

Thermodynamics of readout devices and semiclassical gravity

Samuel Fedida¹ and Adrian Kent^{1,2}

¹*Centre for Quantum Information and Foundations, DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

²*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*



(Received 10 September 2025; accepted 16 December 2025; published 12 January 2026)

We analyze the common claim that nonlinear modifications of quantum theory necessarily violate the second law of thermodynamics. We focus on hypothetical extensions of quantum theory that contain readout devices. These black boxes provide a classical description of quantum states without perturbing them. They allow quantum state cloning, though in a way consistent with the relativistic no-signalling principle. We review the existence of such devices in the context of Møller-Rosenfeld semiclassical gravity, which postulates that the gravitational field remains classical and is sourced by the expectation value of a quantum energy-momentum tensor. We show that the definition of information in the models examined in this paper deviates from that given by von Neumann entropy, and that claims of second law violations based on the distinguishability of nonorthogonal states or on violations of uncertainty principles fail to hold in such theories.

DOI: [10.1103/bpwh-1pkf](https://doi.org/10.1103/bpwh-1pkf)

I. INTRODUCTION

In standard quantum mechanics, the act of measurement extracts classical information from a quantum state while disturbing the latter, with bounds on the tradeoff between information obtained and disturbance. Nonlinear modifications to the Schrödinger equation can, in principle, violate these bounds. For example, quantum theory can be consistently extended by introducing readout devices [1], which are idealized black boxes capable of extracting certain properties of a quantum state without inducing wave-function collapse or otherwise altering the state. One theoretical motivation for considering such extensions is that Møller-Rosenfeld semi-classical gravity effectively introduces finite-precision versions of such devices.

Møller-Rosenfeld semi-classical gravity is based on the hypothesis that spacetime is classical and couples to quantum matter through the expectation value of the energy-momentum tensor of matter. The semi-classical field equations then read

$$G_{\mu\nu} = \frac{8\pi G}{c^4} \langle \hat{T}_{\mu\nu} \rangle. \quad (1)$$

It is not evident that Eq. (1) defines a complete consistent theory. If it does, it is inconsistent with astronomical observation, assuming that there is a universal matter wave function in a superposition significantly different from the observed matter state. It is also inconsistent with table-top experiments [2]. However, neither of these rule out Eq. (1) being valid as an

effective limit [3,4] within a restricted domain, or as part of a theory that modifies standard quantum dynamics (e.g., [5,6]).

One argument [7] often made against nonlinear theories such as semi-classical gravity is that they necessarily introduce superluminal signaling. However, there is a quite general construction that can be used in such theories to avoid superluminal signaling [8].

It has also been argued [9] that the measurement postulates of quantum theory follow from other essential defining postulates of quantum theory, which nonlinear versions of quantum theory, which necessarily introduce nonstandard measurements, must thus violate. These arguments were originally framed with several tacit assumptions [10,11]. When made explicit, careful analysis [9–11] shows that nonlinear theories of the type discussed in [8,10] extend quantum theory by introducing new features in new regimes, without breaking any principle required for the success of quantum theory in experiments to date.

It is also worth noting that, while a semi-classical gravity theory would violate the quantum no-signalling principle (which we stress is distinct from the no-superluminal signalling principle), it also provides an obvious potential mechanism for the relevant light-speed signals, namely, the gravitational degrees of freedom. That said, one might also ask, even in the absence of such a mechanism, whether we should be so surprised by a physical theory in which actions on one subsystem have causal effects on another, given that quantum theory itself includes Bell nonlocal correlations. We think that, for example, Einstein, Podolsky, and Rosen [12] might have found the former less problematic than the latter.

From the perspective of axiomatic relativistic quantum field theory, it also appears that relativistic no-signalling should be understood as an epistemic principle rather than an ontological one, i.e., taking interactions to be local in some ontological sense appears not only unnecessary but also

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

unrealistic. For example, Wightman's theorem [13,14] states that imposing pointwise microcausality (i.e., $[\hat{\phi}(x)\hat{\phi}(y)] = 0$ whenever x and y are spacelike separated points) is too strong for theories of translation-covariant quantum fields which satisfy the spectrum condition (the Hamiltonian is bounded from below) provided there exists a unique translation-invariant vacuum state: it implies that all n -point vacuum expectation values are constant over spacetime. Instead, weaker notions such as Einstein causality (measurements conducted over spacelike separated regions commute) are usually assumed.

These various points give some extra motivation, if any were needed, to the active ongoing experimental efforts (e.g., [15–19]) to determine whether gravity is quantum since models involving semi-classical gravity in some relevant regimes give a concrete alternative that the relevant experiments can distinguish from quantum gravity.

However, there are other independent arguments made against nonlinear theories such as semi-classical gravity, arising from claims that they violate the fundamental principles of thermodynamics, and in particular, the second law. These arguments go back to von Neumann [20] and were elaborated on by Peres [21]. More recently, Hänggi and Wehner showed [22] that, under certain assumptions, violations of some versions of the uncertainty principle implied a violation of the second law of thermodynamics. Nonlinear quantum theories can have [23,24] a classical state space in which nonorthogonal states can be perfectly distinguished, as readout devices [8] illustrate, and so evidently violate standard uncertainty principles. The arguments of the authors of [20–22] thus suggested that they indeed violate the second law and are implausible for this reason. Such lines of thought have been taken (e.g., [25]) to disfavor classical-quantum hybrid models based on nonlinear dynamics.

These arguments are also significant because, regardless of its fundamental status, semi-classical gravity is often taken as a reasonable approximation in many contexts, including for the Hawking radiation of black holes and for perturbation theory in inflationary cosmology, where one considers quantum field theory (QFT) in curved but classical spacetime. In these contexts, understanding the thermodynamic and information theoretic properties of the relevant systems is key.

It should be noted here that the predictions of Møller-Rosenfeld semi-classical gravity deviate in these regimes from those of the semi-classical limit of quantum gravity [26]. Our discussion in the present paper is relevant to the former. It might, depending on the reader's perspective and notwithstanding our conclusions, reinforce the view that the correct approximation is given by the latter. This is important to highlight to distinguish nuances in the literature.

For example, in black hole complementarity [27], deriving a form of quantum state cloning from the assumption that semi-classical gravitational equations hold alongside some fundamental unitarity is not the same as deriving cloning from assuming Møller-Rosenfeld semi-classical gravity. The latter is a straightforward consequence of the fact that Møller-Rosenfeld semi-classical gravity allows the construction of quantum state readout devices [1,26].

In this paper, we consider whether semi-classical gravity and, more generally, nonlinear extensions of quantum theory with readout devices (which we review in the next section) are consistent with thermodynamics. We show that a careful treatment need not contradict appropriately formulated thermodynamic laws.

Interestingly, a second law of thermodynamics can be justified for generalized probabilistic theories [28] and for models combining classical and quantum mechanics [29]. The models analyzed in [28] satisfy two postulates, which do not typically hold in the class of models we consider here. The authors of [29] analyzed models with a specific stochastic coupling between classical and quantum degrees of freedom. The models discussed in [28,29] are linear; we consider nonlinear extensions of quantum theory.

II. READOUT DEVICES

We begin by reviewing the definitions of and some motivations for introducing *readout devices* [1,8,10,19]. These are hypothetical idealized devices that give us classical information about a pure quantum state $|\psi\rangle \in \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_N$ of a set of N systems. The devices we consider cannot be constructed using standard quantum measurements, so considering them implies considering extensions of standard quantum theory.

We will take the relevant $\mathcal{H}_i$, namely, those whose degrees of freedom the devices give information about, to be finite-dimensional Hilbert spaces. Finite precision (FP) models of such devices may also be described. In one simple model, an FP readout device takes a positive integer l as additional input, and outputs the associated information to l digits of binary precision. When the input is a pure state, the action of the FP readout device may be defined either on the ray in Hilbert space or on the corresponding density matrix: the finite precision versions of these are approximately but not generally exactly equivalent.

