

Contextuality in Anomalous Heat Flow

Naim E. Comar^{1,2,*} Danilo Cius^{1,†} Luis F. Santos^{1,‡} Rafael Wagner^{3,4,5,6,§} and
Bárbara Amaral^{1,||}

¹Department of Mathematical Physics, *Institute of Physics, University of São Paulo*, Rua do Matão 1371, São Paulo 05508-090, São Paulo, Brazil

²International Institute of Physics, *Federal University of Rio Grande do Norte*, 59078-970, Natal, Brazil

³Institute of Theoretical Physics, *Ulm University*, Albert-Einstein-Allee 11 89081, Ulm, Germany

⁴International Iberian Nanotechnology Laboratory (INL), 4715-330 Braga, Portugal

⁵Centro de Física, *Universidade do Minho*, 4710-057 Braga, Portugal

⁶Department of Physics “E. Fermi”, *University of Pisa*, I-56127 Pisa, Italy



(Received 20 August 2024; revised 13 March 2025; accepted 19 August 2025; published 22 September 2025)

In classical thermodynamics, heat must spontaneously flow from hot to cold systems. In quantum thermodynamics, the same law applies when considering multipartite product thermal states evolving unitarily. If initial correlations are present, anomalous heat flow can happen, temporarily making cold thermal states colder and hot thermal states hotter. Such an effect can happen due to entanglement but also because of classical randomness, hence lacking a direct connection with nonclassicality. In this work, we introduce scenarios in which anomalous heat flow *does* have a direct link to nonclassicality, defined as the failure of noncontextual models to explain experimental data. We start by extending known noncontextuality inequalities to a setup in which sequential transformations are considered. We then show a class of quantum prepare-transform-measure protocols, characterized by a time interval $(0, \tau_c)$ for a given critical time τ_c , where anomalous heat flow happens only if a noncontextuality inequality is violated. We also analyze a recent experiment from Micadei *et al.* [Nat. Commun. 10, 2456 (2019)] and find the critical time τ_c based on the authors’ experimental parameters. We conclude by investigating heat flow in the evolution of two-qudit systems, showing that our findings are not an artifact of using two-qubit systems.

DOI: [10.1103/f68k-cjx4](https://doi.org/10.1103/f68k-cjx4)

I. INTRODUCTION

The second law of thermodynamics forbids heat to flow from a colder system to a warmer system without consuming any resource, assuming that the systems are isolated [1,2]. The direction of heat flow suggests the notion of an “arrow of time” [3]. If one takes the heat-flow direction and the arrow of time as equivalent phenomena, given two initially correlated interacting systems, a local observer of one of these two systems could have their arrow of time deceived [4,5]. For instance, the heat exchanged (in the

absence of work) between these two systems can initially have a “backflow,” i.e., heat can flow from the colder to the warmer system. Such heat flows are said to be *anomalous* [6]. Crucially, this is *not* a violation of Clausius’s formulation of the second law of thermodynamics, since correlations are *consumed* for such anomalies to happen. Their consumption is responsible for the possibility of a reversal in the heat flow.

An attempt to visualize what is happening is to imagine a “Maxwell demon” [7–9] that has knowledge about the correlations between two physical systems and somehow wants to use this information to perform a thermodynamic task. The demon can use these correlations to make heat flow from the cold to the hot system. This thought experiment was first discussed by Lloyd [10] and then concretely investigated for the case of pure quantum states by Partovi [11], and refined by others [4,12–19]. Recently, the prediction of heat-flow anomaly in quantum theory has been tested [20].

Interestingly, there is no no-go result forbidding the demon (possessing knowledge of initial correlations) to cause anomalous heat flow in microscopic models

*Contact author: naim.comar@usp.br

†Contact author: [cuis@if.usp.br](mailto:cius@if.usp.br)

‡Contact author: luisf@usp.br

§Contact author: rafael.wagner@inl.int

||Contact author: bamaral@if.usp.br

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

described by *classical* statistical mechanics. Indeed, as shown in Ref. [15], one can construct examples of heat-flow reversals caused by correlations without quantum coherence with respect to the local energy basis. Notwithstanding, there is a bound on the *amount* of heat-flow anomaly that can be caused by classical correlations. In such cases, it is possible to surpass this bound only in the presence of entanglement between the initial states [4], this being known as *strong heat backflow*. These results show that, *in general*, anomalous heat flow alone *cannot* indicate a departure from classical explanations, in which case quantifying the anomaly is necessary to separate classical and nonclassical heat transfer.

In this work, we show that experimental scenarios exist in which any anomalous heat flow *does* indicate the failure of a classical explanation. Our working definition of classicality will be the existence of a generalized non-contextual model capable of reproducing the observed data [21]. This notion of classicality is well-motivated conceptually [22]: it emerges under quantum Darwinist dynamics [23] and subsumes other notions of classicality such as Kochen-Specker noncontextuality [24–26], ordinary classical mechanics [27, Sup. Mat. A], or non-negative quasiprobability representations [28–30]. The failure of noncontextual models to explain data can be robustly analyzed [31–34], experimentally tested [35–37], and quantified [38–40].

One of the most useful aspects of generalized non-contextuality is that, even though noncontextual models provide an intuitive classical understanding of (fragments of) physical theories, they can reproduce counterintuitive phenomena such as no-cloning [41], teleportation [41,42], the impossibility of discriminating nonorthogonal states [41], and some single-photon Mach-Zehnder coherent interference patterns [43]. Even rich subtheories of quantum theory can be framed in terms of noncontextual models such as Gaussian quantum mechanics [44] or odd-dimensional (single or multisystem) stabilizer subtheories [45]. The failure of generalized noncontextuality can therefore be considered a *stringent criterion* for nonclassicality. This failure is a strong and rigorous indicator that the system exhibits nonclassical behavior. Moreover, this criterion is considered stringent because contextuality is a clear, broadly applicable, and robust property of models, strongly distinctive from classical behavior.

In this work, we show that anomalous heat flow implies a violation of a generalized noncontextuality inequality in an important class of experimentally meaningful scenarios. This is an indication that contextuality can be further explored as a resource for quantum thermodynamics [46–54]. In our scenario (see Fig. 1), initial bipartite two-qubit states ρ , interacting via an (energy-preserving) unitary $U(t) = e^{-iHt}$ during a time interval $0 < t < \tau_c$, can cause heat backflow (or increase the normal heat flow) *only* when generalized noncontextual models cannot explain the

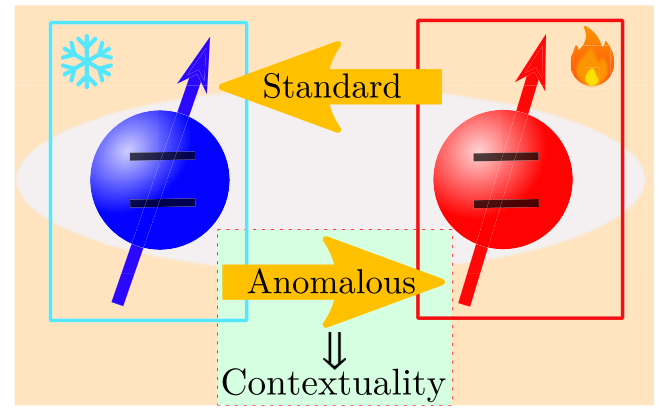


FIG. 1. A sketch of our main result. A “hot” qubit is at temperature T_A and a “cold” qubit is at temperature T_B , i.e., $T_A > T_B$. Each system individually has an associated Hamiltonian, H_A and H_B . We find that for energy-preserving two-qubit interactions $U_I(t)$, i.e., satisfying $[U_I(t), H_A + H_B] = 0$, any anomalous heat flow within a certain interval $0 < t < \tau_c$ leads to generalized contextuality. Our results are not restricted to this simplest case of qubit systems, as we also show.

statistics witnessing this property. We call τ_c the *critical* time. For times $t \geq \tau_c$ outside such intervals, the anomaly does not necessarily imply the failure of generalized non-contextual models to reproduce the data.

To showcase the practical relevance of our findings, we apply our results to the experiment performed in Ref. [20]. We also demonstrate similar results for the partial SWAP interaction between two *qudits* with dimension $d \geq 2$, showcasing that our findings do not depend on the specific Hilbert-space dimension of the physical systems.

Central for demonstrating our results is the work of Ref. [27]. We extend the authors’ findings from a specific transformation process T to one in which sequentially composed transformation processes $T' \circ T$ are considered. We then find that noncontextual models, restricted to obey certain operational equivalences, must be bounded by a noncontextuality inequality that we construct. Our inequality recovers the one found in Ref. [27] as a special case. We believe that our analysis of a concrete sequential scenario will have independent interest in the program of finding noncontextuality inequalities that take into consideration the role of transformation contextuality.

This work is structured as follows. In Secs. II A and II B, we describe the relevant quantum thermodynamic quantities that we investigate and the possibility of reversing the usual direction of heat flow caused by initial correlations. In Sec. II C, we review the concept of generalized noncontextuality and the methods developed by Ref. [27] that are important to us. We present our main results in Sec. III. We start by presenting our Theorem 2, in which the method of Ref. [27] is extended to consider the composition of transformations. In Sec. III A, we show that

noncontextuality inequalities apply to a broad scenario of two interacting qubits (Theorem 3) and that any anomalous heat flow, for dynamical evolutions inside a critical interval $0 < t < \tau_c$, witnesses quantum contextuality. We then use our results to analyze the experiment of Ref. [20] and find an approximate value of τ_c for that experiment using available parameters. In Sec. III B, we consider an interaction given by a partial SWAP unitary and observe the violation of the noncontextuality inequalities for the example of heat flow between two-qudit systems. In Sec. IV, we review our results and make our final remarks. Unless stated otherwise, the dynamics will be evaluated in the interaction picture.

II. BACKGROUND

A. Average heat of local systems during an energy-conserving interaction

Consider a quantum model described by the total Hamiltonian

$$H = H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B + H_I, \quad (1)$$

where H_A and H_B denote the local Hamiltonian operators for quantum systems $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, whereas H_I is the interaction Hamiltonian operator governing the interaction between the two systems. When it is clear from our discussion, we write $H_A \equiv H_A \otimes \mathbb{1}_B$ and $H_B \equiv \mathbb{1}_A \otimes H_B$. Additionally, we suppose that the sum of the average of the free Hamiltonians is conserved during the evolution of the system [55]. Therefore,

$$[U_I(t), H_A + H_B] = 0, \quad (2)$$

where $U_I(t)$ is the unitary time-evolution operator in the interaction picture and thus is given by $U_I(t) = e^{-itH_I}$, setting $\hbar = 1$.

We will only consider initially correlated quantum systems $\mathcal{H}_A$ and $\mathcal{H}_B$, in a global state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, such that each local state ρ_i ($i = A, B$) is a *thermal state* (aka *Gibbs states*). Here, $\mathcal{D}(\mathcal{H})$ denotes the set of all positive trace-1 bounded operators acting on $\mathcal{H}$. This means that

$$\rho_i = \text{Tr}_{\setminus\{i\}} \{\rho\} = \rho_i^{\text{th}} \equiv \frac{e^{-\beta_i H_i}}{Z_i}, \quad (3)$$

where H_i is the local Hamiltonian of $i = A, B$, $\beta_i = 1/T_i$ is the inverse temperature of i with Boltzmann constant $k_B = 1$, and Z_i is the partition function of each thermal state, $Z_i = \text{Tr} \{e^{-\beta_i H_i}\}$.

With the above assumptions, where the Hamiltonian of the joint system is time independent, we can define the average heat flow of the quantum system $\mathcal{H}_A$ as the total energy that it exchanges during the evolution (for the use

of this definition, see Refs. [4, 12–17, 20, 56])

$$\begin{aligned} \langle \mathcal{Q}_A \rangle &\equiv \langle U^\dagger(t) H_A U(t) - H_A \rangle \\ &= \text{Tr} \{ \rho (U^\dagger(t) H_A U(t) - H_A) \}. \end{aligned} \quad (4)$$

Note that, with this definition, a positive $\langle \mathcal{Q}_A \rangle$ means that $\mathcal{H}_A$ receives heat. Similarly, we can define the average heat $\langle \mathcal{Q}_B \rangle$ of the quantum system, $\mathcal{H}_B$. From energy conservation [Eq. (2)], we must have $\langle \mathcal{Q}_A \rangle = -\langle \mathcal{Q}_B \rangle$.

B. A modified Clausius inequality

The dynamics described above imply that the possible average heat flow must satisfy certain inequalities, given the initial state $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. If there are no initial correlations between the systems $\mathcal{H}_A$ and $\mathcal{H}_B$, so that $\rho = \rho_A^{\text{th}} \otimes \rho_B^{\text{th}}$, an immediate consequence is that (see, e.g., Ref. [56, Eq. (15)])

$$(\beta_A - \beta_B) \langle \mathcal{Q}_A \rangle \geq 0, \quad (5)$$

which is equivalent to Clausius's statement of the second law [1, 56–58] for the scenario that we are considering. Equation (5) holds because, for such a case, the left-hand side equals the entropy production, which must always be non-negative.

However, if there are initial correlations between the systems, they can be consumed for thermodynamic tasks. Consequently, the above inequality must be modified [4, 13, 15]. For the kind of dynamics in which we are interested, just introduced in Sec. II A, the inequality becomes

$$(\beta_A - \beta_B) \langle \mathcal{Q}_A \rangle \geq \Delta \mathcal{I}(A : B), \quad (6)$$

where

$$\Delta \mathcal{I}(A : B) = \mathcal{I}_{U\rho U^\dagger}(A : B) - \mathcal{I}_\rho(A : B) \quad (7)$$

is the variation of the mutual information

$$\mathcal{I}_\rho(A : B) = S(\rho_A) + S(\rho_B) - S(\rho), \quad (8)$$

before and after the evolution, and now $\rho \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is any state for which the local states $\rho_i = \rho_i^{\text{th}}$ are Gibbs states. Albeit a well-known fact, we present a detailed derivation of the above in Appendix A.

The inequality in Eq. (5) can be derived from that in Eq. (6) where the parties are initially uncorrelated, since $\Delta \mathcal{I}(A : B)$ is always non-negative in this case. Therefore, as for the heat-flow direction, both inequalities have the same meaning. For the case in which the quantum system $\mathcal{H}_A$ is colder than quantum system $\mathcal{H}_B$, we *must* have $\langle \mathcal{Q}_A \rangle \geq 0$ and heat flows from the hot system to the cold system, as ordinarily expected. However, for initially correlated systems, the variation of the mutual information

can decrease $\Delta\mathcal{I}(A : B) < 0$, allowing for the possibility of *anomalous* heat flow $\langle Q_A \rangle < 0$. This “modified Clausius inequality” includes the consumption of correlations causing heat-flow inversion and still respects the second law of thermodynamics [4,13,15] (something that becomes evident when written in terms of entropy production [56]). Note that if we had assumed quantum system $\mathcal{H}_A$ to be hotter, then the standard heat flow would be described by $\langle Q_A \rangle \leq 0$, i.e., system $\mathcal{H}_A$ would lose heat and the anomaly would be characterized by $\langle Q_A \rangle > 0$, indicating that system $\mathcal{H}_A$ was receiving heat even though it was the hotter system (this will be the case considered in Sec. III A 3).

C. Generalized noncontextuality

Generalized noncontextuality is a constraint on *ontological models* [59] that attempt to explain empirical data predicted by an *operational probabilistic theory* (or fragments thereof) [21,29,60]. Empirical data is obtained by acting on some system with a preparation procedure P , followed by a transformation T and measurement M . Given that an outcome k is obtained, the data are described statistically by some conditional probability distribution $p(k|M, T, P)$. Each procedure, P , M , and T , is defined by a set of laboratory instructions to be performed. They follow a causal order (given two procedures T_1 and T_2 composed in sequence, denoted by $T_2 \circ T_1$, only the first can causally influence the second) and they respect certain laws with regard to which procedures should be applied to which systems (if a transformation T_1 , transforms a system A to some other system A' and a transformation T_2 transforms a system B to some other system B' , they cannot be sequentially composed as $T_2 \circ T_1$ unless $A' = B$; the theory needs to set rules to account for physical systems of different types) [29,61].

Different procedures are associated with different laboratory prescriptions, but they can still yield the same data regardless of any possible operation in that theory. If this happens, these procedures lead to the same possible inferences that can be taken from the data. We then say that the two procedures are *operationally equivalent* [21] or also *inferentially equivalent* [33,61]. Formally, this defines an equivalence relation on the set of procedures in the theory: take any two transformation procedures $T_1, T_2 \in \mathcal{T}$ from the set of all possible transformations $\mathcal{T}$. These are said to be equivalent, and denoted $T_1 \simeq T_2$ if, for all conceivable preparation procedures $P \in \mathcal{P}$ and all conceivable measurements procedures $M \in \mathcal{M}$ having outcomes $k \in \mathcal{K}$,

$$p(k|M, T_1, P) = p(k|M, T_2, P). \quad (9)$$

We term each pair $k|M$ to be a measurement effect. Similar definitions hold for equivalent preparation procedures $P_1 \simeq P_2$ and equivalent effects $k_1|M_1 \simeq k_2|M_2$.

