. E **90**, 032137 (2014).
[12] P. Solinas and S. Gasparinetti, Phys. Rev. E **92**, 042150 (2015).
[13] G. Manzano, J. M. Horowitz, and J. M. R. Parrondo, Phys. Rev. E **92**, 032129 (2015).
[14] J. Aberg, arXiv:1601.01302.
[15] Á. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, Phys. Rev. X **6**, 041017 (2016).
[16] M. Perarnau-Llobet, E. Bäumer, K. V. Hovhannisyan, M. Huber, and A. Acin, Phys. Rev. Lett. **118**, 070601 (2017).
[17] H. J. Miller and J. Anders, New J. Phys. **19**, 062001 (2017).
[18] C. Jarzynski, H. T. Quan, and S. Rahav, Phys. Rev. X **5**, 031038 (2015).
[19] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007).
[20] P. Solinas and S. Gasparinetti, Phys. Rev. A **94**, 052103 (2016).
[21] M. Lostaglio, D. Jennings, and T. Rudolph, Nat. Commun. **6**, 6383 (2015).
[22] P. Talkner and P. Hänggi, Phys. Rev. E **93**, 022131 (2016).
[23] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).
[24] A. A. Clerk, Phys. Rev. A **84**, 043824 (2011).
[25] D. Jennings and M. Leifer, Contemp. Phys. **57**, 60 (2016).
[26] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[27] R. Kunjwal and R. W. Spekkens, Phys. Rev. Lett. **115**, 110403 (2015).
[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.040602 for an extended discussion of the basic notions and the proofs.
[29] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[30] C. Ferrie, Rep. Prog. Phys. **74**, 116001 (2011).
[31] R. W. Spekkens, Found. Phys. **44**, 1125 (2014).
[32] M. S. Leifer, Quanta **3**, 67 (2014).
[33] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[34] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[35] M. J. W. Hall, Phys. Rev. A **69**, 052113 (2004).
[36] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd, Rev. Mod. Phys. **86**, 307 (2014).
[37] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. **60**, 1351 (1988).
[38] J. Dressel, S. Agarwal, and A. N. Jordan, Phys. Rev. Lett. **104**, 240401 (2010).
[39] A. Hosoya and Y. Shikano, J. Phys. A **43**, 385307 (2010).
[40] M. F. Pusey, Phys. Rev. Lett. **113**, 200401 (2014).
[41] M. F. Pusey and M. S. Leifer, Electron. Proc. Theor. Comput. Sci. **195**, 295 (2015).

[42] A. Karanjai, E. G. Cavalcanti, S. D. Bartlett, and T. Rudolph, New J. Phys. **17**, 073015 (2015).

[43] R. W. Spekkens, Phys. Rev. Lett. **101**, 020401 (2008).

[44] S. Kochen and E. P. Specker, in *The Logico-Algebraic Approach to Quantum Mechanics* (Springer, New York, 1975), pp. 293–328.

[45] A. J. Roncaglia, F. Cerisola, and J. P. Paz, Phys. Rev. Lett. **113**, 250601 (2014).

[46] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).

[47] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012).

[48] R. W. Spekkens, in *Quantum Theory: Informational Foundations and Foils* (Springer, New York, 2016), pp. 83–135.

[49] R. Sampaio, S. Suomela, T. Ala-Nissila, J. Anders, and T. Philbin, arXiv:1707.06159.

[50] P. P. Hofer, Quantum **1**, 32 (2017).

[51] H. Margenau and R. N. Hill, Prog. Theor. Phys. **26**, 722 (1961).

[52] H. M. Wiseman, Phys. Rev. A **65**, 032111 (2002).

[53] Y. Aharonov and L. Susskind, Phys. Rev. **155**, 1428 (1967).

[54] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, 11780 (2016).

[55] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).

[56] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).

[57] V. Veitch, S. H. Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).

[58] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, Phys. Rev. X **5**, 021003 (2015).

[59] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, Phys. Rev. Lett. **119**, 120505 (2017).