We define a *state readout device* (RD) to act on subsystems: given subsystems $1, \dots, i \leq n$ of the state $|\psi\rangle$ of a system S it outputs a classical description (in some given basis) of $\rho_{1\dots i} = \text{Tr}_{i+1,\dots,N}(|\psi\rangle\langle\psi|) \in \mathcal{L}(\mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_i)$ where $\mathcal{L}(\mathcal{H}_j)$ denotes the space of linear operators on $\mathcal{H}_j$ for $j \in \{1, \dots, n\}$.¹ That is

$$l, \text{ basis}, \rho_{1\dots i} \xrightarrow{\text{input}} \boxed{\text{(FP)RD}} \xrightarrow{\text{output}} \begin{cases} a_{11} = 0n_{11}^{(1)}n_{11}^{(2)}\dots n_{11}^{(l)}, \\ a_{12} = 0n_{12}^{(1)}n_{12}^{(2)}\dots n_{12}^{(l)} + i0p_{12}^{(1)}p_{12}^{(2)}\dots p_{12}^{(l)}, \\ \dots \\ a_{k-1,k} = 0n_{k-1,k}^{(1)}n_{k-1,k}^{(2)}\dots n_{k-1,k}^{(l)} + i0p_{k-1,k}^{(1)}p_{k-1,k}^{(2)}\dots p_{k-1,k}^{(l)}, \\ a_{kk} = 0n_{kk}^{(1)}n_{kk}^{(2)}\dots n_{kk}^{(l)}, \end{cases} \quad (2)$$

¹In [10] an RD is defined to act on subsystem 1. The definition here is equivalent, up to the redefinition of subsystems, but more convenient for our discussion.

where $k = \prod_{j=1}^i \dim \mathcal{H}_j$. The density matrix $\rho_{1,\dots,i}$ is parameterized by $k^2 - 1$ real parameters, so in the finite precision case ($l < \infty$), this requires $\approx(k^2 - 1)l$ bits of (classical) information storage.

A *probability readout device* (PRD) outputs a classical description of the diagonal elements (Born probabilities) of $\rho_{1,\dots,i}$ in some given basis

$$l, \text{ basis}, \rho_{1,\dots,i} \xrightarrow{\text{input}} \boxed{\text{(FP)PRD}} \xrightarrow{\text{output}} \begin{cases} |a_{11}|^2 = 0n_{11}^{(1)}n_{11}^{(2)} \dots n_{11}^{(l)}, \\ |a_{22}|^2 = 0n_{22}^{(1)}n_{22}^{(2)} \dots n_{22}^{(l)}, \\ \dots \\ |a_{kk}|^2 = 0n_{kk}^{(1)}n_{kk}^{(2)} \dots n_{kk}^{(l)}. \end{cases} \quad (3)$$

In the finite-precision case, this requires $\approx(k - 1)l$ bits of (classical) information storage. Here, note that we only have information about the probabilities not the amplitudes or their relative phases. Using quantum state tomography, the amplitudes (up to an overall phase) can be obtained by repeated use of a PRD, but the information-theoretic analysis is more involved and not relevant for our discussion here.

An *expectation value readout device* (ERD) takes as input some hermitian observable A defined on systems $1 \dots i$, and outputs the expectation value $\text{Tr}(A\rho_{1,\dots,i})$. In the infinite precision case, the expectation values can be obtained from PRDs plus postprocessing when the PRDs work in the eigenbasis of the observable. In the finite precision case, this is approximately true for sufficiently large l , but the detailed propagation of the precision is more subtle and depends on the observable and postprocessing.

Note that an infinite precision RD is more powerful than a perfect cloning device on a pure state with $i = N$, assuming that perfect infinite precision classical and quantum operations are free resources. Reading $|\Psi\rangle$ gives classical information about that state that we may encode in a quantum state in some basis. That is, starting with a blank ancilla state $|*\rangle$, we can implement

$$|\Psi\rangle|*\rangle \longrightarrow \boxed{\text{RD \& enc.}} \longrightarrow |\Psi\rangle|\Psi\rangle, \quad (4)$$

up to a global phase, so that one use of an infinite precision RD allows perfect cloning. However, a perfect cloning device would need to be used uncountably infinitely often, using uncountably many ancillae (which goes beyond standard assumptions about the physically accessible states in quantum field theory) to generate the quantum information required to produce an infinite precision classical description.

Even if quantum theory is not fundamental, it seems hard to imagine that infinite precision state readout devices (or infinite precision versions of the other devices) could be operationally realized in nature since the infinite precision outputs cannot be expressed in a finite time. These devices should rather be thought of as convenient mathematical idealizations.

At first glance, it may also seem hard to motivate the hypothesis that anything resembling a FPRD could be found in nature, given that FPRDs define nonstandard extensions of standard quantum theory that violate quantum no-signalling and no-cloning, albeit without necessarily violating relativistic causality [8]. However, the possibility that semiclassical gravity holds true in some regime provides an explicit example of a type of theory, considered for independent reasons, that would imply FPRDs [1,26]. More generally, nonlinear

extensions of quantum theory generically imply some form of FPRD.

One might take this as an argument against semi-classical gravity and other nonlinear theories. However, to make a compelling argument one needs to show that extensions of quantum theory involving readout devices necessarily have incurable problems. Perhaps they do, but our aim in this paper is to show that the thermodynamic arguments often cited in this context do not identify such problems.

A. Example: ERDs and PRDs in semiclassical gravity

The standard assumption in the context of semi-classical gravity is that the gravitational field could be measured without affecting the quantum matter state. This would imply that a device measuring the classical gravitational field at any point would act as a finite-precision ERD. If this held true, then versions of FPRDs would exist and could, in principle, be constructed [1] in the relevant regime, as we now review. We can simplify the information theoretic analysis by looking at the nonrelativistic limit of the semi-classical Einstein field equations (1), noting again that the construction of readout devices can be defined to avoid superluminal signaling [8].

In this limit, we can define the semi-classical Newtonian potential to follow Poisson's equation

$$\Delta \Phi(|\psi|^2; \mathbf{x}) = 4\pi G \langle \hat{\rho}(\mathbf{x}) \rangle, \quad (5)$$

where $\hat{\rho}(x)$ is the mass density operator of the quantum matter, so that the Schrödinger equation has an extra term from the gravitational Hamiltonian [30] and becomes the Schrödinger-Newton equation (SNE) [31]

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \left(\hat{H}_{\text{matter}} + \int \hat{\rho}(\mathbf{x}) \Phi(|\psi|^2; \mathbf{x}) d^3\mathbf{x} \right) |\psi(t)\rangle. \quad (6)$$

For a single particle of mass m in a superposition of localized positions with quantum state

$$|\psi\rangle = a_0|\mathbf{x}_0\rangle + a_1|\mathbf{x}_1\rangle, \quad (7)$$

the semi-classical gravitational field at some point $\mathbf{y}$ can be found from Poisson's equation (5) and is given by

$$\Phi(\mathbf{y}) = -Gm \left(\frac{|a_0|^2}{|\mathbf{x}_0 - \mathbf{y}|} + \frac{|a_1|^2}{|\mathbf{x}_1 - \mathbf{y}|} \right). \quad (8)$$

Determining the classical gravitational field at some other point $\mathbf{y}'$ provides information on the state $|\psi\rangle$ by giving us

an estimate of the $|a_i|^2$ for $i = 1, 2$. In fact, given that we can take $|\psi\rangle$ to be normalized, just one reading suffices. More generally, given (for simplicity) a normalized pure quantum state

$$|\Psi\rangle = \sum_{i=1}^k a_i |\mathbf{x}_i\rangle, \quad (9)$$

we can estimate the $|a_i|^2$ from readings of the semi-classical gravitational field

$$\Phi(\mathbf{y}_j) = -Gm \sum_{i=1}^k \frac{|a_i|^2}{|\mathbf{x}_i - \mathbf{y}_j|} \quad (10)$$

at $k - 1$ locations $\mathbf{y}_j$. We expect these readings to be finite precision because experimental constraints, including the uncertainty principle, prevent us both from specifying the $\mathbf{y}_j$ and from measuring $\Phi(\mathbf{y}_j)$ to infinite precision.