An ontological model [59] of an operational probabilistic theory attempts to explain the predictions of that theory in the following way. Such a model prescribes a (measurable) space Λ of physical variables $\lambda \in \Lambda$ and assigns probability measures $\mu_P(\lambda)$ for each preparation procedure, stochastic matrices $\Gamma_T(\lambda'|\lambda)$ for each transformation procedure, and response functions $\xi_{k|M}(\lambda)$ for each measurement effect $k|M$ such that they recover the empirical predictions of the operational theory from

$$p(k|M, T, P) = \int_{\Lambda} \int_{\Lambda} \xi_{k|M}(\lambda') \Gamma_T(\lambda'|\lambda) d\mu_P(\lambda). \quad (10)$$

An ontological model of an operational theory is said to satisfy the principle of generalized noncontextuality if it explains the operational indistinguishability of different procedures $T_1 \simeq T_2$ by formally imposing that their counterparts in the model Γ_{T_1} and Γ_{T_2} are *equal*, i.e.,

$$T_1 \simeq T_2 \implies \Gamma_{T_1} = \Gamma_{T_2}, \quad (11)$$

and similarly for preparation procedures and measurement effects. When there exists no generalized noncontextual model that can reproduce the data from an operational probabilistic theory, we will refer to this theory as contextual. Quantum theory, viewed as an operational theory, is contextual. This can be shown from no-go theorems [21,62] or in a quantifiable manner via the violation of generalized noncontextuality inequalities [63–65].

In quantum theory, any quantum channel [66,67] $\mathcal{E} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ defines an equivalence class of all possible physically implementable operational procedures that are described by the channel $\mathcal{E}$. $\mathcal{B}(\mathcal{H})$ denotes the bounded operators acting on $\mathcal{H}$. We write this as $\mathcal{E} = [T_{\mathcal{E}}]$ to represent the fact that there can exist many different laboratory procedures T described in quantum theory by the same channel $\mathcal{E}$ such that $T \simeq T_{\mathcal{E}}$, implying that $T \in [T_{\mathcal{E}}]$. Two operationally equivalent procedures then satisfy the fact that

$$T_1 \simeq T_2 \implies \mathcal{E}_1 = \mathcal{E}_2 \quad (12)$$

and the converse holds as well, but now for every element of the class represented by the quantum mechanical operators

$$\mathcal{E}_1 = \mathcal{E}_2 \implies T_1 \simeq T_2, \quad \forall T_1 \in [T_{\mathcal{E}_1}], T_2 \in [T_{\mathcal{E}_2}]. \quad (13)$$

Noncontextuality inequalities can be used to bound the ability of a noncontextual model to explain experimental tasks of interest, success rates of a given protocol, or other figures of merit. Commonly, such bounds are investigated in a prepare-and-measure setup, where the role of transformations is either not considered [68–72] or is merely used to define novel preparation procedures [23,73,74]. In

other words, most scenarios that *do* consider the role of transformations do so by using transformations to define novel preparations, in the sense that if we have preparations $P \in \mathcal{P}$ and a finite set of transformations $T_1, \dots, T_n$, then we only analyze the contextuality of the preparation procedures defined by new preparations $T_i(P)$, or some compositions $T_i \circ T_j(P)$, and so on. The role that transformation contextuality can play in a scenario *without* viewing these as defining novel preparations has so far been considered only in a handful of scenarios: to investigate the nonclassicality of stabilizer subtheory [45,75], of weak values resulting from weak measurements [76,77], and also of quantum thermodynamics of linear response [27].

1. Noncontextuality inequalities for the average of theory-independent observables

In any operational theory, there is a specific transformation procedure denoted T_{id} that denotes the “identity” transformation. Its action is operationally characterized by

$$p(k | M, T_{\text{id}}, P) = p(k | M, P) \quad (14)$$

for all possible preparation procedures $P \in \mathcal{P}$ and all possible measurement effects $k | M \in \mathcal{K} \times \mathcal{M}$.

In a theory-independent setup, we estimate an observable $\mathcal{A}$ by assigning some values a_k to outcomes k obtained once a measurement M has been performed. We will then define a *theory-independent observable*

$$\mathcal{A} := \{(a_k, k | M)\}_k \quad (15)$$

to be the *finite* family of pairs of values a_k and effects $k | M$, for a given measurement procedure M . Hence, we can *define* the expectation value at an initial instant (where we have not transformed the system of interest) given any fixed preparation $P \in \mathcal{P}$ via

$$\langle \mathcal{A}(0) \rangle := \sum_k a_k p(k | M, P) \quad (16)$$

and the expectation value once a transformation has taken place as

$$\langle \mathcal{A}(t) \rangle := \sum_k a_k p(k | M, T_t, P). \quad (17)$$

Here, t is merely a (suggestive) label for the transformation procedure T_t . Also, in all the discussion that follows, we consider, without loss of generality, that $a_k > 0$ for all k . The expectation for the variation between $\mathcal{A}(0)$ and $\mathcal{A}(t)$, given that preparation P has been performed and that

transformation T_t has preceded it, is then given by

$$\langle \Delta \mathcal{A} \rangle_{P,t} := \sum_k a_k (p(k | M, T_t, P) - p(k | M, P)). \quad (18)$$

When it is clear which P and T_t are being considered, we simplify the notation $\langle \Delta \mathcal{A} \rangle_{P,t} = \langle \Delta \mathcal{A} \rangle$. In Ref. [27], it has been shown that generalized noncontextual models that attempt to reproduce the (theory-independent) averages of observables defined by Eq. (18) must satisfy certain inequalities that can be violated by quantum dynamics approximated by linear-response perturbation theory. To be more specific, by linear response we mean that the evolution (in the interaction picture) is generated by a Hamiltonian interaction $H_I(t')$ with a weak strength parameter g , given in Eq. (19) below, i.e., satisfying $g \ll 1$. In this way, the unitary evolution will have the form

$$U_I(t) = \mathbb{1} - ig \int_0^t H_I(t') dt' + O(g^2). \quad (19)$$

In quantum theory, the term “observable” is usually used to denote Hermitian matrices or, more generally, self-adjoint operators $A = A^\dagger$ in $\mathcal{B}(\mathcal{H})$. In this case, the theory-dependent quantum mechanical averages are given by (in the interaction picture)

$$\langle \Delta A \rangle = \text{Tr}\{A(t)\rho(t)\} - \text{Tr}\{A(t)\rho\}, \quad (20)$$

with respect to some initial state $\rho \in \mathcal{D}(\mathcal{H})$ and where, if we are focusing on the linear-response regime, the interaction unitary is taken to be given by Eq. (19). Note that in quantum theory, the values a_k are the *eigenvalues* of the observable A .

An ontological model should explain the statistics of Eq. (18) via Eq. (10), i.e.,

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle = \sum_k a_k \left[\int_{\Lambda} \int_{\Lambda} \xi_{k|M}(\lambda') \Gamma_{T_t}(\lambda' | \lambda) d\mu_P(\lambda) \right. \\ \left. - \int_{\Lambda} \xi_{k|M}(\lambda) d\mu_P(\lambda) \right]. \end{aligned} \quad (21)$$

In Ref. [27], a bound has been proved for averages described in terms of Eq. (21), when the ontological models are generalized noncontextual, that we improve upon slightly in the following.

Theorem 1. Suppose that the following operational equivalence is satisfied:

$$\alpha T_t + (1 - \alpha) T_t^* \simeq (1 - p_d) T_{\text{id}} + p_d T', \quad (22)$$

where T_t , T_t^* , and T' are transformation procedures, T_{id} is the identity procedure, and $0 \leq p_d \leq 1, 0 < \alpha \leq 1$. Let

$\mathcal{A}$ be an observable. Then, any noncontextual ontological model for the variation of its average, defined by Eq. (18) where $a_k > 0$ for all k , must satisfy

$$-\frac{a_{\max} p_d}{\alpha^2} \leq \langle \Delta \mathcal{A} \rangle \leq \frac{p_d a_{\max}}{\alpha}. \quad (23)$$

Above, $a_{\max} := \max\{a_k\}_k$ is the largest value associated with the observable $\mathcal{A}$.

The proof follows the same reasoning as given in Ref. [27], which we will present in detail for the proof of Theorem 2, and recovers the results from Ref. [27] for the case $\alpha = 1/2$. The condition defined by Eq. (22) is called the *stochastic reversibility* and can be understood as tossing a coin to decide if we transform the system using T_t (with probability α) or T_t^* (with probability $1 - \alpha$). After many realizations, the effective transformation obtained is operationally indistinguishable from doing nothing with the system with some probability $1 - p_d$ and doing some other operation with probability p_d . The probability p_d is interpreted as a “probability of disturbance” from the ideal stochastic reversibility with $p_d = 0$. For example, if we consider T_t^* as inverse evolution and $\alpha = 1/2$, stochastic reversibility means that the average evolution after making an equal ensemble of forward and backward evolutions is very close to not having any evolution in the system. We note that T_t , T_t^* , and T' are *in principle* any triplet of transformations satisfying the operational constraint of stochastic reversibility. We assume no other constraint on such transformations. We also note that each transformation separately (at this operational level) plays no fundamental role and that it is only the relation between these three operations (together with T_{id}) that is of relevance for the characterization of the scenario. Interestingly, in Ref. [27], a method has been suggested to certify if a linear response unitary satisfies

$$\frac{1}{2}\mathcal{U}_t + \frac{1}{2}\mathcal{U}_t^\dagger = (1 - p_d)\text{id} + p_d\mathcal{C}, \quad (24)$$

where $\mathcal{U}_t(\cdot) := U(t)(\cdot)U(t)^\dagger$, with $U(t)$ given by Eq. (19), $\text{id}(X) = X$ is the identity channel, and $\mathcal{C}$ is some quantum channel.

As a final remark, we would like to note that the inequality in Eq. (23) is valid for any fragment of an operational probabilistic theory having preparation procedures $\{P_i\}_i \subseteq \mathcal{P}$, effects $\{k|M\}_k \subseteq \mathcal{M} \times \mathcal{K}$ and at least the transformations $\{T_t, T_t^*, T', T_{\text{id}}\}_t \subseteq \mathcal{T}$, where the transformations satisfy Eq. (22). Here, t is merely a label for the transformations in this fragment. In any such fragment, for each choice of T_t and P , the inequality $-p_d a_{\max}/\alpha^2 \leq \langle \Delta \mathcal{A} \rangle_{P,t} \leq p_d a_{\max}/\alpha$ is a valid noncontextuality inequality.

III. RESULTS

We start our results by generalizing the main theorem of Ref. [27] (Theorem 1, for $\alpha = 1/2$) to the case in which we have two sequential transformations T_1 and T_2 , and under the assumption that both satisfy operational equivalences. Our main goal is to obtain a characterization of a scenario in which the presence of contextuality is due to anomalous heat flow. It is to that end that we make the aforementioned generalization, stated below in Theorem 2.

Theorem 2. Let T_t be a sequential transformation given in terms of other two transformation procedures T_{t_1} and T_{t_2} , i.e., $T_t = T_{t_2} \circ T_{t_1}$. Moreover, let T_t satisfy the operational equivalences

$$\frac{1}{2}T_t + \frac{1}{2}T_t^* \simeq (1 - p_{d_1})T_{t_2} + p_{d_1}T'_{t_1}, \quad (25)$$

where $0 \leq p_{d_1} \leq 1$, and also

$$\frac{1}{2}T_{t_2} + \frac{1}{2}T_{t_2}^* \simeq (1 - p_{d_2})T_{\text{id}} + p_{d_2}T'_{t_2}, \quad (26)$$

where $0 \leq p_{d_2} \leq 1$. Then, any noncontextual ontological model for the average of an observable $\mathcal{A}$ must be bounded by

$$-4a_{\max}b_- \leq \langle \Delta \mathcal{A} \rangle \leq 2a_{\max}b_+, \quad (27)$$

where $b_- := p_{d_1} + 3p_{d_2} - 3p_{d_1}p_{d_2}$ and $b_+ := p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}$, and $a_{\max} := \max\{a_k\}_k$ is the largest value associated with the observable $\mathcal{A}$.

We prove this theorem in Appendix B. Note that $b_-, b_+ \geq 0$ for all $0 \leq p_{d_1}, p_{d_2} \leq 1$. It is worth noting that this result is not limited to discrete systems; as demonstrated in Appendix B, our framework extends to scenarios involving a continuous set of ontic states, opening up avenues for future application in continuous-variable systems. Consequently, this could be a theme for future research. This theorem will be instrumental for cases in which the transformation describing the evolution of the system is not capable of satisfying Eq. (22) but can be described as a composition of transformations satisfying this equation. Note that we recover the theorem of Ref. [27] (or Theorem 1, for $\alpha = 1/2$) when $p_{d_2} = 0$ and also when $T_{t_2} \simeq T_{t_2}^* \simeq T_{\text{id}}$. Also, when convenient, we can write the inequality in Eq. (27) more compactly as $|\langle \Delta \mathcal{A} \rangle| \leq 4a_{\max}b_-$, since $b_- \geq b_+$. Relevantly, t can be any label $t = (t_1, t_2)$ for which $\langle \mathcal{A}(t) \rangle = \sum_k a_k p(k|M, T_{t_2} \circ T_{t_1}, P)$ and we do not assume operational equivalences to be satisfied by T_{t_1} .

Similarly to Theorem 1, our Theorem 2 does not depend on the fact that the operations T_t, T_t^* are the inverses of one another; they work for any set of transformations satisfying these operational equivalences. Also, this result is

not only valid under thermodynamic considerations: it will hold for any observable from which its theory-independent average can be described by an equation of the form given in Eq. (18).

A. Two interacting qubits

Let us consider *Zeeman Hamiltonians* to be the class of single-qubit Hamiltonians defined as linear combinations of the identity operator $\mathbb{1}$ and the Pauli matrix σ_z . We now show that an interesting class of quantum interactions between two-qubit systems must satisfy the operational equivalences defined by stochastic reversibility. The following theorem is made to obtain a practical noncontextuality inequality between *any* two qubits that interact via a unitary evolution preserving their total energy, and hence can be useful and a large range of applications. Here, we shall use this for our main question, which is the influence of contextuality in the heat-flow inversion.

Theorem 3. Let $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ be the Hilbert space describing a two-qubit system. Consider the evolution $U(t) = e^{-itH}$, with H given by Eq. (1), with H_A and H_B being Zeeman Hamiltonians. Suppose also that the interaction preserves energy, i.e., $[H_I, H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B] = 0$. Assuming an interaction-picture representation of the dynamics, for every fixed instant t :

- (1) Nonresonant case ($H_A \neq H_B$). There exists a quantum channel $\mathcal{C}$ such that

$$\frac{1}{2}\mathcal{U}_I + \frac{1}{2}\mathcal{U}_I^\dagger = (1 - p_d)\text{id} + p_d\mathcal{C}, \quad (28)$$

where $\mathcal{U}_I(\cdot) = U_I(t)(\cdot)U_I(t)^\dagger$, $U_I(t) = e^{-itH_I}$, and $0 \leq p_d \leq 1$.

- (2) Resonant case ($H_A = H_B$). The unitary evolution $U_I(t)$ can be written as a composition of two other unitaries $U_I(t) = U_2 \circ U_1$ and there exist quantum channels $\mathcal{C}_1, \mathcal{C}_2$ such that

$$\frac{1}{2}\mathcal{U}_1 + \frac{1}{2}\mathcal{U}_1^\dagger = (1 - p_{d_1})\text{id} + p_{d_1}\mathcal{C}_1, \quad (29)$$

$$\frac{1}{2}\mathcal{U}_2 + \frac{1}{2}\mathcal{U}_2^\dagger = (1 - p_{d_2})\text{id} + p_{d_2}\mathcal{C}_2, \quad (30)$$

where $\mathcal{U}_i(\cdot) = U_i(\cdot)U_i^\dagger$, and $0 \leq p_{d_i} \leq 1$ for $i = 1, 2$.

We prove this result in Appendix C. We now comment on some aspects of this result that can be relevant. First, we *do not* need to assume a linear-response regime for the validity of the stochastic reversibility equations. The values of p_d, p_{d_1} , and p_{d_2} and the quantum channels $\mathcal{C}, \mathcal{C}_1$, and $\mathcal{C}_2$ always exist. This immediately implies that such a dynamical evolution always satisfies the operational assumptions necessary for Theorems 1 and 2.

We can use Theorem 3, together with Theorems 1 and 2, to draw the following conclusion. Let $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ and $U(t) = e^{-itH} = U_0(t)U_I(t)$, where $U_0(t) = e^{-it(H_A+H_B)}$ and $U_I(t) = e^{-itH_I}$ are as described in Theorem 3. Then, *any noncontextual ontological model* for a fragment of quantum theory (viewed as an operational theory) described by states $\{\rho\} \subseteq \mathcal{D}(\mathcal{H})$, POVM elements $\{E_k(t)\}_{k,t} = \{U_0(t)^\dagger E_k U_0(t)\}_{k,t} \subseteq \mathcal{B}(\mathcal{H})^+$ and unitary transformations $\{U_I(t)\}_t \subseteq \mathbb{U}(\mathcal{H})$, such that

$$\Delta A_{\text{NC}} := \sum_k a_k (\text{Tr}\{U_I(t)\rho U_I(t)^\dagger E_k(t)\} - \text{Tr}\{\rho E_k(t)\}),$$

must satisfy the inequality

$$-4a_{\text{max}}b_- \leq \Delta A_{\text{NC}} \leq 2a_{\text{max}}b_+, \quad (31)$$

where $b_- = p_{d_1} + 3p_{d_2} - 3p_{d_1}p_{d_2}$ and $b_+ = p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}$, for some $0 \leq p_{d_1}, p_{d_2} \leq 1$ fixed by the transformation $U(t)$ and where $\{a_k\} \subseteq \mathbb{R}^+$ is any finite set of real positive numbers having $a_{\text{max}} := \max\{a_k\}$. Note that Eq. (31) is an inequality for when we assume the validity of quantum theory, while Eq. (27) is valid for any operational theory. The average $\langle \Delta A(t) \rangle$ predicted by quantum theory for a positive observable A is recovered when we let $E_k := \Pi_k^A$ correspond to the projections onto the eigenspace of A , associated with eigenvalue a_k .