Thus, semi-classical gravity allows us, in principle, to construct an (FP)PRD of the form

$$l, \text{ basis, } |\Psi\rangle \xrightarrow{\text{input}} \boxed{\text{Semiclassical gravity(FP)PRD}} \xrightarrow{\text{output}} \begin{cases} |a_1|^2 = 0n_1^{(1)}n_1^{(2)} \dots n_1^{(l)}, \\ |a_2|^2 = 0n_2^{(1)}n_2^{(2)} \dots n_2^{(l)}, \\ \dots \\ |a_k|^2 = 0n_k^{(1)}n_k^{(2)} \dots n_k^{(l)}. \end{cases} \quad (11)$$

In summary, readings of classical gravitational fields would be equivalent to finite precision estimates of the expectation value of the field generated by the quantum state, i.e., to some form of finite precision ERD. Repeated readings of such ERDs would provide some form of finite precision PRD. The finite precision estimates provided would, in general, include noise arising from gravitational fields extrinsic to the system being measured. Thus, the basis and degree of precision would depend on the underlying theory and on the specific physical context, and the degree of precision may vary for different states and different components of a state. The definitions of finite precision devices above thus also need to be understood as mathematical idealizations of the expected behavior of semi-classical gravity and similar models. Nonetheless they capture the essential features of nonstandard measurements available within these models.

As was shown in [32], no-cloning holds for any generalized probabilistic theory (GPT), including quantum theory, apart from classical probability theory. There is no contradiction here since the fact that semi-classical gravity implies readout devices means that it has classical features (see also discussions in [23,24]). This can be understood in the sense of the theory allowing perfect discrimination of nonorthogonal states. Moreover, the notion of quantum state cloning is closely related to the notion of conservation of quantum information. Again, there is no contradiction: we will see that the notion of information changes in the context of semi-classical gravity, and more broadly, in theories with readout devices.

Given these theoretical motivations, it is interesting to understand the thermodynamical behavior of quantum-like theories that include readout devices. Indeed, several arguments in the literature can be interpreted as ruling such theories out on the basis of thermodynamical instability. In this paper, we address these and show this interpretation is not justified.

B. Entropy in readout device world

Extensions of quantum theory allow different definitions of entropy that reduce to von Neumann entropy in the relevant regimes [33]. One such is the so-called *decomposition entropy*, defined for a d -dimensional system as

$$S_{\text{dec}}(\rho) := \inf_{\substack{\{p_j, \sigma_j\} \\ \rho = \sum_{j=1}^d p_j \sigma_j}} H(\{p_1, \dots, p_d\}), \quad (12)$$

where H is the Shannon entropy and the infimum is taken over decompositions into perfectly distinguishable pure states [22].

Decomposition entropy is proven to reduce to the Shannon and von Neumann entropies in classical probability theory and quantum theory in [33]. Although interesting in discussions of entanglement or purity, decomposition entropy does not satisfy concavity nor subadditivity in general theories (e.g., in a box world [33]), and has a less direct operational interpretation than the *measurement entropy*, which we now review.

First, we discuss the notion of *measurements*, and in particular, of *fine-grained measurements* [33]. A measurement is a set

$$\mathbf{e} = \left\{ (j, e_j) \mid j \in J \text{ \& } \sum_{j \in J} e_j = \mathbb{1} \right\}, \quad (13)$$

where J is an indexing set (the set of outcomes of the measurement) and $e_j : \mathcal{D}(\mathcal{H}) \rightarrow [0, 1]$ are maps, called *effects*, such that $e_j(\rho)$ is the probability of obtaining outcome j given state ρ , and $\sum_{j \in J} e_j(\rho) = 1$ for all $\rho \in \mathcal{D}(\mathcal{H})$.

Some measurements on a system can be more informative than others. Consider two measurements $\mathbf{e}$ and $\mathbf{f}$ for which there exists a map $M : J_{\mathbf{e}} \rightarrow J_{\mathbf{f}}$ such that

$$f_{j'} = \sum_{j \in J_{\mathbf{e}}: M(j)=j'} e_j, \quad \forall j' \in J_{\mathbf{f}}. \quad (14)$$

If M is one-to-one then this is just a relabeling of outcomes, otherwise $\mathbf{f}$ is a *coarse-graining* of $\mathbf{e}$ and, conversely, $\mathbf{e}$ is

said to be a *refinement* of $\mathbf{f}$. A coarse-graining is said to be *trivial* if

$$e_j \propto f_{M(j)}, \quad \forall j \in J_e, \quad (15)$$

in which case the measurement $\mathbf{e}$ is equivalent to that of $\mathbf{f}$, obtaining j' and outputting a randomly selected j satisfying $M(j) = j'$. Thus, the two measurements are equally informative about the state of the system. A measurement $\mathbf{e}$ is said to be *fine-grained* if it has no nontrivial refinements. Fine-grained measurements are thus those measurements which cannot be refined to give more information about the state.

The *measurement entropy* $\hat{H}$ of a state ρ is the infimum of the Shannon entropy of the probabilities associated to effects of fine-grained measurements

$$\hat{H}(\rho) := \inf_{\{e_j\}} H(\{e_j(\rho)\}). \quad (16)$$

It has a clear information-theoretic and operational meaning: it is the infimal output uncertainty of fine-grained measurements on the system. A coarse-graining of a fine-grained measurement can never give more information [33].

The definition extends to uncountable J by replacing sums by integrals and using the Shannon entropy of the probability distribution function.

Infinite-precision readout devices correspond to the collection of effects $\{e_\rho\}_{\rho \in \mathcal{D}(\mathcal{H})}$ where

$$e_\rho(\sigma) = \begin{cases} \delta_{\rho, \sigma} & \text{if } \sigma \text{ is improper,} \\ \sum_i p_i \delta_{\rho, \sigma_i} & \text{if } \sigma = \{p_i, \sigma_i\} \text{ is proper.} \end{cases} \quad (17)$$

In the case of infinite precision readout devices, the measurement entropy is generally different to the decomposition entropy. The “best measurements” are given by the read-

out devices: a measurement that provides all the information about a state cannot be further (nontrivially) refined. In particular, the effects of standard quantum measurements can be reproduced by the output of an infinite-precision readout device together with that of an infinite-precision random number generator. Formally, any quantum effect f_j can be written as

$$f_j = \int_{\rho \in \mathcal{D}(\mathcal{H})} p(\rho) e_\rho d\rho, \quad (18)$$

where $p(\rho)$ are appropriate probability weights and $\int d\rho$ is defined so that for any effect g ,

$$\left(\int g(\rho) e_\rho d\rho \right) (\sigma) = g(\sigma). \quad (19)$$

Taking Eq. (18) as a definition of coarse-graining in this setting, standard quantum effects are coarse-grainings of infinite-precision readout devices.

Thus, given any proper mixture ρ_{proper} associated to an ensemble $\{p_i, |\psi_i\rangle\}$,

$$\begin{aligned} S_{RD}(\{p_i, |\psi_i\rangle\}) &:= \hat{H}(\rho_{\text{proper}}) = H(\{p_i\}) \\ &= - \sum_i p_i \log(p_i). \end{aligned} \quad (20)$$

That is, the measurement entropy of an ensemble in a readout device world is always the Shannon entropy associated to the preparation probabilities. However, for improper mixtures, if one works with infinite-precision readout devices, one can fully determine ρ_{improper} with probability 1 since $e_\rho(\rho_{\text{improper}}) = \delta_{\rho, \rho_{\text{improper}}}$. Hence $S_{RD}(\rho_{\text{improper}}) = 0$.

Finite-precision readout devices can also be characterized in the language of effects and fine-grained measurements. In the basis $\{|\chi_i\rangle\} \subset \mathcal{H}$, the effects are $\{e_s\}_{s \in S_\chi(l)}$ where

$$S_\chi(l) = \left\{ \sum_{i,j} a_{ij} |\chi_i\rangle \langle \chi_j| \mid a_{ij} = 0n_{ij}^{(1)} \dots n_{ij}^{(l)} + i0p_{ij}^{(1)} \dots p_{ij}^{(l)} \right\} \quad (21)$$

is the set of allowed finite precision outputs. Then

$$e_s(\sigma) = \begin{cases} 1 & \text{if } \sigma \approx s \text{ is improper,} \\ p_i & \text{if } \sigma = \{p_i, \sigma_i\} \text{ is proper and } \sigma_i \approx s, \\ 0 & \text{otherwise,} \end{cases} \quad (22)$$

where $\approx$ is to l digits of precision. Thus FPRDs can determine an approximation of ρ_{improper} with probability 1 and so $S_{\text{FPRD}}(\rho_{\text{improper}}) = 0$ as well.

Models of finite precision readout devices whose effects have nontrivial probabilities even for improper mixed states, with the readout for a given state having a probability distribution dominated by nearby states, could also be considered. These are more physically plausible if they arise from a physical theory such as semi-classical gravity. However, any finite precision readout device models should replicate infinite-precision readout devices in the limit as the precision becomes large, so we need not fix on any particular finite-precision readout device model here. For simplicity, we shall not consider these probabilistic finite precision readout devices in our

discussion (although they could be incorporated by altering some of the definitions and statements below).

The information theory underlying the thermodynamic analysis of a model with readout devices depends, *inter alia*, on the resources accessible (in principle or in practice), on the precision and type(s) of readout device(s), and on the precision of quantum state preparations and measurements. In this work, we limit ourselves to the consideration of readout devices that are either infinite precision or can be made arbitrarily (finitely) precise. We also limit ourselves to discussing finite ensembles. Under these assumptions, all quantum measurements can be represented (to arbitrary precision) as coarse-grained measurements using readout devices and random number generators (by assumption arbitrarily precise).

In a model with infinite precision readout devices, microsystems can be understood as being classical in the sense that every possible quantum state represents a different and perfectly distinguishable physical state. If we consider a system of quantum states that is allowed to diffuse and is subject

to readout device measurements (but not quantum measurements), it behaves like a classical ideal gas of chemical species. Under these circumstances, a proper mixture of states behaves thermodynamically as a mixture of distinct chemical species, while an improper mixture behaves as a single species. From this perspective, the measurement entropy (16) is clearly a more natural measure than the decomposition entropy (12).

However, the model also includes quantum measurements of quantum states. These have the usual probabilistic properties, and may either increase or decrease the measurement entropy of an ensemble, depending on the ensemble and the measurement choice. Although the possibility of entropy decrease may seem problematic, a full thermodynamic model would need to consider also the entropy of readout devices and quantum measurement devices.

We will not fix a particular model of quantum measurement in a readout device world. It is enough for our discussion below to note that, in a readout device world, when a thermodynamic cycle includes a quantum measurement, one needs to consider contributions to the measurement entropy from outside the measured system, and that there are sensible models in which these contributions compensate for any loss of system measurement entropy. Moreover, as we will see in the case of von Neumann's engine, although steps of a thermodynamical cycle involving quantum measurements may reduce the measurement entropy (if one does not take into consideration the entropy flow to the quantum measurement device), to close the cycle there must be a step involving a reading of a RD that increases the measurement entropy at least as much, and so there is no violation of the second law.

Finite-precision readout device models can be understood either by considering microsystems as having “smeared classicality” or by reducing the number of states in the state space, depending on the underlying model of finite precision. Again, the significance of information in such models can vary depending on the other available resources. For example, if one works in a model for which the accessible quantum states are qubits, with access to readout devices that indicate, through a single reading, that the quantum state of the system is contained in a specific region of the Bloch ball, then if one can apply arbitrary unitaries and can repeat the readings of the readout device, one may eventually gather enough information by rotating the Bloch ball to determine the quantum state to arbitrary precision.

Both infinite and finite precision readout device models are idealizations of the type of behavior that consistent nonlinear versions of quantum theory would imply. In an actual nonlinear theory that allows some form of effective readout device we generally expect other contributions to the total measurement entropy coming from sources external to the system analyzed. These depend on the details of the theory, but, in general, would include contributions arising from the internal structure and workings of the effective readout devices. For example, they should include the classical phase-space uncertainty of the gravitational field configurations in a semiclassical gravity theory. As noted above, a complete model would also need to describe the interplay between quantum operations and readout devices.

These issues preclude a general model-independent discussion of the thermodynamics of readout devices. However, we will not need to discuss all contributions to entropy flow in all possible models, since the redefinitions of entropy described above suffice to rebut the no-go arguments raised in the literature against nonlinear extensions to quantum theory, as we now explain.

III. PERES' ARGUMENT

As Peres notes [21], at first sight it seems that versions of quantum theory with a nonlinear modification of the Schrödinger equation inevitably violate the second law. To see this, consider two pure states $|\alpha(t)\rangle$ and $|\beta(t)\rangle$. Let

$$\rho(t) = \epsilon |\alpha(t)\rangle \langle \alpha(t)| + (1 - \epsilon) |\beta(t)\rangle \langle \beta(t)| \quad (23)$$

be a mixed state with $0 < \epsilon < 1$ and nonvanishing eigenvalues

$$\lambda_{1,2}(t) = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \epsilon(1 - \epsilon)[1 - x(t)]}, \quad (24)$$

where $x(t) := |\langle \alpha(t) | \beta(t) \rangle|^2$. The von Neumann entropy of this mixture $S_{\text{VN}}(t) = -\sum_i \lambda_i(t) \log[\lambda_i(t)]$ satisfies $\frac{dS_{\text{VN}}}{dx(t)} < 0$ for all ϵ and t . Thus,

$$\forall t, \frac{dS_{\text{VN}}}{dt}(t) = \frac{dS_{\text{VN}}}{dx(t)} \frac{dx(t)}{dt} \geq 0 \quad (25)$$

$$\Leftrightarrow \forall t, \frac{dx(t)}{dt} \leq 0 \quad (26)$$

$$\Rightarrow \forall t \geq 0, |\langle \alpha(t) | \beta(t) \rangle|^2 \leq |\langle \alpha(0) | \beta(0) \rangle|^2. \quad (27)$$

Hence if $\exists t \geq 0$ such that $|\langle \alpha(t) | \beta(t) \rangle|^2 > |\langle \alpha(0) | \beta(0) \rangle|^2$ then $\exists t \geq 0$ such that $\frac{dx(t)}{dt} > 0$ and $\frac{dS_{\text{VN}}}{dt}(t) < 0$. Consider a complete orthogonal set $\{|\alpha_k(t)\rangle\}$ where, for every $|\beta(t)\rangle$,

$$\sum_k |\langle \alpha_k(t) | \beta(t) \rangle|^2 = 1. \quad (28)$$

Thus, if there is some m for which $|\langle \alpha_m(t) | \beta(t) \rangle|^2 < |\langle \alpha_m(0) | \beta(0) \rangle|^2$ then there must also be some n for which $|\langle \alpha_n(t) | \beta(t) \rangle|^2 > |\langle \alpha_n(0) | \beta(0) \rangle|^2$, i.e., the entropy of $\rho_n(t) = \epsilon |\alpha_n(t)\rangle \langle \alpha_n(t)| + (1 - \epsilon) |\beta(t)\rangle \langle \beta(t)|$ would spontaneously decrease and thus the second law is violated unless $|\langle \alpha_k(t) | \beta(t) \rangle|^2 = |\langle \alpha_k(0) | \beta(0) \rangle|^2$ for every $|\alpha_k(t)\rangle$ and $|\beta(t)\rangle$ and all t , which, from Wigner's theorem, implies that time evolution is unitary or antiunitary, with the latter excluded by continuity. Thus, unless the evolution equation is linear, the second law is violated. Of course, in quantum theory, the Schrödinger evolution is linear, and orthogonal states remain orthogonal: i.e., $\langle \alpha(0) | \beta(0) \rangle = 0 \Rightarrow \langle \alpha(t) | \beta(t) \rangle = 0$ for all t . Hence this issue does not arise.