To show why the inequality in Eq. (31) holds, for the *nonresonant* case, we use Eq. (28) as the quantum theoretical version of the operational equivalence from Theorem 1, from which we obtain Eq. (31) with $p_{d_2} = 0$. As for the *resonant* case, we can apply $\mathcal{U}_2$ to both sides of Eq. (29) and obtain a quantum theoretical version of the operational equivalence from Theorem 2, since $\mathcal{U}_2$ is linear. This is an important use of Theorem 2, which can be applied in a broad scope of different situations. Indeed, given *any* noncontextual ontological model for a fragment of quantum theory, if two unitary transformations, $\mathcal{U}_1$ and $\mathcal{U}_2$, in this fragment, satisfy Eqs. (29) and (30), respectively, then the above procedure can be carried out to obtain the inequality in Eq. (31).

Let us make a clarification regarding the interaction picture and the results in Theorems 1 and 2. Note that in the interaction picture, when assuming the validity of quantum theory, it is sufficient that the unitary evolution satisfying Eq. (22) is the one related to the *interaction Hamiltonian* (regardless of energy preservation) and one simply evolves the observable with the noninteracting evolution, hence considering $E_k(t)$ instead of E_k . In the theory-independent setting, one can simply define the average observable to be estimated with respect to a new set of measurements $k_t | M_t$ and values a_{k_t} for a fixed label t . This is, however, not necessary for the conclusions following from Theorems 1 and 2 and it is a feature of applying these findings when assuming quantum theory as an operational theory.

Therefore, we have the average

$$\begin{aligned}
 \langle \Delta A(t) \rangle &= \\
 &= \sum_k a_k \left(\text{Tr}\{U_I \rho U_I^\dagger U_0^\dagger \Pi_k^A U_0\} - \text{Tr}\{\rho U_0^\dagger \Pi_k^A U_0\} \right) \\
 &= \text{Tr}(U_I(t) \rho U_I(t)^\dagger A(t)) - \text{Tr}(\rho A(t)) \\
 &= \text{Tr}(\rho(t) A(t)) - \text{Tr}(\rho A(t)). \tag{32}
 \end{aligned}$$

We can see why the inequality in Eq. (31) bounds non-contextual explanations for such fragments of quantum theory. In any such fragment, the dynamics characterizing any transformation procedure represented in quantum theory by the interaction unitary U_I as in Theorem 3 satisfy the operational equivalences from Theorem 2. These equivalences are the requirements for the noncontextuality inequality from Theorem 2 to hold for theory-independent observables given by Eq. (18). When applied to any fragment of quantum theory respecting these equivalences, such observables take the form described by the inequality in Eq. (31). As a final remark, note that in general, p_d , p_{d_1} , and p_{d_2} will depend on the parameters (and the time parameter t) describing the interaction Hamiltonian H_I . Their exact dependence is given in Appendix C.

1. Generalized contextuality without measurement incompatibility

The above calculations allow us to draw a simple yet remarkable conclusion. We can consider a fragment of quantum theory that has: (i) unitary evolutions $U(t)$ and $U(t)^\dagger$, together with the identity map and some other channel $\mathcal{C}$, for a fixed time t ; (ii) a single measurement protocol, characterized by the effects $\{\Pi_k(t)\}_k$; and (iii) a tomographically complete set of states ρ for $\mathcal{H}$. Even though such a fragment obviously has no *measurement incompatibility* (as there is, in fact, a single measurement), such a fragment may lead to a violation of the above inequality (as we will show later). Indeed, that proofs of the failure of generalized noncontextuality can be found with a single measurement is a theoretical prediction from Ref. [78]. Our construction provides a concrete fragment of quantum mechanics in which this can be verified by the violation of a noncontextuality inequality, obtained from Theorem 2.

2. Contextuality as a necessary condition for anomalous heat flow

We will now show that contextuality allows for anomalous quantum heat transfer beyond what *any* noncontextual model is capable of. Using Theorem 3, we now study violations of the generalized noncontextuality obtained when there is heat exchanged between two qubits, $\mathcal{H}_A = \mathbb{C}^2 = \mathcal{H}_B$. They are described locally in terms of Zeeman Hamiltonians, which, without loss of generality, we choose to be $H_A = \omega_A/2(\mathbb{1} - \sigma_z)$ and $H_B = \omega_B/2(\mathbb{1} - \sigma_z)$, and

they interact via an energy-preserving unitary. From the definition of Eq. (4), we see that if we choose $A := H_A \otimes \mathbb{1}_B$, then

$$\langle \mathcal{Q}_A \rangle = \langle A(t) - A(0) \rangle = \langle H_A \otimes \mathbb{1}_B(t) - H_A \otimes \mathbb{1}_B \rangle \tag{33}$$

and we can investigate the average heat flow from quantum system $\mathcal{H}_B$ to system $\mathcal{H}_A$ (or vice versa). From Theorem 3, the noncontextuality inequalities (Eqs. (23) and (27)) bound the heat received by A explainable by noncontextual models. The inequality in Eq. (23) bounds the nonresonant case, while that in Eq. (27) bounds the resonant case.

For the nonresonant case $\omega_A \neq \omega_B$, the study is trivial, since no heat is transferred. This happens because by a direct calculation, for all times,

$$\langle \mathcal{Q}_A \rangle = 0. \tag{34}$$

It is simple to see that this holds because the most general nonresonant interaction Hamiltonian $H_{I_{nr}}$ in this case that is capable of satisfying $[H_{I_{nr}}, H_A + H_B] = 0$ must have the form $H_{I_{nr}} = g(|01\rangle\langle 01| + |10\rangle\langle 10|)$ (see Appendix C).

For the resonant case $\omega_A = \omega_B = \omega$ and

$$H_A = \frac{\omega}{2}(\mathbb{1} - \sigma_z^A). \tag{35}$$

We now consider the most general case of an initial two-qubit density matrix (in the eigenbasis of $\sigma_z^A \otimes \sigma_z^B$), written as

$$\rho = \begin{pmatrix} p_{00} & v_1^* & v_2^* & \gamma^* \\ v_1 & p_{01} & \eta e^{i\xi} & v_3^* \\ v_2 & \eta e^{-i\xi} & p_{10} & v_4^* \\ \gamma & v_3 & v_4 & p_{11} \end{pmatrix}, \tag{36}$$

where $0 \leq p_{00}, p_{01}, p_{10}, p_{11} \leq 1$, η , and ξ are real numbers and the remaining parameters are complex numbers, constrained to satisfy the fact that ρ is positive semidefinite and $p_{00} + p_{01} + p_{10} + p_{11} = 1$. Note that we choose a commonly used basis representation to make our calculations but our results are basis-independent, as is usually true for proofs of the failure of noncontextuality [64]. Here, v^* denotes the complex conjugate of the complex number v . The most general energy-conserving unitary interaction for the resonant case is given by $U_{I_r}(t) = e^{-iH_{I_r}t}$, where H_{I_r} is a Hamiltonian given by

$$H_{I_r} = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & e^{i\theta} & 0 \\ 0 & e^{-i\theta} & a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{37}$$

where $g > 0$ and $0 \leq \theta \leq 2\pi$. As demonstrated in Appendix C in the proof of Theorem 3, we find that $U_{I_r} = U_2 \circ U_1$ and that the unitary evolutions characterized by such Hamiltonians satisfy the operational equivalences shown in Theorem 3 for $p_{d_1} = \sin^2[(a-1)gt/2]$ and $p_{d_2} = \sin^2(gt)$. Note that in both cases $p_{d_i} = O(g^2 t^2)$ when $gt \ll 1$.

For this situation, a direct computation of the average heat [Eq. (4)] results in

$$\langle Q_A \rangle = \omega [(p_{01} - p_{10}) \sin^2(gt) + \eta \sin(2gt) \sin(\xi - \theta)]. \quad (38)$$

This equation asserts that, for the case of resonant qubits, transformation contextuality happens for sufficiently small gt if and only if the presence of coherence in the initial density matrix has some effect on the heat flow. This is because, using Theorems 3 and 2 with $A := H_A \otimes \mathbb{1}_B$, the absolute value of the heat must be bounded by a quantity of order $O(g^2 t^2)$, to allow for a noncontextual explanation of the dynamics (the inequality in Eq. (27)). However, the heat contribution caused by coherence, namely, the term $\eta \sin(2gt) \sin(\theta + \xi)$, is $O(gt)$, which can always violate the noncontextual bound for small enough gt . In Ref. [27], this has been called the contextuality of quantum linear response.

For small enough gt , whenever the variation of the observable (in this case, the energy) is $O(gt)$, a quantum violation of the noncontextuality inequality will always exist. More concretely, we can find the range for values gt in which an anomaly must imply quantum contextuality. Since, in an experimental setup, g is normally a fixed parameter, this means that we are interested in finding the interval $0 < t < \tau_c$, for some *critical time* τ_c such that any anomalous heat flow cannot be explained by noncontextual models. Note that in general, $t = 0$ starts with no heat flow and we let $t = \tau_c$ be the regime in which we lose a violation of the noncontextuality inequality in Eq. (27).

While our main focus has been on the earliest critical time τ_c at which contextuality due to anomalous heat flow ceases to be guaranteed, this is by no means the *only* interval in which noncontextuality inequalities can be violated. In the most general setting, our results imply that, for certain evolutions, multiple disjoint time windows exist in which a violation of the noncontextuality inequality signals anomalous heat flow. For concreteness—and to facilitate comparison with experimental data—we then concentrate on this specific first time interval.

If we impose for the initial state ρ that the local states are thermal with inverse temperatures β_A and β_B (see Eq. (3)),

then the most general global state ρ will have the form

$$\rho = \begin{pmatrix} v_0 & v_1^* & v_2^* & \gamma^* \\ v_1 & \frac{1}{Z_A} - v_0 & \eta e^{i\xi} & -v_2^* \\ v_2 & \eta e^{-i\xi} & \frac{1}{Z_B} - v_0 & -v_1^* \\ \gamma & -v_2 & -v_1 & \frac{e^{-\omega\beta_A - \omega\beta_B} - 1}{Z_A Z_B} + v_0 \end{pmatrix}, \quad (39)$$

where v_0 is a real number and $Z_i = 1 + e^{-\omega\beta_i}$, $i = A, B$. This is the case considered in Sec. II A, with two-qubit systems. The average heat [Eq. (38)], after some manipulations, becomes

$$\langle Q_A \rangle = \omega \left\{ \frac{1}{2} \sin^2(gt) \left[\tanh\left(\frac{\omega\beta_A}{2}\right) - \tanh\left(\frac{\omega\beta_B}{2}\right) \right] + \eta \sin(2gt) \sin(\xi - \theta) \right\}. \quad (40)$$

The term $(\omega/2) \sin^2(gt) [\tanh(\omega\beta_A/2) - \tanh(\omega\beta_B/2)]$ in the above equation is of order $O(g^2 t^2)$ and is responsible for the standard heat flow in the absence of correlations. For instance, if $T_A < T_B$, then $\beta_A = 1/T_A > 1/T_B = \beta_B$ and therefore $\tanh(\omega\beta_A/2) > \tanh(\omega\beta_B/2)$, meaning that quantum system $\mathcal{H}_A$ —in this case, the colder system—receives heat as expected.

Concurrently, the term $\omega\eta \sin(2gt) \sin(\xi - \theta)$ has sign and magnitude depending only on the coherence term $\eta e^{i\xi}$ and the interaction parameters g, θ . This allows the latter term to be capable of causing anomalous heat flow or to increase the standard heat flow depending on the type of coherence and interactions. Importantly, this last term is of order $O(gt)$, which (from Theorem 3) implies that, whenever this term is nonzero, there will be a contextuality witness for small enough gt or, equivalently, for *some* interval of time $0 < t < \tau_c$. Since any heat-flow inversion is possible only due to this last term, we conclude that the following holds.

Corollary 1. For any two qubits (having local Zeeman Hamiltonians and associated Gibbs states) interacting via an energy-preserving unitary, anomalous heat flow is possible for $0 < t < \tau_c$ only if noncontextual models fail to explain the data.

This corollary is our main result. It shows a contextuality signature for a fairly broad class of thermodynamic scenarios. It is also important to note that the term v_0 in Eq. (39) carries all the information about the incoherent correlations in the initial density matrix of the two qubits. However, this term has no role in the heat exchanged [Eq. (40)]. This leads us to conclude that *only* coherent correlations and entanglement are responsible for the anomalous

heat flow, for the interacting two-qubit case. This relation between heat-flow anomaly and the aforementioned quantum resources is already known from Ref. [14]. We add the statement that it is both these resources, together with the dependency on gt and the validity of the operational equivalences that we investigate, that make these correlations inexplicable by means of noncontextual ontological models.

In addition, we can point out a relation between the violation of the standard Clausius inequality [Eq. (5)] and the presence of contextuality when $gt \ll 1$. That is, for short times, the same heat-flow term contributing to the anomaly, and hence to the violation to the noncontextuality inequality, is responsible for the initial negativity of the mutual information change in Eq. (6) (see Appendix D). This suggests a direct relationship between contextuality and the emergence of correlation effects in the second law of thermodynamics.

3. Connection with experimental results

We have claimed that for fixed values of g , there are critical times τ_c that can be found, such that, for any $0 < t < \tau_c$, the noncontextual inequality is violated for this class of two-qubit interactions that we have just studied. Small gt bounds the amount of time in which this contextuality witness happens in real experiments and we wish this time interval to be experimentally accessible. Therefore, we now estimate such a critical time τ_c using the parameters from the recent experimental result of Ref. [20].

In this experiment, the two qubits are realized by spin-1/2 systems in a nuclear magnetic resonance (NMR) setup and are resonant with local Hamiltonians in the form of Eq. (35) with $\omega = h\nu_{\text{exp}}$, where $h \approx 4.135 \times 10^{-15} \text{ eV}\cdot\text{Hz}^{-1}$ is the Planck constant and $\nu_{\text{exp}} = 1 \text{ kHz}$. The interaction is set by the following effective interaction Hamiltonian:

$$H_I = \frac{\pi J}{2} (\sigma_x^A \otimes \sigma_y^B - \sigma_y^A \otimes \sigma_x^B), \quad (41)$$

where $J = 215.1 \text{ Hz}$. The initial temperature (in units of energy) of the qubit $\mathcal{H}_A$ is $T_A = 4.3 \text{ peV}$ and of the qubit $\mathcal{H}_B$ is $T_B = 3.66 \text{ peV}$ (recalling that $T_i = \beta_i^{-1}$). Note, therefore, that in the experiment of Ref. [20], $\mathcal{H}_A$ is the *hotter* system, so any anomalous heat flow happens when this system *receives heat*, which in our convention is characterized by $\langle Q_A \rangle > 0$.

Additionally, the experiment considers the initial state in the form of Eq. (39) with $\gamma = \xi = \nu_0 = \nu_1 = \nu_2 = 0$, and $\eta = -0.19$. This situation corresponds to the resonant case with heat given by Eq. (40) and an interaction Hamiltonian having parameters $g = J\pi$, $a = 0$, and $\theta = \pi/2$.

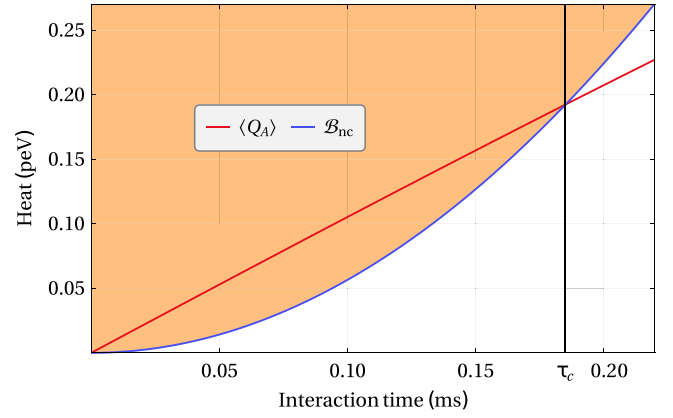


FIG. 2. An anomalous heat flow with $0 < t < \tau_c$ implies a violation of a noncontextuality inequality. The average heat flow predicted by quantum theory (red line) and the noncontextuality bound (blue curve). The noncontextuality bound is given by the inequality in Eq. (27). The region above the curve (orange) corresponds to the values in which the heat transfer *cannot* be described by a noncontextual model. Any anomalous heat flow violates the noncontextuality inequality for $0 < t < \tau_c \approx 1.85 \times 10^{-4} \text{ s}$. In this interval, heat averages predicted by quantum theory are greater than those achievable by noncontextual models. The interaction Hamiltonian is given by Eq. (41) and the initial quantum state is given by Eq. (39). The parameters are as follows: $J = 215.1 \text{ Hz}$, $\omega = 4.135 \times 10^{-12} \text{ eV}$, $T_A = 4.3 \text{ peV}$, $T_B = 3.66 \text{ peV}$, $\gamma = \xi = \nu_0 = \nu_1 = \nu_2 = 0$, $g = J\pi$, $a = 0$, $\theta = \pi/2$, and $\eta = -0.19$, taken from Ref. [20].

Theorem 2 implies that the absolute value of the heat received by the qubit $\mathcal{H}_A$ must be bounded by the function

$$\begin{aligned} \mathcal{B}_{\text{nc}} &:= 2a_{\text{max}}b_+ = 2a_{\text{max}}(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) \\ &= 2h\nu_0[\sin^2(J\pi t) + 2\sin^2(J\pi t/2) \\ &\quad - 2\sin^2(J\pi t)\sin^2(J\pi t/2)] \\ &= 2h\nu_0[1 - \cos^3(J\pi t)], \end{aligned} \quad (42)$$

defining the noncontextual bound. In Fig. 2, we see the region where the heat flow predicted by quantum theory (red curve), given by Eq. (40) with the parameters of the experiment of Ref. [20], violates such a bound for small t (the anomalous heat flow here is positive since $T_A > T_B$; a standard heat flow would be negative). Therefore, we obtain the approximate value for critical time to be $\tau_c \approx 1.85 \times 10^{-4} \text{ s}$.