However, this argument for second-law violation in nonlinear theories assumes that other postulates that apply to quantum theory remain unchanged in such theories. In particular, it assumes that von Neumann entropy represents thermodynamical entropy. This need not necessarily be the case in nonlinear modifications of quantum theory. It also assumes that the time evolution for density operators follows Eq. (23). This need not be the case either for improper mixtures, as we now review.

A. Nonlinear time evolution in readout device world and the mixture equivalence principle

There are two kinds of state mixing in quantum theory: improper mixing, which arises as a description of a subsystem obtaining from tracing out a degree of freedom of an entangled multipartite system, and proper mixing, which arises from (classical) statistical considerations of an ensemble of states. It is important to highlight that, in nonlinear extensions of quantum theory, proper and improper mixtures are generally distinguishable: the so-called mixture equivalence principle [26] fails to hold.

In the case of proper mixtures, as was argued in [26], a proper mixture $\rho(0)$ based on an ensemble $\{p_i, |\psi_i(0)\rangle\}$ indeed evolves to a proper mixture at time t to $\rho(t)$ based on an ensemble $\{p_i, |\psi_i(t)\rangle\}$, even if the time evolution is nonlinear. That is, Eq. (23) does hold if $\rho(0)$ is a proper mixture of $|\alpha(0)\rangle$ and $|\beta(0)\rangle$.

However, an initial improper mixture whose density matrix $\rho(0)$ can be written in the form $\rho(0) = \sum_i p_i |\psi_i(0)\rangle\langle\psi_i(0)|$ will *not* generally evolve to $\sum_i p_i |\psi_i(t)\rangle\langle\psi_i(t)|$. The time evolution of an improper mixture ρ depends on that of the pure state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ of which it is the reduced density operator, with $\rho(t) = \text{Tr}_{\mathcal{H}_2}(|\psi(t)\rangle\langle\psi(t)|)$. Theories with readout devices trivially allow nonlinearity [8], meaning $\rho(t)$ will generally not be of the form of Eq. (23) for $t > 0$ when ρ is improper.

B. Peres' argument, revisited

Peres' argument relies on subtle assumptions that do not hold in nonlinear extensions of quantum theory. Foremost, we must now use measurement entropy instead of von Neumann entropy as the relevant entropic quantifier for a second law. In the case of proper mixtures, we can use the fact that measurement entropy in a readout device world is always

$$S_{\text{RD}}(\{p_i, |\psi_i(t)\rangle\}) = - \sum_i p_i \log(p_i) \quad (29)$$

and deduce that, in Peres' thought experiment, S_{RD} is constant and equal to $-\epsilon \log(\epsilon) + (1 - \epsilon) \log(1 - \epsilon)$ for all $t \geq 0$. Importantly, note that, in this case, the measurement and decomposition entropy of ρ are not equal, the latter being irrelevant for the discussion relevant to operationality and the second law. There is thus no violation of a second law of thermodynamics in the proper mixed state case.

In the case of improper mixtures, the analysis is different. As argued above, an improper mixture $\rho(0)$ does not generally evolve to the $\rho(t)$ defined by Eq. (23) at later times. Nevertheless, regardless of whether one has access to an infinite-precision or a finite-precision RD, $S_{(\text{FP})\text{RD}}[\rho_{\text{improper}}(t)] = 0$ for all times. There is thus no violation of a generalized second law in this case either.

IV. VON NEUMANN'S ARGUMENT

We now review the thought experiment originally discussed by von Neumann [20,21]. It takes the form of a thermodynamic cycle, shown in Fig. 1. It relies on the existence of semipermeable walls: boundaries between the two regions that allow some states to go through but not

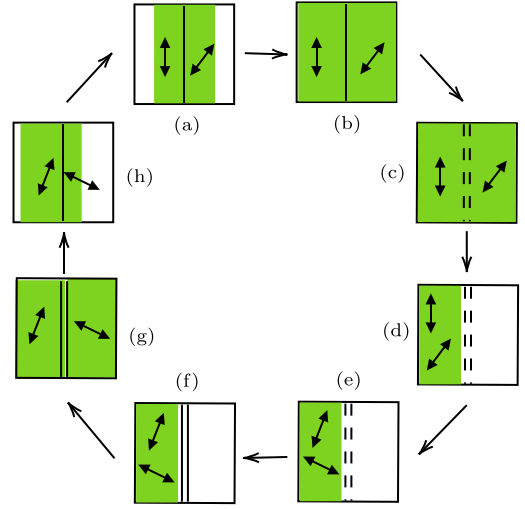


FIG. 1. Cycle extracting heat from an isothermal reservoir and converting it into work using a semi-permeable nonquantum membrane that separates a specific pair of nonorthogonal states as well as a semi-permeable quantum membrane that separates a specific pair of orthogonal states. The quantum membranes, which let through one of two orthogonal states, are depicted as solid double lines, while the semi-permeable membranes, which can be implemented by readout devices and let through one of two nonorthogonal states, are shown in dashed double lines. Walls, which do not let any states through, are depicted with single lines.

others. In von Neumann's words [20], for two quantum states $|\phi\rangle, |\psi\rangle \in \mathcal{H}$, "if $|\phi\rangle, |\psi\rangle$ are not orthogonal then the assumed existence of such a wall would contradict the second law of thermodynamics." To quote Peres [21], we have become "wily inventor[s] [who] claim having produced semipermeable partitions which unambiguously distinguish nonorthogonal states. [We] can thereby convert into work an unlimited amount of heat extracted from an isothermal reservoir [...] Will you invest your money in this new technology?"

Suppose that there are N photons in a box that has two chambers of equal volume. Half of the photons are prepared with vertical linear polarization and occupy one chamber, the other half with linear polarization at 45° from the vertical and occupy the other chamber, as seen in Fig. 1(a). These chambers undergo an isothermal expansion which doubles their volume Fig. 1(b), supplying a work $W^+ = Nk_B T \log(2)$, where T is the temperature of the reservoir. The walls separating the photon gases are then replaced by semipermeable membranes in Fig. 1(c) which can select nonorthogonal states (*).

One of the membranes is transparent to vertically polarized photons and reflects those polarized at 45° ; the other membrane does the opposite. We use a double frictionless piston as in Fig. 1(d) to obtain a mixture

$$\rho = \frac{1}{2} \left(|\uparrow\rangle\langle\uparrow| + \left| \frac{\pi}{4} \right\rangle\left\langle \frac{\pi}{4} \right| \right) = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}. \quad (30)$$

The eigenvalues of ρ correspond to photons polarized at $\frac{\pi}{8}$ radians from the vertical and those orthogonal [Fig. 1(e)]. We then replace the semi-permeable membranes with quantum

mechanical ones [Fig. 1(f)], which separate the two orthogonal polarization states [Fig. 1(g)]. We replace the quantum mechanical membranes by a wall and isothermally compress the gas [Fig. 1(h)] where both chambers have the same pressure and the same total volume as those in Fig. 1(a). The isothermal compression requires an expenditure of work of $W^- = Nk_B T S(\rho) < W^+$ (**). We finish the cycle by unitarily rotating the polarization of the photons in each chamber so as to complete the cycle [Fig. 1(a)].² Thus, we have a net decrease of the von Neumann entropy over a full cycle.

However, this argument for second-law violation again assumes that other postulates of quantum theory remain unchanged and that the information-theoretic properties of the theory follow that of quantum theory. We now show that, although readout device models indeed allow perfect discrimination of nonorthogonal states, the von Neumann cycle does not demonstrate problematic thermodynamical behavior.

A. Selection of nonorthogonal states in readout device world

Readout devices allow for what von Neumann argues is the key impossible step in his engine that stops quantum theory from violating a second law, that is, the selection of nonorthogonal states. This is easy to see with, e.g., an (FP)PRD: given a basis and two nonorthogonal states $|\psi\rangle$ and $|\phi\rangle$, one retrieves the classical description of both.

Provided the precision l is sufficient to distinguish both, an agent (a Maxwell demon) can then determine which of the vertical linear polarization or those at 45° from the vertical are incoming towards a membrane, and let them through or not at will.

For example, in semi-classical gravity, step (*) can be implemented as follows.

(1) Convert the polarization degrees of freedom of the photons into position degrees of freedom (say, with a beam splitter), this is (up to reversible nonlinear corrections) a unitary process which does not generate any entropy or work.

(2) Read out the position states using readout devices. This produces classical records of the states in the device outputs, with each readout corresponding to one of two distinct classical strings representing the two original polarization states. The proportions of these strings in the list of readouts are the same as those of the polarization states, i.e., equal. In the case where one must distinguish between two polarization states, this generates one bit of information, the corresponding work is $W^{\text{RD}} = Nk_B T \log(2)$ by Landauer's principle.

(3) Reconvert the position degrees of freedom into polarization degrees of freedom.

This amounts to performing a Cavendish experiment, writing down the value of the gravitational field sourced by the superposition of position eigenstates, and deducing the resulting expression of the state.

²If $|\psi\rangle$ and $|\phi\rangle$ are two orthogonal states of a quantum system, then $U = \exp[i\lambda(|\psi\rangle\langle\phi| - |\phi\rangle\langle\psi|)]$, with $\lambda \in \mathbb{R}$, is a unitary which rotates the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$: $U|\psi\rangle = \cos(\lambda)|\psi\rangle - \sin(\lambda)|\phi\rangle$ and $U|\phi\rangle = \sin(\lambda)|\psi\rangle + \cos(\lambda)|\phi\rangle$ [21].

B. von Neumann's argument, revisited

From these readings of an RD, we see already that

$$W^+ - (W^- + W^{\text{RD}}) \leq 0, \quad (31)$$

i.e., the second law of thermodynamics is not violated should the rest of the analysis cycle be maintained. Thus, one could argue that, through a “realistic” implementation of a Maxwell demon, one already evades this violation of the second law in a similar fashion to how one evades Maxwell's demon in classical thermodynamics and that in quantum thermodynamics in Szilard's engine [34].

However, for the sake of argument, we can check whether the rest of von Neumann's analysis still holds in readout device world. In step (**), specifically in the assumption that Fig. 1(d) is physically equivalent to Fig. 1(e), it is also assumed that the entropy of ρ is the Shannon entropy given by its eigendecomposition. Neither assumption holds in a readout device world, where distinct ensembles that result in the same proper mixture are distinguishable, so that Figs. 1(d) and 1(e) are not physically equivalent. The entropy of the ensemble $\{p_i, |\psi_i\rangle\}$ is the Shannon entropy of that ensemble, *even if it is not the eigendecomposition of the associated mixture*. We thus have that

$$S_{\text{RD}}(\rho) = -\frac{1}{2} \log\left(\frac{1}{2}\right) - \frac{1}{2} \log\left(\frac{1}{2}\right) = \log(2). \quad (32)$$

That is, $W^- = Nk_B T \log(2) = W^+$, i.e., even without taking into account the entropy production of the readings (*), we would have that $W^+ - W^- = 0$, i.e., no second law of thermodynamics is broken. This is true even if ρ is an improper mixture: given a state readout device (whether finite precision or not), though the measurement entropy associated to the mixture would be 0, we would also have that $W^+ = W^- = 0$ and thus we still have from step (*) that $W^+ - (W^- + W^{\text{RD}}) \leq 0$, so operationally no work is generated after one cycle.

C. Entropy reductions and the second law

Because Figs. 1(d) and 1(e) are not equivalent in a readout device world, the cycle should properly be understood as involving a step from Figs. 1(d) to 1(g). To implement this requires quantum measurements to implement the required semi-permeable quantum membrane. However, these measurements decrease the entropy of the system in a readout device world. Indeed, if we start in Fig. 1(a) with a proper mixture, it follows that going from Figs. 1(d) to 1(g), i.e., from a state which is “more mixed” to one which is “less mixed,” reduces the measurement entropy.

As discussed above, a complete model needs to consider the (readout device appropriate, i.e., measurement) entropy that is transferred to the quantum measurement device by the measurements. Without a complete description of the RD and quantum semi-permeable membranes and the entropy they transfer to the associated devices one cannot correctly calculate the entropy change over the cycle.

However, even without a complete model of quantum measurements in RD world, we see that step (*) would compensate any decrease of the thermodynamic entropy that could arise from Figs. 1(d) to 1(g) to complete the cycle, i.e., there is no violation of the second law over a cycle. We expect that,

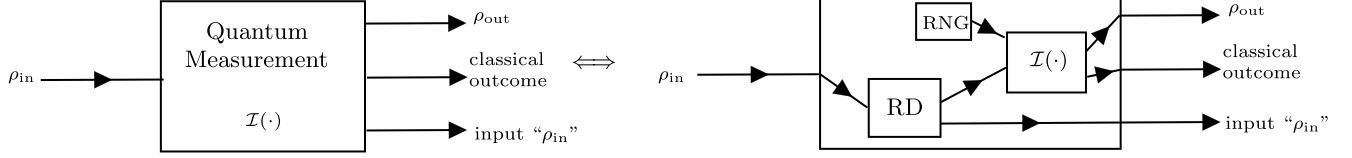


FIG. 2. Quantum measurements as a black box in RD world. A possible model of quantum measurements in this postquantum setting is as a composite process of (i) an RD capturing the information of the input state and placing it in some register (perhaps within the measurement apparatus), (ii) the quantum instrument $\mathcal{I}(\cdot)$ associated to the quantum measurement, applying the nonselective quantum operation, and (iii) a random number generator (RNG), selecting one specific outcome at random.

in more general cycles, steps with quantum measurements can generate work, but completing any cycle requires some use of a RD that generates information about the quantum state and thus requires at least as much work.

One may be uneasy at the idea that decreases of the measurement entropy are possible in a fundamental theory (even if, over a cycle, there is no decrease). One view is that this is just a feature of RD theories, and that there is no possibility of overall work generation, so the second law is not violated.

A second view is that temporary information (understood in the context of measurement entropy) nonconservation is unaesthetic and unsatisfactory, albeit not immediately contradictory. On this view, one should give a precise (RD world model-dependent) model of quantum measurements that shows such decreases never happen when one considers the full setup, because a transfer of information from the system to the measurement device compensates for any apparent decrease.

For example, one could model quantum measurements as a composite black-box process which includes an RD measurement device, for which there is always a register of the initial quantum state after the measurement that is inaccessible if one only has access to standard quantum resources. Such a setup is depicted in Fig. 2. The information about the initial quantum state is then always conserved, and so there is never any measurement entropy decrease upon quantum measurements.

Note that in this model information is still generated in (almost) trivial quantum measurements, for which the quantum instrument acts (approximately) as the identity on the quantum state. This implies a continuous production of information about the quantum state during continuous measurements, and indeed during free evolution if this is modelled as a continuous measurement process. This is arguably not completely unphysical, however. For example, in semi-classical gravity, during free evolution, the classical gravitational field still continuously reads the local quantum state, acting as a continuous readout device.

Again, this discussion depends on the model and on the accessible physical resources. Also, finite precision models for both quantum measurements and readout devices may be more appropriate in some contexts. Our point is that there are sensible models whereby even entropy reductions under quantum measurements can be averted.

In summary, on either view, von Neumann's engine is implementable in a readout device world but does not demonstrate the violation of any thermodynamical law there.