As mentioned previously, during the reversal of heat flow, the initial correlations are “consumed” and, as a result, the variation of mutual information $\Delta\mathcal{I}(A:B)$ is negative. A clear image for this consumption is presented in Micadei *et al.* [20, Fig. 2] for this experimental case. As we show in Appendix D, this consumption has a negative slope boosted by the same heat-flow term that violates the noncontextuality inequality.

We can also see that, from Fig. 2, when $t \geq \tau_c$, anomalous heat flow can happen, yet without the violation of the noncontextuality inequality (it is important to stress that respecting this noncontextuality inequality is a necessary but not a sufficient condition for noncontextuality). We can also comment that the plot in Fig. 2 is enlarged, so that the values of $\langle \mathcal{Q}_A \rangle$ appear to be a line when they are described by a periodic function. This implies that, in fact, there are various critical time intervals $\Delta\tau_{ij} = (\tau_c^i, \tau_c^j)$ in which the reversal of heat flow witnesses quantum contextuality. Finally, while we have focused on the linear regime due to the experimental parameters, *in principle* different choices of parameters could allow for the anomaly to witness contextuality *beyond* a linear-response regime.

B. Two-interacting-qudit systems: The partial SWAP as a case study

Our main theoretical result from Theorem 2 is not dependent on the dimensionality of the physical system considered. To highlight this fact, we explore similar conclusions as in the two-qubit case, using the methods of Theorems 1 or 2, by finding interactions that respect the stochastic reversibility condition [Eq. (22)] valid to any bipartite system $\mathbb{C}^d \otimes \mathbb{C}^d$ for some dimension d . In what follows, we make statements that are valid for any Hilbert space $\mathcal{H}$ and specify to the case of dynamics governed by the partial SWAP acting on $\mathcal{H} = \mathbb{C}^d \otimes \mathbb{C}^d$. For concreteness, we conclude with an analysis of the case $d = 3$.

1. A simple family of unitary evolutions satisfying stochastic reversibility

We say that a Hermitian operator $K = K^\dagger$ is also *involutory* when $K^2 = \mathbb{1}$. From Stone's theorem, any family of unitary operators $\{U(t)\}_t$ can be written as $U = e^{-itK}$ for some self-adjoint K and for the cases in which K is also involutory, we can show that it satisfies stochastic reversibility.

Proposition 1. Let $\mathcal{H}$ be any Hilbert space. Let K be an involutory self-adjoint operator from $\mathcal{B}(\mathcal{H})$ and $U_K(t) := e^{-igtK}$. Then, $\mathcal{U}_K(\cdot) := U_K(t)(\cdot)U_K(t)^\dagger$ satisfies the stochastic reversibility condition given in Eq. (22), with $\mathcal{C}(\cdot) = K(\cdot)K$, $\alpha = 1/2$, and $p_d = \sin^2(gt)$.

Proof. Note that $\mathcal{C}(X) = KXK^\dagger$ and $K^\dagger K = KK = \mathbb{1}$ and hence $\mathcal{C}$ is in a Kraus representation form, and therefore it is a quantum channel. The validity of stochastic reversibility with this specific p_d follows from Taylor expansion, since K is bounded and involutory. ■

Examples of Hermitian involutory matrices are the SWAP operator, the Pauli matrices, the identity matrix, controlled-NOT (CNOT) gates, conjugations [79], and the Hadamard matrix.

2. Partial SWAP

A unitary that satisfies this condition and energy conservation [Eq. (2)] is the *partial* SWAP, defined as

$$U_{PS}(t) := e^{-igtS}, \quad (43)$$

where S is the SWAP operator [66] and $g > 0$. Since the SWAP operator is described by an involutory matrix, i.e., $S^2 = \mathbb{1}$, Proposition 1 applies to e^{-igtS} , where

$$U_{PS}(t) = \cos(gt)\mathbb{1} - i\sin(gt)S. \quad (44)$$

By direct calculation, we have

$$\frac{1}{2}\mathcal{U}_{PS} + \frac{1}{2}\mathcal{U}_{PS}^\dagger = \cos^2(gt)\text{id} + \sin^2(gt)\mathcal{C}_S, \quad (45)$$

where $\mathcal{U}_{PS}[X] := U_{PS}(t)XU_{PS}^\dagger(t)$ and $\mathcal{U}_{PS}^\dagger[X] = U_{PS}(t)^\dagger X^\dagger U_{PS}(t)$, $p_d = \sin^2(gt)$, and the quantum channel $\mathcal{C}_S[X] = SX S$. Note that the partial SWAP dynamics are valid for any S acting on $\mathbb{C}^d \otimes \mathbb{C}^d$. In what follows, we investigate heat flow in this general case.

3. Heat flow during a partial SWAP of two qudits

The partial SWAP unitary can be useful to obtain non-contextuality inequalities for higher-dimensional systems, since it satisfies the stochastic reversibility condition in any dimension in which S is well defined. An example that may find fruitful applications is any two qudits interacting via a partial SWAP. For these cases, we find conclusions similar to the two interacting qubits of Sec. III A.

Consider two single-qudit quantum systems $\mathcal{H}_A$ and $\mathcal{H}_B$, each with a Hilbert space of dimension $d \geq 2$. The qudit $\mathcal{H}_A$ has a local Hamiltonian

$$H_A = \sum_{k=0}^{d-1} \omega_k |k\rangle_A \langle k|_A, \quad (46)$$

where the ω_k are positive numbers (the so-called Bohr frequencies) and the $|k\rangle_A$ are the eigenvectors of H_A . If the global state of the composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is initially

$$\rho = \sum_{n,m,n',m'} p_{n,m,n',m'} |n,m\rangle \langle n',m'|, \quad (47)$$

where $p_{n,m,n',m'}$ are the density-matrix elements with respect to $\rho > 0$ and $\text{tr}\{\rho\} = 1$, with marginals $p_{n,n'}^A = \sum_m p_{n,m,n',m'}$ and $p_{m,m'}^B = \sum_n p_{n,m,n',m'}$. Note that we have defined $|n,m\rangle := |n\rangle_A \otimes |m\rangle_B$, where $|n\rangle_A$ is an eigenvector of H_A with eigenvalue ω_n and $|m\rangle_B$ is an eigenvector of the local Hamiltonian acting on the Hilbert space of B . Given these conditions, we show (see Appendix E) that the

heat *received* by $\mathcal{H}_A$, after interacting with $\mathcal{H}_B$ via a partial SWAP (given by Eq. (44)) is

$$\langle \mathcal{Q}_A \rangle = \sin^2(gt) \left(\sum_m p_{m,m}^B \omega_m - \sum_n p_{n,n}^A \omega_n \right) - \sin(2gt) \sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n. \quad (48)$$

This equation maintains the same characteristics as the heat exchanged between any two qubits [Eq. (38)] and, consistently, it is equivalent to Eq. (38) for $d = 2$ and $\theta = 0$. The following three aspects are of main importance. First, the first term of the right-hand side of the above equation, namely, $\sin^2(gt) (\sum_m p_{m,m}^B \omega_m - \sum_n p_{n,n}^A \omega_n)$, which depends only on the diagonal terms of the global density matrix (and, more specifically, only on the diagonal terms of the reduced density matrices), is responsible for the standard heat flow. For instance, if the local states are thermal $\rho_A = (e^{-\beta_A H_A} / Z_A)$ and $\rho_B = (e^{-\beta_B H_B} / Z_B)$, for $\beta_{A(B)}$ inverse temperatures, $Z_{A(B)}$ partition functions, and the local Hamiltonians $H_{A(B)} = \sum_{k=0}^{d-1} \omega_k |k\rangle_{A(B)} \langle k|_{A(B)}$, with $\omega_k = k\Delta\omega > 0$ (equally spaced Böhr frequencies), this first term will be

$$\Delta\omega \frac{\sin^2(gt)}{2} [\coth(\beta_B \Delta\omega/2) - d \coth(d\beta_B \Delta\omega/2) - (\coth(\beta_A \Delta\omega/2) - d \coth(d\beta_A \Delta\omega/2))], \quad (49)$$

which is always positive for $\beta_A > \beta_B$, meaning that it contributes positively to the heat received by $\mathcal{H}_A$ when the quantum system $\mathcal{H}_A$ is locally in a colder state than the quantum system $\mathcal{H}_B$. Clearly, this term is also consistent with the case of two qubits (for $d = 2$ and $\theta = 0$, this gives the first term on the right-hand side of Eq. (40)).

Second, the second term on the right-hand side of Eq. (48), namely, $-\sin(2gt) \sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n$, is the term responsible for the possibility of anomalous heat flow and depends only on the coherent terms of the initial density matrix (in the energy eigenbasis).

Third, this second term on the right-hand side of Eq. (48) is of type $\mathcal{O}(gt)$ for small values of gt , while the first term (responsible for the standard heat flow) is of type $\mathcal{O}(g^2 t^2)$. Therefore, this second coherent term is always capable of violating a noncontextuality inequality for a sufficiently small gt . This follows directly from Proposition 1 (which is valid for the partial SWAP unitary) and Theorem 1.

Hence, we have shown that the density-matrix coherence terms can cause an effect on the heat flow exchanged between two qudits, interacting via a partial SWAP, *if and only if* a noncontextuality inequality can be violated. Also, for this case, the anomalous heat flow happens

only if a noncontextuality inequality can be violated. These facts happen in complete analogy with the heat exchanged between two qubits of Sec. III A and show that the same conclusions can be extended for a case of higher dimension.

Let us investigate the violation of the noncontextuality inequality (Theorem 1) using Eq. (48). Let $\omega_{\max} = \max\{\omega_k\}_{k=0,d-1}$. Then, we can find the critical times τ_c^l, τ_c^u , assuming that $\omega_{\max}, g$ are all fixed parameters and $\sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n \neq 0$, given by

$$\tau_c^u = \frac{1}{g} \cot^{-1} \left(\frac{2\omega_{\max} - (\sum_m p_{m,m}^B \omega_m - \sum_n p_{n,n}^A \omega_n)}{2 \sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n} \right), \quad (50)$$

$$\tau_c^l = \frac{1}{g} \cot^{-1} \left(\frac{-4\omega_{\max} - (\sum_m p_{m,m}^B \omega_m - \sum_n p_{n,n}^A \omega_n)}{2 \sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n} \right), \quad (51)$$

where above we have considered τ_c^u the critical time associated with the noncontextuality inequality upper bound, and τ_c^l for the lower bound. These bounds can be used to directly infer the relationship between contextuality and anomalous heat flow. For a concrete example, we will focus on $d = 3$ in what follows.

4. Exemplification: Heat flow during a partial SWAP of two qutrits

The partial SWAP unitary can be useful to obtain noncontextuality inequalities for higher-dimensional systems since it satisfies the stochastic reversibility condition in any dimension in which S is well defined. An intriguing model for examination involves the interplay between two qutrits mediated by a SWAP gate. This framework finds utility in elucidating experimental implementations of quantum gates and quantum engines, while concurrently facilitating extensions to qudit systems [80–82]. Each qutrit manifests solely three perfectly distinguishable states, denoted as $|0\rangle$, $|1\rangle$, and $|2\rangle$, which we write as the canonical basis

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (52)$$

It is ensured that the composite state of the system adheres to the form $|i\rangle \otimes |j\rangle$ ($i, j = 0, 1, 2$). Consequently, the SWAP operator S_3 is characterized by the conditions

$$S_3 |i\rangle \otimes |j\rangle = |j\rangle \otimes |i\rangle, \quad S_3^2 = \mathbb{1}. \quad (53)$$

Thus, the SWAP operator assumes the form

$$S_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (54)$$

and the unitary operator governing the dynamics of the system arising from this interaction is expressed as (see Eq. (44))

$$U_3(t) = e^{-igtS_3} = \cos(gt) \mathbb{1} - i \sin(gt) S_3. \quad (55)$$

As before, g denotes the strength of the interaction. The local Hamiltonian governing a single-qutrit system is expressed as

$$H_A = \omega_0 |0\rangle \langle 0| + \omega_1 |1\rangle \langle 1| + \omega_2 |2\rangle \langle 2|, \quad (56)$$

where $\omega_i \in \mathbb{R}^+$ denotes the energy associated with eigenstate $|i\rangle$, this Hamiltonian being the realization of the qudit Hamiltonian of Eq. (46). Also, we consider both local Hamiltonians to be equal, i.e., $H_A = H_B$. Consequently, a density matrix, with thermal local states [Eq. (3)], describing the composite system while conserving energy, is generically represented as

$$\rho_{3 \otimes 3} = \begin{pmatrix} p_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & p_1 & 0 & \eta_{31} e^{i\theta_{31}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_2 & 0 & 0 & 0 & \eta_{62} e^{i\theta_{62}} & 0 & 0 \\ 0 & \eta_{31} e^{-i\theta_{31}} & 0 & p_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_5 & 0 & \eta_{75} e^{i\theta_{75}} & 0 \\ 0 & 0 & \eta_{62} e^{-i\theta_{62}} & 0 & 0 & 0 & p_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \eta_{75} e^{-i\theta_{75}} & 0 & p_7 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p_8 \end{pmatrix}, \quad (57)$$

where the off-diagonal elements denote correlations between the qutrits (note that the nonzero terms are those relevant to the heat received by system $\mathcal{H}_A$, according to Eq. (48)) and the diagonal elements, denoted as $\{p_i\}_i$, form a classical probability distribution contingent upon the inverse temperatures β_A and β_B of qutrits $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively (for their specific formulations, see Appendix F). Note that the parameters η_{31} , η_{62} , η_{75} , θ_{31} , θ_{62} , and θ_{75} are *not* free and must be such that $\rho_{3 \otimes 3}$ is positive semidefinite, together with the fact that $\{p_i\}_{i=0}^8$ must define a probability distribution. The heat content of qutrit $\mathcal{H}_A$, expressed as $\langle Q_A \rangle$, can be evaluated directly from Eq. (4) or using Eq. (48):

$$\begin{aligned} \langle Q_A \rangle &= \text{Tr} \left\{ \rho_{3 \otimes 3} (U_3(t) H_A \otimes \mathbb{1}_B U_3^\dagger(t) - H_A \otimes \mathbb{1}_B) \right\} \\ &= \zeta \sin^2(gt) + \xi \sin(gt) \cos(gt), \\ &= \zeta \sin^2(gt) + \frac{\xi}{2} \sin(2gt), \end{aligned} \quad (58)$$

where ζ is equivalent to the sum $\sum_m p_{m,m}^B \omega_m - \sum_n p_{n,n}^A \omega_n$ of Eq. (48) and depends on the inverse temperatures β_i and energies ω_i of the qutrits, while ξ is equivalent to

the sum $\sum_{n,m} \text{Im}(p_{n,m,m,n}) \omega_n / 2$ of Eq. (48) and encapsulates the correlations η_{ij} and phases θ_{ij} (for details, see Appendix F). Exemplifying our discussion on qudits, the term ζ is responsible for the standard heat flow (meaning that $\beta_A < \beta_B$ implies that $\zeta > 0$), while ξ is the term that makes the heat-flow inversion possible. For small values of gt , contextual effects are assured for non-null values of ξ . It is noteworthy that while ξ can be utilized to reverse the heat flux, such inversion is not a prerequisite for contextuality, since this term can also cause an increase in the standard heat flow. Let us investigate the violation of the noncontextuality inequality (found in Ref. [27]) using Eq. (58). Let $\omega_{\max} = \max\{\omega_0, \omega_1, \omega_2\}$. Then, we can find the critical times τ_c^l, τ_c^u assuming that $\zeta, \xi, \omega_{\max}$, and g are all fixed parameters and $\xi \neq 0$, given by

$$\tau_c^u = \frac{1}{g} \cot^{-1} \left(\frac{2\omega_{\max} - \zeta}{\xi} \right), \quad (59)$$

$$\tau_c^l = \frac{1}{g} \cot^{-1} \left(\frac{-4\omega_{\max} - \zeta}{\xi} \right), \quad (60)$$

where above we have considered τ_c^u as the critical time associated with the noncontextuality inequality upper bound, and τ_c^l for the lower bound. These bounds can be

used to directly infer the relationship between contextuality and anomalous heat flow.

Moreover, the change of the mutual information in this situation has an analogous result as in the two-qubit case. Again, the heat term responsible for the anomalous heat flow and the violation of our noncontextuality inequality is the same one that causes the change of the mutual information to be negative for $gt \ll 1$ (for details, see Appendix F). Hence, the effects of correlations on the Clausius inequality also have a direct connection with contextuality certification in this case, which can indicate the same pattern in the higher-dimensional case.

IV. SUMMARY AND OUTLOOK

In this paper, we have demonstrated that heat-flow inversion, caused by initial correlations between two-qubit unitary interactions that conserve total energy, *cannot* be described by generalized noncontextual models when the interaction happens in certain intervals of time $0 < t < \tau_c$. These results introduce the notion of critical times τ_c governing dynamical nonclassicality.

For the kind of qubit interactions that we have investigated, a coherent effect in the heat flow (either reversing the flow or increasing it in the standard direction) must be present to allow for a violation of the noncontextual inequalities that we have introduced. We also show analogous results for two qudits, with any dimension d , interacting via a partial SWAP, indicating that similar conclusions in higher-dimensional systems are true for a large class of applicable interactions and can be further explored for more general interactions.

As an application of our findings, we use the experimental parameters of Ref. [20] to show that the connection we find between contextuality and reversal of the spontaneous direction of heat flow can be tested by existing quantum hardware and state-of-the-art manipulation of quantum resources. For an experimental test, a robust account on the critical time τ_c could be found by extending Theorems 1, 2, and 3 to consider experimental imperfections.

It is worth pointing out that in Ref. [14], connections have been shown between the reversal of the direction of heat flow and another form of nonclassicality, defined to be the negativity of the real part of Kirkwood-Dirac quasiprobability distributions [83–87], known as the Terletsky-Margenau-Hill quasiprobability distribution [14,86,88,89]. As the authors have pointed out, negativity in such distributions is a proxy for generalized contextuality [28,29,45] but is known to *not* be sufficient for contextuality in general [30].