V. HÄNGGI AND WEHNER'S ARGUMENT

We now review an argument of Hänggi and Wehner which shows that, under several assumptions reviewed below, a violation of the spin uncertainty principle implies a violation of the second law of thermodynamics [22]. The “impossible cycle” relies on the notion of *entropic* or *fine-grained* uncertainty relations [35–37] which reformulate the standard Heisenberg uncertainty relations in a systematic fashion such that the bound does not depend on the state used for the computation of the expectation values of the standard uncertainty relations

$$\Delta_\psi \hat{A} \Delta_\psi \hat{B} \geq \frac{1}{2} |\langle [\hat{A}\hat{B}] \rangle_\psi|, \quad (33)$$

where

$$\Delta_\psi \hat{O} = \sqrt{\langle \hat{O}^2 \rangle_\psi - \langle \hat{O} \rangle_\psi^2}. \quad (34)$$

Let $\mathbf{f}$ and $\mathbf{g}$ be two eigenbases for projective measurements with pure effects $\{f_0, f_1\}$ and $\{g_0, g_1\}$, respectively. For example, in the X and Z bases these are $\{|+\rangle\langle+|, |-\rangle\langle-|\}$ and $\{|0\rangle\langle0|, |1\rangle\langle1|\}$, respectively. Fine-grained uncertainty relations for such measurements are then defined to be the following four inequalities:

$$\frac{1}{2}(p(f_0|\psi)) + p(g_0|\psi)) \leq \zeta_{f_0, g_0}, \quad (35)$$

$$\frac{1}{2}(p(f_0|\psi)) + p(g_1|\psi)) \leq \zeta_{f_0, g_1}, \quad (36)$$

$$\frac{1}{2}(p(f_1|\psi)) + p(g_0|\psi)) \leq \zeta_{f_1, g_0}, \quad (37)$$

$$\frac{1}{2}(p(f_1|\psi)) + p(g_1|\psi)) \leq \zeta_{f_1, g_1}, \quad (38)$$

for any pure quantum state $|\psi\rangle$, where for a projective measurement in the Z basis with outcome $|0\rangle$, $p(|0\rangle|\psi) = |\langle 0|\psi\rangle|^2$ by the Born rule, and $\zeta_{f_i, g_j} > 0$ are suprema over states for the measurement outcomes $\{f_i, g_j\}$. For $\zeta < 1$ these inequalities quantify uncertainty since if the outcome is certain for one outcome, then it cannot be for the other. For quantum theory which respects the Heisenberg uncertainty principle, the constraint on ζ is stronger, with $\zeta \leq \frac{1}{2} + \frac{1}{2\sqrt{2}}$.

A. Readout devices and the uncertainty principle

First, we highlight that readout devices indeed violate the uncertainty principle. It is easy to see that, given any two states, the associated fine-grained uncertainty relations are upper bounded by 1. For example, given $\rho_{0_X} = |+\rangle\langle+|$ and $\rho_{1_Z} = |1\rangle\langle1|$, there exists an effect e such that

$$\frac{1}{2} + \frac{1}{2\sqrt{2}} \leq \frac{1}{2}[p(e|\rho_{0_X}) + p(e|\rho_{1_Z})] = 1, \quad (39)$$

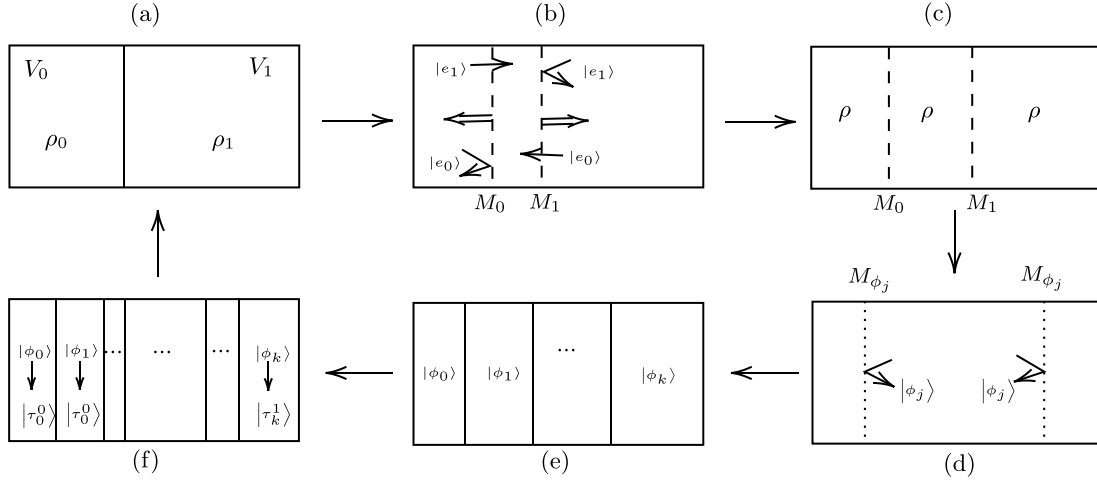


FIG. 3. The Hänggi-Wehner cycle which can extract net work if the uncertainty principle is violated and other principles are left unchanged. As discussed in the text, this cycle fails to extract net work for the nonlinear models we consider. First (a) the system is prepared in two states ρ_1 and ρ_2 in separated volumes, each with $N/2$ particles. (b) We then replace the wall by semi-transparent membranes that let through one state but block the other. (c) These membranes move apart until equilibrium ρ is reached. (d) Insert new membranes to separate the pure components of ρ . (e) Subdivide these into smaller regions such that the resulting pure states are building blocks of ρ_1 and ρ_2 . (f) Unitarily transform these states into the pure state decomposition of ρ_1 and ρ_2 .

which can be achieved, e.g., by an RD, or by letting the states evolve nonlinearly to orthogonal states and measuring quantum mechanically in an appropriate basis, which allows one to get arbitrarily close to the upper bound of 1. Note that in readout device models, unlike the kind of theory considered by Hänggi and Wehner, the converse uncertainty relations defined by fixing a state and considering “complementary” measurements (e.g., through projectors $|+\rangle\langle+|$ and $|0\rangle\langle 0|$) cannot generally be formulated. Indeed, in a readout device world, pure effects are not generally dual to pure states.

B. Hänggi-Wehner cycle

Let us outline the steps of this engine using quantum notation. The cycle is presented in Fig. 3.

We start the engine with a box which contains two types of particles: ρ_0 and ρ_1 , given by

$$\rho_0 = \frac{1}{2}(|+\rangle\langle+| + |0\rangle\langle 0|) = \frac{\mathbb{1} + \frac{X+Z}{2}}{2}, \quad (40)$$

$$\rho_1 = \frac{1}{2}(|-\rangle\langle-| + |1\rangle\langle 1|) = \frac{\mathbb{1} - \frac{X+Z}{2}}{2}, \quad (41)$$

which are in two separated volumes.

Figure 3(a): There are $N/2$ particles in the ρ_0 state and $N/2$ particles in the ρ_1 state.

Figure 3(b): We then make a projective measurement in the $\{|e_0\rangle, |e_1\rangle\}$ basis where $|e_0\rangle$ and $|e_1\rangle$ are the eigenstates of $\frac{X+Z}{\sqrt{2}}$ (the measurement thus only has two outcomes), and replace the wall separating these two volumes by semi-transparent membranes M_0 and M_1 which let through $|e_0\rangle$ and $|e_1\rangle$, respectively, but block the other state.

Figure 3(c): These membranes then move apart until they reach equilibrium: we have the maximally mixed state

$$\rho = \frac{1}{2}(\rho_0 + \rho_1) = \frac{\mathbb{1}}{2} \quad (42)$$

in the three regions. The work that can be extracted from this process is

$$W_1^+ = Nk_B T \log(2) \left[1 - \frac{1}{2}H(\zeta_{|+ \rangle, |0 \rangle}) - \frac{1}{2}H(\zeta_{| - \rangle, |1 \rangle}) \right]. \quad (43)$$

Figure 3(d): We then insert new membranes on the side so as to separate the pure components of ρ into regions which are only populated by pure states $|\phi_j\rangle$. This process requires a work

$$W^- = Nk_B T \log(2) S(\rho) = Nk_B T \log(2) \quad (44)$$

as $S(\rho) = 1$.