Furthermore, one can also use the combination of Theorems 2 and 3 to explore the violation of noncontextuality inequalities for the variation of other observables, rather than energy, for energy-conserving interactions between qubits. For instance, one can study the population inversion by choosing the observable $A = \sigma_z$ (for density

matrices in the eigenbasis of σ_z) and obtain noncontextuality inequalities in this case. This scenario is useful for certifying the presence of contextuality in open quantum systems with the collisional-models approach [90–92], where the models often involve two-qubit interactions (for examples, see Refs. [92–97]).

Similarly to Ref. [4], which has shown that the occurrence of *strong heat backflow* is a signature of entanglement, our results argue that the presence of heat-flow inversion caused by correlations—or, more generally, any coherent correlation effect in the heat flow—in energy-conserving two-qubit interactions, and partial SWAP interactions between two qudits, imply a signature of generalized contextuality for a time interval between the beginning of the interaction and a critical time τ_c , with the value of τ_c depending on the parameters of the scenario. Similar scenarios are ubiquitous in quantum thermodynamics [46–52]. We believe that our results can be useful to certify contextuality in a variety of models as well as to indicate contextuality as a resource for genuinely nonclassical phenomena in quantum thermodynamics.

ACKNOWLEDGMENTS

We acknowledge Matteo Lostaglio for his valuable suggestions, including the estimation of the critical time τ_c in comparison with experimental parameters. We acknowledge the suggestions from PRX Quantum reviewers, which, among other things, led us to the results concerning the two interacting qudits via a partial SWAP. Furthermore, we acknowledge Rodrigo Isaú Ramos Moreira for helping us with the SWAP operator for qudits. B.A. and N.E.C. acknowledge financial support from the Instituto Serrapilheira, Chamada No. 4 2020. N.E.C. also acknowledges the support of the Financiadora de Estudos e Projetos (Grant No. 1699/24 IIF-FINEP). B.A. and D.C. acknowledge Pró-Reitoria de Pesquisa e Inovação (PRPI) from the Universidade de São Paulo (USP) for financial support through the Programa de Estímulo à Supervisão de Pós-Doutorandos por Jovens Pesquisadores. B.A. also acknowledges the Fundação de Amparo à Pesquisa do Estado de São Paulo, Auxílio à Pesquisa—Jovem Pesquisador, through Grant No. 2020/06454-7. R.W. acknowledges support by the Fundação para a Ciência e a Tecnologia (FCT, Portugal) through Ph.D. Grant No. SFRH/BD/151199/2021. L.F.S. acknowledges the financial support of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)—Brazil, Finance Code 001. This work was also supported by the Digital Horizon Europe project FoQa-CiA, Grant Agreement No. no.101070558, funded by the European Union, the Natural Sciences and Engineering Research Council of Canada (NSERC, Canada), and UK Research and Innovation (UKRI, United Kingdom). R.W. acknowledges support from European Research Council (ERC) under the European Union’s Horizon 2020 research

and innovation programme (Grant agreement No. 856432, HyperQ).

DATA AVAILABILITY

All data analyzed in this study were obtained from the work of [Micadei *et al.*, 2019], as explicitly cited in our manuscript (Ref. [20]). This data is public available, and obtained from <https://www.nature.com/articles/s41467-019-10333-7>. No new data were generated by our team.

APPENDIX A: PROOF OF EQ. (6)

This proof follows the pedagogical compendium of Ref. [56]. Assume the conditions presented in Sec. II A. Let $\rho' \equiv U\rho U^\dagger$ and $\rho'_i := \text{Tr}_{\setminus i} \{\rho'\}$. The quantity

$$\mathcal{S} = S(\rho'_A || \rho_A) + S(\rho'_B || \rho_B) \quad (\text{A1})$$

is always non-negative, since it is a sum of relative entropies $S(\rho || \sigma) = \text{Tr} \{\rho \log(\rho) - \rho \log(\sigma)\} \geq 0$. Now, we can rewrite

$$\begin{aligned} S(\rho'_A || \rho_A) &= \text{Tr} \{ \rho'_A \log(\rho'_A) - \rho'_A \log(\rho_A) \} \\ &= S(\rho_A) - S(\rho'_A) + \text{Tr} \{ \rho_A \log(\rho_A) - \rho'_A \log(\rho_A) \} \\ &= -\Delta S_A + \text{Tr} \{ \rho_A \log(\rho_A) - \rho'_A \log(\rho_A) \}, \end{aligned}$$

where we have used the definition of the von Neumann entropy $S(\rho) = -\text{Tr} \{\rho \log(\rho)\}$ and defined $\Delta S_A = S(\rho'_A) - S(\rho_A)$.

Furthermore, using Eq. (3), we have $\log(\rho_A) = -\beta_A H_A - \log(Z_A)$. Using this in the above equation, we obtain

$$\begin{aligned} S(\rho'_A || \rho_A) &= -\Delta S_A + \beta_A \text{Tr} \{ (\rho'_A - \rho_A) H_A \} \\ &= -\Delta S_A + \beta_A \langle Q_A \rangle, \end{aligned} \quad (\text{A2})$$

where in the last equation we have used Eq. (4). Analogously, we have

$$S(\rho'_B || \rho_B) = -\Delta S_B + \beta_B \langle Q_B \rangle, \quad (\text{A3})$$

where $\Delta S_B = S(\rho'_B) - S(\rho_B)$.

Given the mutual information $\mathcal{I}_\rho(A : B) = -S(\rho) + S(\rho_A) + S(\rho_B)$ between the systems $\mathcal{H}_A$ and $\mathcal{H}_B$, for a unitary evolution, the variation of the mutual information $\Delta \mathcal{I}(A : B) = \mathcal{I}_{\rho'}(A : B) - \mathcal{I}_\rho(A : B)$ will be

$$\Delta \mathcal{I}(A : B) = \Delta S_A + \Delta S_B,$$

since $S(\rho) = S(\rho')$. Therefore, by summing Eqs. (A2) and (A3) and using the above equation, we obtain

$$\begin{aligned} \mathcal{S} &= S(\rho'_A || \rho_A) + S(\rho'_B || \rho_B) \\ &= \beta_A \langle Q_A \rangle + \beta_B \langle Q_B \rangle - \Delta \mathcal{I}(A : B) \geq 0, \end{aligned} \quad (\text{A4})$$

which implies

$$(\beta_A - \beta_B) \langle Q_A \rangle \geq \Delta \mathcal{I}(A : B), \quad (\text{A5})$$

where we have used energy conservation $\langle Q_A \rangle = -\langle Q_B \rangle$.

Note that in order to deduce Eq. (6), we have only used the fact that the unitary evolution does not change the von Neumann entropy, energy conservation, and $\mathcal{S} \geq 0$. This points out that, to enunciate a second law, we may only need to find a suitable always nondecreasing “irreversible” quantity, which, in this case, has been $\mathcal{S}$. The fact that this quantity is always non-negative is a mathematical fact. However, it implies some bounds on the physical quantities, given the specific evolution studied.

APPENDIX B: PROOF OF THEOREM 2

In what follows, we simply write $T_{t_1} \equiv T_1$ and $T_{t_2} \equiv T_2$ for simplicity. We start by considering the operational equivalence

$$\frac{1}{2} T_2 \circ T_1 + \frac{1}{2} T_2 \circ T_1^* \simeq (1 - p_{d_1}) T_2 + p_{d_1} T_2 \circ T_1',$$

which we rewrite as

$$\frac{1}{2} T + \frac{1}{2} T_{A_1} \simeq (1 - p_{d_1}) T_2 + p_{d_1} T_{A_2}, \quad (\text{B1})$$

where $T = T_2 \circ T_1$, $T_{A_1} = T_2 \circ T_1^*$ and $T_{A_2} = T_2 \circ T'$. Our proof will follow the exact same methodology as that of Ref. [27, Theorem 1]. We divide the proof into two cases, where we first consider ontological models in which Λ is a finite set of ontic states and then, later, we consider a more involved proof for the case in which Λ is a continuous measurable set of ontic states.

1. Finite set Λ of ontic states

Consider the ontological-models description for the average difference $\langle \Delta \mathcal{A} \rangle$ in the observable $\mathcal{A} = \{(a_k, k | M)\}_k$ given by

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle &= \sum_k a_k (p(k | M, T, P) - p(k | M, P)) \\ &= \sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_T(\lambda' | \lambda) \xi_M(k | \lambda') \right. \\ &\quad \left. - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right), \end{aligned} \quad (\text{B2})$$

where the $a_k > 0$ are the possible values assigned to $\mathcal{A}$ from effects $k | M$, $\mu_P(\lambda)$ is the probability distribution associated by the model with the preparation procedure P , $\Gamma_T(\lambda' | \lambda)$ is the transition matrix representing

the probability that the ontic states will change from λ to λ' , and $\xi_M(k|\lambda)$ is the response function associated with $k|M$, which gives the probability distribution of obtaining the outcome k , given that the ontic state was λ and that a measurement M has been performed. Without loss of generality, we consider $\Lambda = \Lambda'$.

Since $\Gamma_T(\lambda'|\lambda) \leq 1, \forall \lambda, \lambda' \in \Lambda$ in the above equation, we have

$$\langle \Delta \mathcal{A} \rangle \leq \sum_k a_k \left(\sum_{\lambda \neq \lambda'} \mu_P(\lambda) \Gamma_T(\lambda'|\lambda) \xi_M(k|\lambda') \right). \quad (\text{B3})$$

We now consider that the ontological models must respect Eq. (B1), constraining the model to be noncontextual, implying that for all $\lambda, \lambda' \in \Lambda$,

$$\begin{aligned} \frac{1}{2} \Gamma_T(\lambda'|\lambda) + \frac{1}{2} \Gamma_{A_1}(\lambda'|\lambda) \\ = (1 - p_{d_1}) \Gamma_{T_2}(\lambda'|\lambda) + p_{d_1} \Gamma_{A_2}(\lambda'|\lambda). \end{aligned} \quad (\text{B4})$$

Moreover, the assumption of transformation noncontextuality applied to Eq. (26) (stochastic reversibility for T_2) also requires that for all $\lambda, \lambda' \in \Lambda$,

$$\begin{aligned} \frac{1}{2} \Gamma_{T_2}(\lambda'|\lambda) + \frac{1}{2} \Gamma_{T_2}^*(\lambda'|\lambda) \\ = (1 - p_{d_2}) \delta_{\lambda',\lambda} + p_{d_2} \Gamma_{C_2}(\lambda'|\lambda), \end{aligned} \quad (\text{B5})$$

where $\delta_{\lambda,\lambda'} = \Gamma_{T_{\text{id}}}(\lambda|\lambda')$ is the transition-matrix representation of the identity transformation by the ontological model. Using the above two equations, and the fact that

since $a_{\max} \geq a_k, \forall k$ and $\sum_k p(k|M, T, P) = 1, \forall T \in \mathcal{T}, P \in \mathcal{P}$. The same is also valid for the term $\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{A_2}(\lambda'|\lambda) \xi_M(k|\lambda') = p(k|M, T_{A_2}, M)$ and therefore, using these inequalities in the one in Eq. (B6), we obtain

$$\langle \Delta \mathcal{A} \rangle \leq (4(1 - p_{d_1})p_{d_2} + 2p_{d_1}) a_{\max} = 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) a_{\max}. \quad (\text{B8})$$

To complete the proof for the finite- Λ case, we analyze the converse. From equation Eq. (B2), we have

$$-\langle \Delta \mathcal{A} \rangle = -\langle \mathcal{A}(t) - \mathcal{A}(0) \rangle = -\sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_T(\lambda'|\lambda) \xi_M(k|\lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k|\lambda) \right).$$

Proceeding similarly as before, isolating Γ_T using Eqs. (B4) and (B5) and substituting it in the definition of $\langle \Delta \mathcal{A} \rangle$, we obtain

$$\begin{aligned} -\langle \Delta \mathcal{A} \rangle = -\sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \left[-2(1 - p_{d_1}) \Gamma_{T_2}^*(\lambda'|\lambda) - \Gamma_{A_1}(\lambda'|\lambda) + 4(1 - p_{d_1})(1 - p_{d_2}) \delta_{\lambda, \lambda'} \right. \right. \\ \left. \left. + 4(1 - p_{d_1})p_{d_2} \Gamma_{C_2}(\lambda'|\lambda) + 2p_{d_1} \Gamma_{A_2}(\lambda'|\lambda) \right] \xi_M(k|\lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k|\lambda) \right) \end{aligned}$$

the transition matrices are always non-negative, we obtain

$$\begin{aligned} \Gamma_T(\lambda'|\lambda) \leq 4(1 - p_{d_1})(1 - p_{d_2}) \delta_{\lambda, \lambda'} \\ + 4(1 - p_{d_1})p_{d_2} \Gamma_{C_2}(\lambda'|\lambda) + 2p_{d_1} \Gamma_{A_2}(\lambda'|\lambda). \end{aligned}$$

Using this inequality in the one in Eq. (B3), we obtain

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle \leq \sum_k a_k \left(\sum_{\lambda \neq \lambda'} \mu_P(\lambda) (4(1 - p_{d_1})p_{d_2} \Gamma_{C_2}(\lambda'|\lambda) \right. \\ \left. + 2p_{d_1} \Gamma_{A_2}(\lambda'|\lambda)) \xi_M(k|\lambda') \right) \\ \leq \sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) (4(1 - p_{d_1})p_{d_2} \Gamma_{C_2}(\lambda'|\lambda) \right. \\ \left. + 2p_{d_1} \Gamma_{A_2}(\lambda'|\lambda)) \xi_M(k|\lambda') \right), \end{aligned} \quad (\text{B6})$$

where in the last inequality, we have summed back the terms $\lambda = \lambda'$, which are always non-negative.

We now note that the term $\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{C_2}(\lambda'|\lambda) \xi_M(k|\lambda') = p(k|M, T_{C_2}, M)$ is just a probability distribution of having an outcome k under a given evolution. Therefore,

$$\begin{aligned} \sum_k a_k \sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{C_2}(\lambda'|\lambda) \xi_M(k|\lambda') \\ = \sum_k a_k p(k|M, T_{C_2}, M) \leq a_{\max} \sum_k p(k|M, T_{C_2}, M) \\ = a_{\max}, \end{aligned} \quad (\text{B7})$$

$$\begin{aligned}
&\leq \sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \left[2(1 - p_{d_1}) \Gamma_{T_2}^*(\lambda' | \lambda) + \Gamma_{A_1}(\lambda' | \lambda) - 4(1 - p_{d_1})(1 - p_{d_2}) \delta_{\lambda, \lambda'} \right] \xi_M(k | \lambda') + \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right) \\
&= \sum_k a_k \left[2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) + 2(1 - p_{d_1}) \right. \\
&\quad \times \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{T_2}^*(\lambda' | \lambda) \xi_M(k | \lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right) \\
&\quad \left. + \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{A_1}(\lambda' | \lambda) \xi_M(k | \lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right) \right], \tag{B9}
\end{aligned}$$

where in the inequality we have used the fact that the terms $-\sum_k a_k \sum_{\lambda, \lambda'} \mu_P(\lambda) 4(1 - p_{d_1}) p_{d_2} \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda')$ and $-\sum_k a_k \sum_{\lambda, \lambda'} \mu_P(\lambda) 4p_{d_1} \Gamma_{A_2}(\lambda' | \lambda) \xi_M(k | \lambda')$ are never positive, and in the last equality we summed over the Kronecker delta and redistributed the terms. Now, note that the term $\sum_k a_k (\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{A_1}(\lambda' | \lambda) \xi_M(k | \lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda))$ is analogous to the right-hand side of Eq. (B2) by exchanging $\Gamma_T(\lambda' | \lambda)$ for $\Gamma_{A_1}(\lambda' | \lambda)$, and from Eqs. (B4) and (B5), we have that

$$\begin{aligned}
\Gamma_{A_1}(\lambda' | \lambda) &\leq 4(1 - p_{d_1})(1 - p_{d_2}) \delta_{\lambda, \lambda'} \\
&\quad + 4(1 - p_{d_1}) p_{d_2} \Gamma_{C_2}(\lambda' | \lambda) + 2p_{d_1} \Gamma_{A_2}(\lambda' | \lambda),
\end{aligned}$$

from which, analogous to the deduction of the inequality in Eq. (B8), we obtain

$$\begin{aligned}
\sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{A_1}(\lambda' | \lambda) \xi_M(k | \lambda') \right. \\
\left. - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right) &\leq 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) a_{\max}. \tag{B10}
\end{aligned}$$

In a similar way, the term $\sum_k a_k (\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{T_2}(\lambda' | \lambda) \xi_M(k | \lambda') - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda))$ is also analogous to the right-hand side of Eq. (B2), and from Eq. (B5), we obtain

$$\Gamma_{T_2}(\lambda' | \lambda) \leq 2(1 - p_{d_2}) \delta_{\lambda', \lambda} + 2p_{d_2} \Gamma_{C_2}(\lambda' | \lambda),$$

from which, using the same arguments that lead to Eq. (B8), we obtain

$$\begin{aligned}
\sum_k a_k \left(\sum_{\lambda, \lambda'} \mu_P(\lambda) \Gamma_{T_2}(\lambda' | \lambda) \xi_M(k | \lambda') \right. \\
\left. - \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) \right) &\leq 2p_{d_2} a_{\max}. \tag{B11}
\end{aligned}$$

Using the inequalities in Eqs. (B10) and (B11) in the inequality in Eq. (B9), we have

$$\begin{aligned}
-\langle \Delta \mathcal{A} \rangle &\leq 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) \\
&\quad \times \sum_k a_k \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) + 4(1 - p_{d_1}) p_{d_2} a_{\max} \\
&\quad + 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) a_{\max}. \tag{B12}
\end{aligned}$$

Finally, noting that the term $\sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) = p(k | M, P)$, we have as before $\sum_k a_k \sum_{\lambda} \mu_P(\lambda) \xi_M(k | \lambda) = \sum_k a_k p(k | M, P) \leq a_{\max}$. Using this in the inequality in Eq. (B12), we obtain

$$\begin{aligned}
-\langle \Delta \mathcal{A} \rangle &\leq 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) a_{\max} \\
&\quad + 4(1 - p_{d_1}) p_{d_2} a_{\max} + 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) a_{\max} \\
&= 4(p_{d_1} + 3p_{d_2} - 3p_{d_1}p_{d_2}) a_{\max}, \tag{B13}
\end{aligned}$$

which completes the proof for the finite- Λ case.

2. Continuous set Λ of ontic states

For an *infinite* (and possibly *continuous*) set Λ , we have the following version of Eq. (B2):

$$\begin{aligned}
\langle \Delta \mathcal{A} \rangle &= \langle \mathcal{A}(t) - \mathcal{A}(0) \rangle \\
&= \sum_k a_k \left(\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' | \lambda) \xi_M(k | \lambda') \right. \\
&\quad \left. - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right), \tag{B14}
\end{aligned}$$

where Λ is a measure space with measure λ . The transition matrices are probability densities defined on this space, and the integrals are Lebesgue integrals. Moreover, generalized noncontextuality requirements (such as $P \simeq P' \implies \mu_P = \mu_{P'}$) are now defined up to sets of measure zero. Then, we define the set

$$\bar{\Lambda}_k(\lambda) = \{\lambda' \mid \xi(k \mid \lambda') > \xi(k \mid \lambda)\}, \quad (\text{B15})$$

for each outcome k and ontological state λ . Additionally, we have $\bar{\Lambda}_k^c(\lambda) = \Lambda / \bar{\Lambda}_k(\lambda)$. Now, Eq. (B14) implies that

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle &= \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda') \right. \\ &\quad \left. + \int_{\bar{\Lambda}_k^c(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k \mid \lambda) \right) \\ &\leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda') \right. \\ &\quad \left. + \int_{\bar{\Lambda}_k^c(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda) - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k \mid \lambda) \right), \end{aligned} \quad (\text{B16})$$

where in the inequality we have used the fact that $\xi_M(k \mid \lambda') \leq \xi_M(k \mid \lambda)$ for $\lambda' \in \bar{\Lambda}_k^c(\lambda)$.

Therefore,

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle &\leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda') \right. \\ &\quad \left. + \int_{\bar{\Lambda}_k^c(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda) - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k \mid \lambda) \right) \\ &= \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) (\xi_M(k \mid \lambda') - \xi_M(k \mid \lambda)) \right. \\ &\quad \left. + \int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda) - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k \mid \lambda) \right) \\ &= \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) (\xi_M(k \mid \lambda') - \xi_M(k \mid \lambda)) \right), \end{aligned} \quad (\text{B17})$$

where the term $\sum_k a_k \int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_T(\lambda' \mid \lambda) \xi_M(k \mid \lambda)$ has been summed and subtracted in the first equality above, and in the last equality, we have used the fact that $\int_{\Lambda} d\lambda' \Gamma_T(\lambda' \mid \lambda) = 1$.

At this point, we make the *transformation-noncontextuality* assumption in the operational equivalences of Eqs. (26) and (B1). This results in equations identical to Eqs. (B4) and (B5), where the transition matrices are now probability densities, and where $\delta_{\lambda, \lambda'}$ is substituted by $\delta(\lambda - \lambda')$, which is a Dirac delta function. Using the combination of the continuous form of Eqs. (B4) and (B5) and the fact that $\Gamma_{A_1}(\lambda' \mid \lambda) \geq 0$, we have

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle &\leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) (4(1 - p_{d_1})(1 - p_{d_1})\delta(\lambda - \lambda') \right. \\ &\quad \left. + 4(1 - p_{d_1})p_{d_2}\Gamma_{C_2}(\lambda' \mid \lambda) + 2p_{d_1}\Gamma_{A_2}(\lambda' \mid \lambda)) (\xi_M(k \mid \lambda') - \xi_M(k \mid \lambda)) \right). \end{aligned} \quad (\text{B18})$$

The term with the Dirac delta in the above inequality is null after performing the double integral, since the space $\bar{\Lambda}_k(\lambda)$ has only terms with $\lambda' \neq \lambda$. Additionally, since $\xi_M(k \mid \lambda) \geq 0$, we have

$$\begin{aligned} \langle \Delta \mathcal{A} \rangle &\leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) (4(1 - p_{d_1})p_{d_2}\Gamma_{C_2}(\lambda' \mid \lambda) + 2p_{d_1}\Gamma_{A_2}(\lambda' \mid \lambda)) \xi_M(k \mid \lambda') \right) \\ &\leq \sum_k a_k \left(\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) (4(1 - p_{d_1})p_{d_2}\Gamma_{C_2}(\lambda' \mid \lambda) + 2p_{d_1}\Gamma_{A_2}(\lambda' \mid \lambda)) \xi_M(k \mid \lambda') \right), \end{aligned} \quad (\text{B19})$$

where in the last inequality we have summed the term $\int_{\tilde{\Lambda}_k^c(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) ((1 - p_d)\epsilon\Gamma'(\lambda' | \lambda) + p_d\Gamma_C(\lambda' | \lambda)) \xi_M(k | \lambda')$, which is non-negative.

Noting that $\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P\Gamma_{C_2}(\lambda' | \lambda)\xi_M(k | \lambda') = p(k | M, T_{C_2}, P)$ and $\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P\Gamma_{A_2}(\lambda' | \lambda)\xi_M(k | \lambda') = p(k | M, T_{A_2}, P)$ are probabilities for obtaining outcome k , given a set of operational equivalences, we have

$$\sum_k a_k \int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P\Gamma_{C_2(A_2)}(\lambda' | \lambda)\xi_M(k | \lambda') = \sum_k a_k p(k | M, T_{C_2(A_2)}, P) \leq a_{\max}. \quad (\text{B20})$$

Using the inequality in Eq. (B19), we obtain

$$\langle \Delta\mathcal{A} \rangle \leq (4(1 - p_{d_1})p_{d_2} + 2p_{d_1})a_{\max} = 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2})a_{\max}. \quad (\text{B21})$$

To prove the converse, we start again from Eq. (B14), which implies that

$$-\langle \Delta\mathcal{A} \rangle = -\langle \mathcal{A}(t) - \mathcal{A}(0) \rangle = \sum_k a_k \left(\int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) - \int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_T(\lambda' | \lambda) \xi_M(k | \lambda') \right).$$

We now make the transformation-noncontextuality assumption, using the infinite version of Eqs. (B4) and (B5). Hence, we obtain

$$\begin{aligned} -\langle \Delta\mathcal{A} \rangle = \sum_k a_k & \left(\int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) - \int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \left[-2(1 - p_{d_1})\Gamma_{T_2^*}(\lambda' | \lambda) - \Gamma_{A_1}(\lambda' | \lambda) \right. \right. \\ & \left. \left. + 4(1 - p_{d_1})(1 - p_{d_2})\delta(\lambda - \lambda') + 4(1 - p_{d_1})p_{d_2}\Gamma_{C_2}(\lambda' | \lambda) + 2p_{d_1}\Gamma_{A_2}(\lambda' | \lambda) \right] \xi_M(k | \lambda') \right). \end{aligned} \quad (\text{B22})$$

From the fact that $-4(1 - p_{d_1})p_{d_2} \int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{C_2} \xi_M(k | \lambda')$ and $-2p_{d_1} \int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{A_2}(\lambda' | \lambda) \xi_M(k | \lambda')$ are never positive, we conclude that

$$\begin{aligned} -\langle \Delta\mathcal{A} \rangle & \leq \sum_k a_k \left(\int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) - \int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \left(-2(1 - p_{d_1})\Gamma_{T_2^*}(\lambda' | \lambda) - \Gamma_{A_1}(\lambda' | \lambda) \right. \right. \\ & \left. \left. + 4(1 - p_{d_1})(1 - p_{d_2})\delta(\lambda - \lambda') \right) \right) \\ & = \sum_k a_k \left[2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2}) \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right. \\ & \left. + 2(1 - p_{d_1}) \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \right. \\ & \left. + \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{A_1}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \right], \end{aligned} \quad (\text{B23})$$

where in the last step we have only evaluated the integral over the $\delta(\lambda - \lambda')$ and reorganized the terms.

Moreover, the term $\sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_{A_1}}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right)$ is analogous to the right-hand side of Eq. (B14) with the exchange of $\Gamma_T(\lambda' | \lambda)$ by $\Gamma_{T_{A_1}}(\lambda' | \lambda)$. Since these terms are symmetric in the noncontextuality condition (the continuous version of Eq. (B4)), we can repeat the same arguments from Eq. (B14) for the inequality in Eq. (B21) and conclude that

$$\sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{A_1}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \leq 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2})a_{\max}. \quad (\text{B24})$$

Similarly, the term $\sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right)$ is also analogous to the right-hand side of Eq. (B14), with the exchange of $\Gamma_T(\lambda' | \lambda)$ by $\Gamma_{T_2^*}(\lambda' | \lambda)$. Therefore, we repeat the same arguments from Eq. (B14) to the inequality of Eq. (B17), to obtain

$$\begin{aligned} & \sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \\ & \leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) (\xi_M(k | \lambda') - \xi_M(k | \lambda)) \right). \end{aligned} \quad (\text{B25})$$

We now use the continuous version of the noncontextuality condition of Eq. (B5) and use the fact that $\Gamma_{T_2}(\lambda' | \lambda) \geq 0$ to obtain

$$\begin{aligned} & \sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \\ & \leq \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) (2(1 - p_{d_2})\delta(\lambda - \lambda') + 2p_{d_2} \Gamma_{C_2}(\lambda' | \lambda)) (\xi_M(k | \lambda') - \xi_M(k | \lambda)) \right). \end{aligned} \quad (\text{B26})$$

Again, the double integral over the $\delta(\lambda - \lambda')$ will be null due to the integration over the space $\bar{\Lambda}_k(\lambda)$ and we use the fact that $\xi_M(k | \lambda) \geq 0$ to obtain

$$\begin{aligned} & \sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \\ & \leq 2p_{d_2} \sum_k a_k \left(\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda') \right) \\ & \leq 2p_{d_2} \sum_k a_k \left(\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda') \right), \end{aligned} \quad (\text{B27})$$

where in the last inequality we have used the fact that $\int_{\bar{\Lambda}_k(\lambda)} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda') \leq \int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda)$. Now, again we note that $\int_{\Lambda} d\lambda' \int_{\Lambda} d\lambda \mu_P(\lambda) \Gamma_{C_2}(\lambda' | \lambda) \xi_M(k | \lambda') = p(k | M, T_{C_2}, P)$ is a probability of obtaining the outcome k . Hence, this quantity must satisfy an inequality similar to Eq. (B20), from which we obtain

$$\sum_k a_k \left(\int_{\Lambda} d\lambda \int_{\Lambda} d\lambda' \mu_P(\lambda) \Gamma_{T_2^*}(\lambda' | \lambda) \xi_M(k | \lambda') - \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) \right) \leq 2p_{d_2} a_{\max}. \quad (\text{B28})$$

Finally, since $\int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) = p(k | M, P)$ is the probability distribution for obtaining the outcome k in the prepare-and-measure case (P, M) , it must also satisfy

$$\sum_k a_k \int_{\Lambda} d\lambda \mu_P(\lambda) \xi_M(k | \lambda) = \sum_k a_k p(k | M, P) \leq a_{\max}.$$

Using the above inequality and the inequalities in Eqs. (B24) and (B28) in Eq. (B23), we obtain

$$-\langle \Delta \mathcal{A} \rangle \leq 2(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2})a_{\max} + 4(1 - p_{d_1})p_{d_2}a_{\max} + 3(p_{d_1} + 2p_{d_2} - 2p_{d_1}p_{d_2})a_{\max} \leq 4(p_{d_1} + 3p_{d_2} - 3p_{d_1}p_{d_2}),$$

which completes the proof for the inequality for infinite Λ .

APPENDIX C: PROOF OF THEOREM 3

We start by showing that a general interaction Hamiltonian between two *nonresonant* qubits, which conserves the total energy of the qubits, must be of the following form:

$$H_{I_{\text{nr}}} = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{C1})$$

for a real number g . In the case of an interaction Hamiltonian between two *resonant* qubits, it must have the following form:

$$H_{I_{\text{r}}} = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & e^{i\theta} & 0 \\ 0 & e^{-i\theta} & a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{C2})$$

where a , g , and θ are real parameters.

Note that the most general interaction Hamiltonian between two qubits is the most general linear combination of tensor products of Pauli matrices acting on the Hilbert spaces $\mathcal{H}_A = \mathbb{C}^2$ and $\mathcal{H}_B = \mathbb{C}^2$:

$$\begin{aligned} H_I = & \alpha_0 \mathbb{1}^A \otimes \mathbb{1}^B + \alpha_{11} \sigma_x^A \otimes \sigma_x^B + \alpha_{12} \sigma_x^A \otimes \sigma_y^B \\ & + \alpha_{13} \sigma_x^A \otimes \sigma_z^B + \alpha_{21} \sigma_y^A \otimes \sigma_x^B + \alpha_{22} \sigma_y^A \otimes \sigma_y^B \\ & + \alpha_{23} \sigma_y^A \otimes \sigma_z^B + \alpha_{31} \sigma_z^A \otimes \sigma_x^B \\ & + \alpha_{32} \sigma_z^A \otimes \sigma_y^B + \alpha_{33} \sigma_z^A \otimes \sigma_z^B, \end{aligned} \quad (\text{C3})$$

where all coefficients must be real numbers.

Furthermore, the assumption that the interaction Hamiltonian conserves the sum of the two local energies is equivalent to

$$[H_I, H_A \otimes \mathbb{1}_B + \mathbb{1}_A \otimes H_B] = 0, \quad (\text{C4})$$

where H_A is the local Hamiltonian of the system $\mathcal{H}_A$ and H_B is the local Hamiltonian of the system $\mathcal{H}_B$.

We suppose that the qubits have local Hamiltonians H_A and H_B described by what we refer to as *Zeeman Hamiltonians*. If the two qubits are *nonresonant*, without loss of generality, we say that $H_A = \omega_A/2(c_1 \mathbb{1}^A + c_2 \sigma_z^A)$ and $H_B = \omega_B/2(c_3 \mathbb{1}^B + c_4 \sigma_z^B)$, for $\omega_A \neq \omega_B$ and $c_i \in \mathbb{R}$. With these assumptions, Eq. (C4) results in a set of 16 equations, the solutions of which imply the following interaction Hamiltonian:

$$H_{I_{\text{nr}}} = \begin{pmatrix} \alpha_0 + \alpha_{33} & 0 & 0 & 0 \\ 0 & \alpha_0 - \alpha_{33} & 0 & 0 \\ 0 & 0 & \alpha_0 - \alpha_{33} & 0 \\ 0 & 0 & 0 & \alpha_0 + \alpha_{33} \end{pmatrix}.$$

Since the sum of identity terms on the Hamiltonian does not change the dynamics, we are free to select the value of

α_0 . We choose $\alpha_0 = -\alpha_{33}$ and define $g = -2\alpha_{33}$, which recovers Eq. (C1).

Similarly, if the qubits are *resonant*, we suppose that $H_A = H_B = \omega/2(c_1 \mathbb{1}^{A(B)} + c_2 \sigma_z^{A(B)})$. For this case, Eq. (C4) results in a set of 16 equations, the solutions of which imply the following interaction Hamiltonian:

$$H_{I_{\text{r}}} = \begin{pmatrix} \alpha_0 + \alpha_{33} & 0 & 0 & 0 \\ 0 & \alpha_0 - \alpha_{33} & 2\alpha_{22} - 2i\alpha_{21} & 0 \\ 0 & 2\alpha_{22} + 2i\alpha_{21} & \alpha_0 - \alpha_{33} & 0 \\ 0 & 0 & 0 & \alpha_0 + \alpha_{33} \end{pmatrix}.$$

Again, we select $\alpha_0 = -\alpha_{33}$ and define a , g , and θ , such that $ga = -2\alpha_{33}$ and $ge^{-i\theta} = 2\alpha_{22} + 2i\alpha_{21}$, which imply Eq. (C2).

To prove the main result of the theorem, we start with the *nonresonant*-qubit case. We desire to prove that the unitary $U_{I_{\text{nr}}} = e^{-iH_{I_{\text{nr}}}}$ satisfies an equation in the form of Eq. (28). To do this, we write the most general form of a two-qubit density matrix,

$$\rho_{\text{gen}} = \begin{pmatrix} p_{00} & v_1^* & v_2^* & \gamma^* \\ v_1 & p_{01} & \eta e^{i\xi} & v_3^* \\ v_2 & \eta e^{-i\xi} & p_{10} & v_4^* \\ \gamma & v_3 & v_4 & p_{11} \end{pmatrix}, \quad (\text{C5})$$

with $p_{00} + p_{01} + p_{10} + p_{11} = 1$, where the diagonal terms are real non-negative numbers while v_1, v_2, v_3, v_4 , and γ are complex numbers, and η and ξ are real; all parameters are constrained such that $\rho_{\text{gen}} \geq 0$. And now, given the unitary evolution, $U_{\text{nr}} = e^{-iH_{I_{\text{nr}}}}$, generated by the interaction Hamiltonian given by Eq. (C1), we have

$$\begin{aligned} \frac{1}{2} \mathcal{U}_{\text{nr}}[\rho_{\text{gen}}] + \frac{1}{2} \mathcal{U}_{\text{nr}}^\dagger[\rho_{\text{gen}}] &= \frac{1}{2} U_{\text{nr}} \rho_{\text{gen}} U_{\text{nr}}^\dagger + \frac{1}{2} U_{\text{nr}}^\dagger \rho_{\text{gen}} U_{\text{nr}} \\ &= \begin{pmatrix} p_{00} & v_1^* \cos(gt) & v_2^* \cos(gt) & \gamma^* \\ v_1 \cos(gt) & p_{01} & \eta e^{i\xi} & v_3^* \cos(gt) \\ v_2 \cos(gt) & \eta e^{-i\xi} & p_{10} & v_4^* \cos(gt) \\ \gamma & v_3 \cos(gt) & v_4 \cos(gt) & p_{11} \end{pmatrix}. \end{aligned} \quad (\text{C6})$$

This can be factorized as

$$\frac{1}{2} \mathcal{U}_{\text{nr}}[\rho_{\text{gen}}] + \frac{1}{2} \mathcal{U}_{\text{nr}}^\dagger[\rho_{\text{gen}}] = (1 - p_{d_{\text{nr}}}) \rho_{\text{gen}} + p_{d_{\text{nr}}} C_{\text{nr}}[\rho_{\text{gen}}], \quad (\text{C7})$$

where $p_{d_{\text{nr}}} = \sin^2(gt/2)$ and

$$C_{\text{nr}}[\rho_{\text{gen}}] := \begin{pmatrix} p_{00} & -v_1^* & -v_2^* & \gamma^* \\ -v_1 & p_{01} & \eta e^{i\xi} & -v_3^* \\ -v_2 & \eta e^{-i\xi} & p_{10} & -v_4^* \\ \gamma & -v_3 & -v_4 & p_{11} \end{pmatrix}. \quad (\text{C8})$$

To prove that Eq. (C7) is really the same as Eq. (28), we must prove that the channel $\mathcal{C}_{\text{nr}}(\bullet)$ defined above is a completely positive trace-preserving (CPTP) map. To verify *complete positivity*, we can compute its Choi-Jamiołkowski matrix [66,67] and prove that this matrix is positive semidefinite. The Choi-Jamiołkowski matrix for a map $\mathcal{C}(\bullet)$ is defined as follows:

$$\begin{aligned}\Lambda_{\mathcal{C}} &:= \mathcal{I}^{\mathcal{R}} \otimes \mathcal{C}^{\mathcal{S}} \left(\left| \tilde{\Psi} \right\rangle \left\langle \tilde{\Psi} \right| \right) \\ &= \sum_{i=0}^{3,3} \sum_{j=0}^{} |i\rangle \langle j|^{\mathcal{R}} \otimes \mathcal{C}(|i\rangle \langle j|)^{\mathcal{S}},\end{aligned}\quad (\text{C9})$$

where the index $\mathcal{R}$ denotes the channel acting on an auxiliary Hilbert space $\mathcal{H}_{\mathcal{R}}$ with dimension 4, the index $\mathcal{S}$ denotes the channel acting on the global two-qubit system Hilbert space $\mathcal{H}_{\mathcal{S}}$, $\mathcal{I}$ is the identity channel, and $|\tilde{\Psi}\rangle = \sum_{j=0}^3 |j\rangle^{\mathcal{R}} \otimes |j\rangle^{\mathcal{S}}$ is the unnormalized maximally entangled state in $\mathcal{H}_{\mathcal{R}} \otimes \mathcal{H}_{\mathcal{S}}$.

Computing the Choi-Jamiolkowski matrix $\Lambda_{C_{\text{nr}}}$ explicitly for the map $C_{\text{nr}}(\bullet)$, we obtain

[illegible]

This matrix has eigenvalues $\{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4\}$ and therefore it is positive semidefinite. Hence, the channel $\mathcal{C}_{\text{nr}}(\bullet)$ is completely positive. Since the trace of the map described in Eq. (C8) is the same as the initial density matrix in Eq. (C5), the channel is also clearly *trace preserving*. Therefore, Eq. (28) holds, as we wanted to show.

To prove the resonant case, we start by noting that the interaction Hamiltonian of Eq. (C2) can be written in the following form:

$$H_{I_r} = H_\theta + H_a,$$

where

$$H_\theta = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & e^{i\theta} & 0 \\ 0 & e^{-i\theta} & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$H_a = g \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a-1 & 0 & 0 \\ 0 & 0 & a-1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since $U_r = e^{-itH_r}$, because $[H_\theta, H_a] = 0$ we have that U_r can be decomposed as

$$U_r = U_\theta U_a,$$

where $U_\theta = e^{-itH_\theta}$ and $U_a = e^{-itH_a}$. For the rest of the proof, we show that U_a and U_θ each satisfy Eqs. (29) and (30), respectively.

Again, we suppose that the initial global state starts in its general form Eq. (C5). For the case of U_a , note that H_a has the same form as H_{nr} [Eq. (C1)] with the exchange of g by $g(a - 1)$. Therefore, a similar equation to Eq. (C7) is immediately valid, namely,

$$\frac{1}{2}\mathcal{U}_\theta[\rho_{\text{gen}}] + \frac{1}{2}\mathcal{U}_\theta^\dagger[\rho_{\text{gen}}] = (1 - p_1)\rho_{\text{gen}} + p_1\mathcal{C}_1[\rho_{\text{gen}}],$$

where $p_1 = \sin^2((a-1)gt/2)$, and $\mathcal{C}_1[\rho_{\text{gen}}]$ is defined exactly as in Eq. (C8), which we have already proved is a CPTP map.

As for U_θ , we compute

$$\begin{aligned} \frac{1}{2}\mathcal{U}_\theta[\rho_{\text{gen}}] + \frac{1}{2}\mathcal{U}_\theta^\dagger[\rho_{\text{gen}}] &= \frac{1}{2}U_\theta\rho_{\text{gen}}U_\theta^\dagger + \frac{1}{2}U_\theta^\dagger\rho_{\text{gen}}U_\theta \\ &= \begin{pmatrix} p_{00} & f_\theta(2,1)^* & f_\theta(3,1)^* & \gamma^* \\ f_\theta(2,1) & f_\theta(2,2) & f_\theta(3,2)^* & f_\theta(4,2)^* \\ f_\theta(3,1) & f_\theta(3,2) & f_\theta(3,3) & f_\theta(4,3)^* \\ \gamma & f_\theta(4,2) & f_\theta(4,3) & p_{11} \end{pmatrix}, \end{aligned} \quad (\text{C10})$$

where we now have a more complicated matrix defined via the functions

$$\begin{aligned} f_\theta(2,2) &= \frac{1}{2}((p_{01} - p_{10})\cos(2gt) + p_{01} + p_{10}), \\ f_\theta(3,3) &= \frac{1}{2}(p_{01} + p_{10} + (p_{10} - p_{01})\cos(2gt)), \\ f_\theta(2,1) &= v_1\cos^2(gt) - v_2e^{i\theta}\sin^2(gt), \\ f_\theta(3,1) &= v_2\cos^2(gt) - v_1e^{-i\theta}\sin^2(gt), \\ f_\theta(3,2) &= \eta e^{-i\theta}(\cos(\xi - \theta) - i\cos(2gt)\sin(\xi - \theta)), \\ f_\theta(4,2) &= v_3\cos^2(gt) - v_4e^{-i\theta}\sin^2(gt), \\ f_\theta(4,3) &= v_4\cos^2(gt) - v_3e^{i\theta}\sin^2(gt). \end{aligned}$$

Eq. (C10) can be factorized as

$$\frac{1}{2}\mathcal{U}_\theta[\rho_{\text{gen}}] + \frac{1}{2}\mathcal{U}_\theta^\dagger[\rho_{\text{gen}}] = (1 - p_{d_2})\rho_{\text{gen}} + p_{d_2}\mathcal{C}_2[\rho_{\text{gen}}], \quad (\text{C11})$$

where

$$p_{d_2} = \sin^2(gt)$$

and

$$\mathcal{C}_2[\rho_{\text{gen}}] = \begin{pmatrix} p_{00} & -v_2^*e^{-i\theta} & -v_1^*e^{i\theta} & \gamma^* \\ -v_2e^{i\theta} & p_{10} & \eta e^{-i\xi}e^{i2\theta} & -v_4^*e^{i\theta} \\ -v_1e^{-i\theta} & \eta e^{i\xi}e^{-i2\theta} & p_{01} & -v_3^*e^{-i\theta} \\ \gamma & -v_4e^{-i\theta} & -v_3e^{i\theta} & p_{11} \end{pmatrix}. \quad (\text{C12})$$

Eq. (C11) gives us the desired form of factorization (i.e., the form of Eq. (26)). The remaining step is to prove that the map $\mathcal{C}_2(\bullet)$, described in Eq. (C12), is a CPTP map. Again, trace preservation follows trivially, so we proceed to show complete positivity.

Computing $\Lambda_{\mathcal{C}_2}$ (defined in Eq. (C9)) explicitly, we obtain

$$\Lambda_{\mathcal{C}_2} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -e^{i\theta} & 0 & 0 & -e^{-i\theta} & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -e^{-i\theta} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & e^{-2i\theta} & 0 & 0 & 0 & 0 & 0 & -e^{-i\theta} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -e^{i\theta} & 0 & 0 & 0 & 0 & 0 & e^{2i\theta} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -e^{i\theta} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -e^{i\theta} & 0 & 0 & -e^{-i\theta} & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

This matrix has eigenvalues $\{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4\}$. Therefore, it is positive semidefinite and the channel $\mathcal{C}_2(\bullet)$ is completely positive, and this concludes the proof.

APPENDIX D: $\Delta\mathcal{I}(A : B)$ FOR QUBITS AND THE CLAUSIUS INEQUALITY

The relation between heat flow and the change in mutual information comes from

$$\Delta\mathcal{I}(A : B) = (\beta_A - \beta_B)\langle\mathcal{Q}_A\rangle - S(\rho'_A||\rho_A) - S(\rho'_B||\rho_B), \quad (\text{D1})$$

which is a consequence of Eq. (A4), derived in Appendix A. To show that, for short times, the main term contributing to an anomalous heat flow is also the one that contributes to a negative mutual-information difference $\Delta\mathcal{I}$, we proceed as follows. Let the initial global two-qubit density matrix be given by Eq. (39) (with $\nu_0 = \nu_1 =$

$\nu_2 = 0$ for simplicity, since they are irrelevant for heat exchange). In this case, if the global state evolves according to the general resonant energy-conserving Hamiltonian [Eq. (37)], then for short times, we obtain the following expression for the change in mutual information:

$$\begin{aligned} \Delta\mathcal{I}(A : B) = & (\beta_A - \beta_B)2\eta\omega \sin(\xi - \theta)gt \\ & + \left(\omega(\beta_A - \beta_B) \left(\frac{1}{e^{\beta_B\omega} + 1} - \frac{1}{e^{\beta_A\omega} + 1} \right) - 4\eta^2 \sin^2(\theta - \xi)(\cosh(\omega\beta_A) + \cosh(\omega\beta_B) + 2) \right) g^2 t^2 + \mathcal{O}(g^3 t^3). \end{aligned} \quad (\text{D2})$$

It is impractical to display the full analytic expression for the mutual information here, but for our analysis it is sufficient to consider the first two orders in gt .

The first-order term, $(\beta_A - \beta_B)2\eta\omega \sin(\xi - \theta)gt$, is the leading contribution responsible for the negativity of $\Delta\mathcal{I}$ and is exactly $(\beta_A - \beta_B)$ times the first-order term of the heat average shown in Eq. (40). This comes from the heat contribution to the mutual-information change, as given in Eq. (D1), and is of the same order as the gt heat-flow term responsible for violating the noncontextuality inequality. Therefore, the same term that signals quantum contextuality and the reversal of heat flow is the one that contributes, in leading order, to the negative change in the mutual information at short times.

APPENDIX E: PROOF OF EQ. (48)

As described in the main text, we compute the heat received by a system A , which is a qudit with Hilbert-space dimension $d \geq 2$. The local Hamiltonian of the qudit A is given by Eq. (46). This qudit interacts with another qudit B ,

with Hilbert-space dimension d , via a partial SWAP unitary, given by Eq. (44). Let the initial density matrix of the joint system AB be given by Eq. (47); then the heat received by A , after it interacts during an interval of time t with B , is

$$\langle \mathcal{Q}_A \rangle = \text{tr}\{H_A(U_{PS}\rho U_{PS}^\dagger - \rho)\}. \quad (\text{E1})$$

To compute this quantity, we make the following computation, using Eq. (44):

$$\begin{aligned} U_{PS}\rho U_{PS}^\dagger = & \cos^2(gt)\rho + i \sin(gt) \cos(gt)\rho S \\ & - i \sin(gt) \cos(gt)S\rho + \sin^2(gt)S\rho S. \end{aligned} \quad (\text{E2})$$

This implies that

$$\begin{aligned} U_{PS}\rho U_{PS}^\dagger - \rho = & \sin^2(gt) (S\rho S - \rho) \\ & + i \sin(gt) \cos(gt) (\rho S - S\rho). \end{aligned} \quad (\text{E3})$$

Using this and Eqs. (46) and (47) in Eq. (E1), we obtain

$$\begin{aligned} \langle \mathcal{Q}_A \rangle = & \text{tr} \left(\sum_{k=0}^{d-1} \omega_k |k\rangle_A \langle k|_A \otimes \mathbb{1}_B \left[\sin^2(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (S|n,m\rangle \langle n',m'| S - |n,m\rangle \langle n',m'|) \right. \right. \\ & \left. \left. + i \sin(gt) \cos(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (|n,m\rangle \langle n',m'| S - S|n,m\rangle \langle n',m'|) \right] \right) \\ = & \text{tr} \left(\sum_{k=0}^{d-1} \omega_k |k\rangle_A \langle k|_A \otimes \mathbb{1}_B \left[\sin^2(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (|m,n\rangle \langle m',n'| - |n,m\rangle \langle n',m'|) \right. \right. \\ & \left. \left. + i \sin(gt) \cos(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (|n,m\rangle \langle m',n'| - |m,n\rangle \langle n',m'|) \right] \right), \end{aligned} \quad (\text{E4})$$

where in the last equality we have just used $S|a, b\rangle = |b, a\rangle$ and $\langle a, b|S = \langle b, a|$. Continuing with the computation, applying $\sum_k \omega_k |k\rangle_A \langle k|_A \otimes \mathbb{1}_B$ in the operators at its right and summing over the Kronecker deltas, we have

$$\begin{aligned} \langle \mathcal{Q}_A \rangle &= \text{tr} \left(\sin^2(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (\omega_m |m, n\rangle \langle m', n'| - \omega_n |n, m\rangle \langle n', m'|) \right. \\ &\quad \left. + i \sin(gt) \cos(gt) \sum_{n,m,n',m'} p_{n,m,n',m'} (\omega_n |n, m\rangle \langle m', n'| - \omega_m |m, n\rangle \langle n', m'|) \right) \\ &= \sin^2(gt) \sum_{n,m} p_{n,m,n,m} (\omega_m - \omega_n) + i \sin(gt) \cos(gt) \sum_{n,m} p_{n,m,m,n} (\omega_n - \omega_m), \end{aligned}$$

where in the last equality we have used the fact that $\text{tr}(|a, b\rangle \langle c, d|) = \delta_{a,c} \delta_{b,d}$. The above equation is the same as Eq. (48): this becomes clear if we use the marginals $p_{n,n}^A = \sum_m p_{n,m,n,m}$ and $p_{m,m}^B = \sum_n p_{n,m,n,m}$ in the first term on the right-hand side of the equation. For the second term, we use the fact that ρ is Hermitian. Therefore, $p_{n,m,m,n}^* = p_{m,n,n,m}$ and this implies that

$$\begin{aligned} \sum_{n,m} p_{n,m,m,n} (\omega_n - \omega_m) &= \sum_{n,m} \omega_n (p_{n,m,m,n} - p_{n,m,m,n}^*) \\ &= 2i \sum_{n,m} \omega_n \text{Im}(p_{n,m,m,n}), \end{aligned} \quad (\text{E5})$$

where in the first equality we have used the fact that $\sum_{n,m} p_{n,m,m,n} \omega_m = \sum_{m,n} p_{m,n,n,m} \omega_n$.

APPENDIX F: RELEVANT PARAMETERS FOR THE TWO QUTRITS AND PARTIAL SWAP INTERACTIONS CASE

In Sec. III B, we have employed local thermal states. This implies that the diagonal terms p_i ($i = 0, 1, \dots, 8$) correspond to those in the thermal-density matrix $\rho_{3 \otimes 3}$, where $\rho_J = (e^{-\beta_J H_A} / Z_J)$, with $H_A = H_B$ given by Eq. (56). Consequently, we can explicitly express p_i as follows:

$$\begin{aligned} p_0 &= \frac{e^{-(\beta_A + \beta_B)\omega_0}}{Z_A Z_B}, & p_1 &= \frac{e^{(-\beta_A \omega_0 - \beta_B \omega_1)}}{Z_A Z_B}, \\ p_2 &= \frac{e^{(-\beta_A \omega_0 - \beta_B \omega_2)}}{Z_A Z_B}, & p_3 &= \frac{e^{(-\beta_B \omega_0 - \beta_A \omega_1)}}{Z_A Z_B}, \\ p_4 &= \frac{e^{-(\beta_A + \beta_B)\omega_1}}{Z_A Z_B}, & p_5 &= \frac{e^{(-\beta_A \omega_1 - \beta_B \omega_2)}}{Z_A Z_B}, \\ p_6 &= \frac{e^{(-\beta_B \omega_0 - \beta_A \omega_2)}}{Z_A Z_B}, & p_7 &= \frac{e^{(-\beta_B \omega_1 - \beta_A \omega_2)}}{Z_A Z_B}, \\ p_8 &= \frac{e^{-(\beta_A + \beta_B)\omega_2}}{Z_A Z_B}. \end{aligned}$$

Utilizing these results in Eq. (58), we obtain that

$$\langle \mathcal{Q}_A \rangle = \zeta \sin^2(gt) + \xi \sin(gt) \cos(gt), \quad (\text{F1})$$

where ζ is given by

$$\begin{aligned} \zeta &= -\omega_0 (p_3 + p_6 - p_1 - p_2) + \omega_1 (p_1 - p_3 - p_5 + p_7) \\ &\quad + \omega_2 (p_2 + p_5 - p_6 - p_7), \end{aligned}$$

indicating that ζ is solely a function of the energies ω_i and temperatures β_J . In contrast,

$$\begin{aligned} \xi &= \eta_{31} (\omega_2 - \omega_1) \sin(\theta_{31}) + \eta_{62} (\omega_3 - \omega_1) \sin(\theta_{62}) \\ &\quad + \eta_{75} (\omega_3 - \omega_2) \sin(\theta_{75}) \end{aligned}$$

encodes the correlations of the system, represented by η_{ij} and phases θ_{ij} . Notably, if we select three states separated by the same quantum number, such that $\omega_1 - \omega_0 = \omega_2 - \omega_1 \equiv \Delta\omega$, and set the phases $\theta_{ij} = \pi/2$, we achieve the simplest form

$$\xi = \Delta\omega (\eta_{31} + 2\eta_{62} + \eta_{75}),$$

which indicates that the inversion of heat flux is strictly determined by negative values of the correlations η_{ij} .

We have also computed the first-order term, in gt , for the change of the mutual information (the higher-order terms are too large and give no additional information for our purpose):

$$\Delta \mathcal{I}(A : B) = 2gt(\beta_A - \beta_B)\xi + \mathcal{O}(g^2 t^2). \quad (\text{F2})$$

This term is, as in the two-qubit case, equal to $(\beta_A - \beta_B)$ times the first-order term of the heat average of Eq. (F1). It comes from the heat contribution for the change of the mutual information of Eq. (D1) and is the leading-order term causing the initial negativity in the change of the mutual information.

- [1] H. Callen, *Thermodynamics and an Introduction to Thermostatistics*, Student Edition, (Wiley India Pvt. Limited, 2006).
- [2] S. Blundell and K. Blundell, *Concepts in Thermal Physics* (Oxford University Press, Oxford, 2010).
- [3] A. Eddington, *The Nature of the Physical World* (Creative Media Partners, LLC, 2018).
- [4] D. Jennings and T. Rudolph, Entanglement and the thermodynamic arrow of time, *Phys. Rev. E* **81**, 061130 (2010).
- [5] V. Vedral, The Arrow of Time and Correlations in Quantum Physics, [arXiv:1605.00926](https://arxiv.org/abs/1605.00926) [quant-ph].
- [6] P. Lipka-Bartosik, G. F. Diotallevi, and P. Bakhshinezhad, Fundamental limits on anomalous energy flows in correlated quantum systems, *Phys. Rev. Lett.* **132**, 140402 (2024).
- [7] H. Leff and A. Rex, *Maxwell's Demon: Entropy, Information, Computing* (Adam Hilger, 1990).
- [8] H. Leff and A. Rex, *Maxwell's Demon 2 Entropy, Classical and Quantum Information, Computing* (CRC Press, 2002).
- [9] A. Rex, Maxwell's demon—a historical review, *Entropy* **19**, 240 (2017).
- [10] S. Lloyd, Use of mutual information to decrease entropy: Implications for the second law of thermodynamics, *Phys. Rev. A* **39**, 5378 (1989).
- [11] M. H. Partovi, Entanglement versus *Stosszahlansatz*: Disappearance of the thermodynamic arrow in a high-correlation environment, *Phys. Rev. E* **77**, 021110 (2008).
- [12] M. N. Bera, A. Riera, M. Lewenstein, and A. Winter, Generalized laws of thermodynamics in the presence of correlations, *Nat. Commun.* **8**, 2180 (2017).
- [13] S. Jevtic, D. Jennings, and T. Rudolph, Maximally and minimally correlated states attainable within a closed evolving system, *Phys. Rev. Lett.* **108**, 110403 (2012).
- [14] A. Levy and M. Lostaglio, Quasiprobability distribution for heat fluctuations in the quantum regime, *PRX Quantum* **1**, 010309 (2020).
- [15] S. Jevtic, T. Rudolph, D. Jennings, Y. Hirono, S. Nakayama, and M. Murao, Exchange fluctuation theorem for correlated quantum systems, *Phys. Rev. E* **92**, 042113 (2015).
- [16] T. Sagawa and M. Ueda, Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics, *Phys. Rev. Lett.* **109**, 180602 (2012).
- [17] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola, Experimental observation of the role of mutual information in the nonequilibrium dynamics of a Maxwell demon, *Phys. Rev. Lett.* **113**, 030601 (2014).
- [18] J. Cai, Y.-J. Xia, and Z.-X. Man, Reversal of heat flow and extraction of work by means of initial correlations within open quantum systems, *Phys. Lett. A* **501**, 129389 (2024).
- [19] I. Henao and R. M. Serra, Role of quantum coherence in the thermodynamics of energy transfer, *Phys. Rev. E* **97**, 062105 (2018).
- [20] K. Micadei, J. P. S. Peterson, A. M. Souza, R. S. Sarthour, I. S. Oliveira, G. T. Landi, T. B. Batalhão, R. M. Serra, and E. Lutz, Reversing the direction of heat flow using quantum correlations, *Nat. Commun.* **10**, 2456 (2019).
- [21] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, *Phys. Rev. A* **71**, 052108 (2005).
- [22] R. W. Spekkens, The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein, [arXiv:1909.04628](https://arxiv.org/abs/1909.04628) [physics.hist-ph].
- [23] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, Emergence of noncontextuality under quantum Darwinism, *PRX Quantum* **2**, 030351 (2021).
- [24] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-A. Larsson, Kochen-Specker contextuality, *Rev. Mod. Phys.* **94**, 045007 (2022).
- [25] R. Kunjwal, Contextuality beyond the Kochen-Specker theorem, [arXiv:1612.07250](https://arxiv.org/abs/1612.07250) [quant-ph].
- [26] M. S. Leifer and O. J. E. Maroney, Maximally epistemic interpretations of the quantum state and contextuality, *Phys. Rev. Lett.* **110**, 120401 (2013).
- [27] M. Lostaglio, Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response, *Phys. Rev. Lett.* **125**, 230603 (2020).
- [28] R. W. Spekkens, Negativity and contextuality are equivalent notions of nonclassicality, *Phys. Rev. Lett.* **101**, 020401 (2008).
- [29] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, A structure theorem for generalized-noncontextual ontological models, *Quantum* **8**, 1283 (2024).
- [30] D. Schmid, R. D. Baldijão, Y. Ying, R. Wagner, and J. H. Selby, Kirkwood-Dirac representations beyond quantum states (and their relation to noncontextuality), [arXiv:2405.04573](https://arxiv.org/abs/2405.04573) [quant-ph].
- [31] M. Khoshbin, L. Catani, and M. Leifer, Alternative robust ways of witnessing nonclassicality in the simplest scenario, *Phys. Rev. A* **109**, 032212 (2024).
- [32] J. H. Selby, E. Wolfe, D. Schmid, A. B. Sainz, and V. P. Rossi, Linear program for testing nonclassicality and an open-source implementation, *Phys. Rev. Lett.* **132**, 050202 (2024).
- [33] V. P. Rossi, D. Schmid, J. H. Selby, and A. B. Sainz, Contextuality with vanishing coherence and maximal robustness to dephasing, *Phys. Rev. A* **108**, 032213 (2023).
- [34] A. M. Fonseca, V. P. Rossi, R. D. Baldijão, J. H. Selby, and A. B. Sainz, Robustness of contextuality under different types of noise as quantifiers for parity-oblivious multiplexing tasks, [arXiv:2406.12773](https://arxiv.org/abs/2406.12773) [quant-ph].
- [35] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, *Nat. Commun.* **7**, 1 (2016).
- [36] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories, *PRX Quantum* **2**, 020302 (2021).
- [37] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi, N. Spagnolo, R. Osellame, E. F. Galvão, and F. Sciarrino, Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor, *Sci. Adv.* **9**, eadj4249 (2023).
- [38] C. Duarte and B. Amaral, Resource theory of contextuality for arbitrary prepare-and-measure experiments, *J. Math. Phys.* **59**, 062202 (2018).
- [39] R. Wagner, R. D. Baldijão, A. Tezzin, and B. Amaral, Using a resource theoretic perspective to witness and engineer quantum generalized contextuality for

- prepare-and-measure scenarios, *J. Phys. A: Math. Theor.* **56**, 505303 (2023).
- [40] L. Catani, T. D. Galley, and T. Gonda, Resource-theoretic hierarchy of contextuality for general probabilistic theories, [arXiv:2406.00717 \[quant-ph\]](#).
- [41] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, *Phys. Rev. A* **75**, 032110 (2007).
- [42] L. Hardy, Disentangling nonlocality and teleportation, [arXiv:quant-ph/9906123](#).
- [43] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, Why interference phenomena do not capture the essence of quantum theory, *Quantum* **7**, 1119 (2023).
- [44] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction, *Phys. Rev. A* **86**, 012103 (2012).
- [45] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey, Uniqueness of noncontextual models for stabilizer subtheories, *Phys. Rev. Lett.* **129**, 120403 (2022).
- [46] S. Deffner and S. Campbell, Quantum thermodynamics: An introduction to the thermodynamics of quantum information, [arXiv:1907.01596 \[quant-ph\]](#).
- [47] F. Binder, L. Correa, C. Gogolin, J. Anders, and G. Adesso, *Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions*, Fundamental Theories of Physics (Springer International Publishing, 2019).
- [48] J. Goold, M. Huber, A. Riera, L. d. Rio, and P. Skrzypczyk, The role of quantum information in thermodynamics—a topical review, *J. Phys. A: Math. Theor.* **49**, 143001 (2016).
- [49] S. Vinjanampathy and J. Anders, Quantum thermodynamics, *Contemp. Phys.* **57**, 545 (2016).
- [50] R. Alicki and R. Kosloff, Introduction to quantum thermodynamics: History and prospects, [arXiv:1801.08314 \[quant-ph\]](#).
- [51] R. Kosloff, Quantum thermodynamics: A dynamical viewpoint, *Entropy* **15**, 2100 (2013).
- [52] J. Millen and A. Xuereb, Perspective on quantum thermodynamics, *New J. Phys.* **18**, 011002 (2016).
- [53] T. Upadhyaya, J. a. William F. Braasch, G. T. Landi, and N. Y. Halpern, What happens to entropy production when conserved quantities fail to commute with each other, [arXiv:2305.15480 \[quant-ph\]](#).
- [54] M. Lostaglio, Quantum fluctuation theorems, contextuality, and work quasiprobabilities, *Phys. Rev. Lett.* **120**, 040602 (2018).
- [55] M. Lostaglio, An introductory review of the resource theory approach to thermodynamics, *Rep. Prog. Phys.* **82**, 114001 (2019).
- [56] G. T. Landi and M. Paternostro, Irreversible entropy production: From classical to quantum, *Rev. Mod. Phys.* **93**, 035008 (2021).
- [57] R. Clausius, Ueber eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie, *Ann. Phys.* **169**, 481 (1854).
- [58] R. Clausius, Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie, *Ann. Phys.* **201**, 353 (1865).
- [59] N. Harrigan and R. W. Spekkens, Einstein, incompleteness, and the epistemic view of quantum states, *Found. Phys.* **40**, 125 (2010).
- [60] R. Kunjwal, Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements, *Quantum* **3**, 184 (2019).
- [61] D. Schmid, J. H. Selby, and R. W. Spekkens, Unscrambling the omelette of causation and inference: The framework of causal-inferential theories, [arXiv:2009.03297 \[quant-ph\]](#).
- [62] M. Banik, S. S. Bhattacharya, S. K. Choudhary, A. Mukherjee, and A. Roy, Ontological models, preparation contextuality and nonlocality, *Found. Phys.* **44**, 1230 (2014).
- [63] D. Schmid, R. W. Spekkens, and E. Wolfe, All the non-contextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, *Phys. Rev. A* **97**, 062103 (2018).
- [64] R. Wagner, R. S. Barbosa, and E. F. Galvão, Inequalities witnessing coherence, nonlocality, and contextuality, *Phys. Rev. A* **109**, 032220 (2024).
- [65] M. F. Pusey, Robust preparation noncontextuality inequalities in the simplest scenario, *Phys. Rev. A* **98**, 022112 (2018).
- [66] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
- [67] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
- [68] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation contextuality powers parity-oblivious multiplexing, *Phys. Rev. Lett.* **102**, 010401 (2009).
- [69] D. Saha and A. Chaturvedi, Preparation contextuality as an essential feature underlying quantum communication advantage, *Phys. Rev. A* **100**, 022108 (2019).
- [70] D. Schmid and R. W. Spekkens, Contextual advantage for state discrimination, *Phys. Rev. X* **8**, 011015 (2018).
- [71] K. Flatt, H. Lee, C. R. I. Carceller, J. B. Brask, and J. Bae, Contextual advantages and certification for maximum-confidence discrimination, *PRX Quantum* **3**, 030337 (2022).
- [72] S. Mukherjee, S. Naonit, and A. K. Pan, Discriminating three mirror-symmetric states with a restricted contextual advantage, *Phys. Rev. A* **106**, 012216 (2022).
- [73] M. Lostaglio and G. Senno, Contextual advantage for state-dependent cloning, *Quantum* **4**, 258 (2020).
- [74] R. Wagner, A. Camillini, and E. F. Galvão, Coherence and contextuality in a Mach-Zehnder interferometer, *Quantum* **8**, 1240 (2024).
- [75] P. Lillystone, J. J. Wallman, and J. Emerson, Contextuality and the single-qubit stabilizer subtheory, *Phys. Rev. Lett.* **122**, 140405 (2019).
- [76] M. F. Pusey, Anomalous weak values are proofs of contextuality, *Phys. Rev. Lett.* **113**, 200401 (2014).
- [77] R. Kunjwal, M. Lostaglio, and M. F. Pusey, Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts, *Phys. Rev. A* **100**, 042116 (2019).
- [78] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Contextuality without incompatibility, *Phys. Rev. Lett.* **130**, 230201 (2023).
- [79] J. Miyazaki and K. Matsumoto, Imaginarity-free quantum multiparameter estimation, *Quantum* **6**, 665 (2022).
- [80] H. E. D. Scovil and E. O. Schulz-DuBois, Three-level masers as heat engines, *Phys. Rev. Lett.* **2**, 262 (1959).

- [81] J. C. Garcia-Escartin and P. Chamorro-Posada, A SWAP gate for qudits, [Quantum Inf. Process.](#) **12**, 3625 (2013).
- [82] C. M. Wilmott, On swapping the states of two qudits, [Int. J. Quantum Inf.](#) **09**, 1511 (2011).
- [83] R. Wagner, Z. Schwartzman-Nowik, I. L. Paiva, A. Te'eni, A. Ruiz-Molero, R. S. Barbosa, E. Cohen, and E. F. Galvão, Quantum circuits for measuring weak values, Kirkwood-Dirac quasiprobability distributions, and state spectra, [Quantum Sci. Technol.](#) **9**, 015030 (2024).
- [84] M. Lostaglio, A. Belenchia, A. Levy, S. Hernández-Gómez, N. Fabbri, and S. Gherardini, Kirkwood-Dirac quasiprobability approach to the statistics of incompatible observables, [Quantum](#) **7**, 1128 (2023).
- [85] D. R. M. Arvidsson-Shukur, W. F. Braasch Jr., S. D. Bievre, J. Dressel, A. N. Jordan, C. Langrenez, M. Lostaglio, J. S. Lundeen, and N. Y. Halpern, Properties and applications of the Kirkwood-Dirac distribution, [arXiv:2403.18899 \[quant-ph\]](#).
- [86] N. Y. Halpern, B. Swingle, and J. Dressel, Quasiprobability behind the out-of-time-ordered correlator, [Phys. Rev. A](#) **97**, 042105 (2018).
- [87] S. Gherardini and G. D. Chiara, Quasiprobabilities in quantum thermodynamics and many-body systems: A tutorial, [arXiv:2403.17138 \[quant-ph\]](#).
- [88] H. Margenau and R. N. Hill, Correlation between measurements in quantum theory, [Prog. Theor. Phys.](#) **26**, 722 (1961).
- [89] Y. P. Terletsy, The limiting transition from quantum to classical mechanics, [J. Exp. Theor. Phys.](#) **7**, 1290 (1937).
- [90] J. Rau, Relaxation phenomena in spin and harmonic oscillator systems, [Phys. Rev.](#) **129**, 1880 (1963).
- [91] S. Campbell and B. Vacchini, Collision models in open system dynamics: A versatile tool for deeper insights? [Europhys. Lett.](#) **133**, 60001 (2021).
- [92] F. Ciccarello, S. Lorenzo, V. Giovannetti, and G. M. Palma, Quantum collision models: Open system dynamics from repeated interactions, [Phys. Rep.](#) **954**, 1 (2022).
- [93] V. Scarani, M. Ziman, P. Štelmachovič, N. Gisin, and V. Bužek, Thermalizing quantum machines: Dissipation and entanglement, [Phys. Rev. Lett.](#) **88**, 097905 (2002).
- [94] M. Ziman, P. Štelmachovič, V. Bužek, M. Hillery, V. Scarani, and N. Gisin, Diluting quantum information: An analysis of information transfer in system-reservoir interactions, [Phys. Rev. A](#) **65**, 042105 (2002).
- [95] S. Cusumano, Quantum collision models: A beginner guide, [Entropy](#) **24**, 1258 (2022).
- [96] N. E. Comar and G. T. Landi, Correlations breaking homogenization, [Phys. Rev. A](#) **104**, 032217 (2021).
- [97] S. Campbell, B. Cakmak, O. E. Mustecaplıoğlu, M. Paternostro, and B. Vacchini, Collisional unfolding of quantum Darwinism, [Phys. Rev. A](#) **99**, 042103 (2019).