Figure 3(e): We then subdivide the volumes containing the pure states $|\phi_j\rangle$ into smaller volumes such that the number of particles contained in these smaller volumes are proportional to $\frac{r_j^0}{2}$ and $\frac{r_j^1}{2}$, where

$$\rho_0 = \sum_j r_j^0 |\tau_j^0\rangle\langle\tau_j^0|, \quad (45)$$

$$\rho_1 = \sum_j r_j^1 |\tau_j^1\rangle\langle\tau_j^1|. \quad (46)$$

Figure 3(f): The pure state $|\phi_j\rangle$ contained in each small volume is then unitarily transformed into the pure state $|\tau_j^0\rangle$ or $|\tau_j^1\rangle$, which are also pure and thus do not require work. We finally mix the different components of ρ_0 together (and likewise for ρ_1) so as to complete the cycle [Fig. 3(a)], which extracts work

$$W_2^+ = Nk_B T \log(2) \sum_{i=0}^1 \frac{1}{2} S(\rho_i). \quad (47)$$

Since ρ_0 and ρ_1 both have one positive eigenvalue $\frac{1}{2} + \frac{1}{2\sqrt{2}}$, we obtain

$$S(\rho_{0,1}) = H\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) \quad (48)$$

$$\equiv -\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) \log\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right). \quad (49)$$

Thus, the total work which can be extracted is

$$\Delta W = W_1^+ + W_2^+ - W^- \quad (50)$$

$$= Nk_B T \log(2) \left[H\left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) - \frac{1}{2}H(\zeta_{|+,|0}) - \frac{1}{2}H(\zeta_{|- ,|1}) \right]. \quad (51)$$

It is here evident that quantum theory does not violate the second law for such a cycle, but that any theory with quantum-like features that satisfies certain assumptions, for which $\zeta > \frac{1}{2} + \frac{1}{2\sqrt{2}}$, does violate the second law. However, again, we shall show that these assumptions are violated in the extensions of quantum theory that we consider here.

C. Readout devices and Hänggi and Wehner's argument

As was highlighted by Hänggi and Wehner in their original formulation of the engine [22], this cycle is based on several assumptions that the theory should respect the following.

- (A1) The state space is convex.
- (A2) Effects are linear functionals.
- (A3) Pure states are dual to pure effects. Uncertainty relations can thus be stated equivalently in terms of states or measurements.³
- (A4) Pure effects are projective.
- (A5) If $f_0 + f_1 = u$ for two effects f_0 and f_1 , then the dual states $\rho_{f_0} + \rho_{f_1} = \rho_u$ and $e(\rho_u/2) = \frac{1}{2}$ for all pure effects e .
- (A6) Let $\rho = \sum_{j=1}^d q_j \sigma_j$ be a decomposition of ρ into perfectly distinguishable pure states σ_j . Let h_{σ_j} denote the pure effect dual to σ_j . Then $\sum_{j=1}^d h_{\sigma_j} = u$ and $h_k(\sigma_j) = \delta_{kj}$ for all j and k .
- (A7) Let ρ and σ be pure states. Then the transformation $\rho \rightarrow \sigma$ is reversible (and thus does not require any work, neither can any work be gained from it).

In a readout device world, most of these assumptions, which are necessary for the derivation of the engine, are violated.

(1) Assumption (A2) is violated: readout devices output a classical description of density operators with probability 1, i.e., they do not respect the convex structure of the state (while the Born rule does).

(2) Assumptions (A3), (A5), and (A6) are violated: pure states are not necessarily dual to pure effects because effects from readout devices are all pure. Consequently, (A5) and (A6) also fail to hold.

(3) Assumption (A7) is violated: since we work with nonlinear quantum mechanics, it need not be the case that if ρ and σ are any pure states, the transformation $\rho \rightarrow \sigma$ is reversible. Hence, it may be that the transformation between two pure states requires work, or that work can be extracted from it.

Further note that the failure of assumption (A6) to hold also highlights the failure of the equivalence between decomposition entropy and measurement entropy, which is key in

this derivation. Again, if one starts with a proper mixture given by an ensemble $\{1/2, \rho_0; 1/2, \rho_1\}$, the entropy will be constantly that of the Shannon entropy of that specific ensemble. The subdivision of ρ into different pure states does not induce a new value for the entropy because decomposition entropy is not the relevant quantity here. Likewise, if ρ is an improper mixture, then the measurement entropy ought to be 0 regardless of whether one works with infinite-precision or finite-precision RDs across the whole process unless one injects work into it. Thus, there is no violation of a generalized second law in this context.

VI. CONCLUSION

We have seen that claims that nonlinear extensions of quantum theory necessarily violate the second law of thermodynamics need not hold once a generalization to the measurement entropy, suitable for the theory, is considered. Focusing on a class of radical modifications, namely, state readout devices, which can be realized, for example, by Møller-Rosenfeld semi-classical gravity, we demonstrated that several no-go arguments rest on implicit assumptions tied to standard quantum information theory. These include the identification of thermodynamic entropy with von Neumann entropy, the equivalence between proper and improper mixtures, and some form of duality between states and effects.

We reevaluated classic arguments by Peres and von Neumann, showing that their conclusions do not hold in the context of readout device theories. In particular, von Neumann's thought experiment, which purportedly extracted net work by separating nonorthogonal states, fails not only once one accounts for the work cost of information erasure via readout devices as required by Landauer's principle (which is a similar resolution to "standard" Maxwellian and Szilardian demons), but also does not account for a more operational notion of entropy relevant to the nonlinear theory. Similarly, we addressed the Hänggi-Wehner engine, which links violations of the uncertainty principle to thermodynamic violations. We showed that the derivation of this result depends on several assumptions, such as the linearity of effects and the equivalence between pure states and effects, which are violated in readout device models. As such, violations of entropic uncertainty relations in these models need not lead to thermodynamic paradoxes.

These results suggest that nonlinear extensions of quantum theory, when carefully constructed and interpreted, can remain thermodynamically consistent, though we do not provide a model-independent proof. In particular, semi-classical gravity, which admits nonstandard readout mechanisms without necessarily violating relativistic causality, provides a framework in which nonstandard physical phenomena can arise without (at least immediately) contradicting thermodynamic principles. Thermodynamic arguments to date thus give no reason to reject such models.

ACKNOWLEDGMENTS

We acknowledge financial support from the UK Quantum Communications Hub Grant No. EP/T001011/1 and Project No. OPP640, funded by the Science and Technology

³It was noted in [22] that this assumption is satisfied by any theory which satisfies so-called bit symmetry [38]: every logical bit can be mapped to any other logical bit by a reversible transformation.

Facilities Council's International Science Partnerships Fund. S.F. is funded by a studentship from the Engineering and Physical Sciences Research Council, Grant No. 2882481. A.K. was supported, in part, by the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science, and Economic Development and by

the Province of Ontario through the Ministry of Research, Innovation, and Science. A.K. thanks Gerard Milburn and Sally Shrapnel for helpful discussions.

DATA AVAILABILITY

No data were created or analyzed in this study.

-
- [1] A. Kent, Quantum state readout, collapses, probes, and signals, *Phys. Rev. D* **103**, 064061 (2021).
 - [2] D. N. Page and C. D. Geilker, Indirect evidence for quantum gravity, *Phys. Rev. Lett.* **47**, 979 (1981).
 - [3] J. Wilson-Gerow and P. C. E. Stamp, Propagators in the correlated worldline theory of quantum gravity, *Phys. Rev. D* **105**, 084015 (2022).
 - [4] C. I. Kuo and L. H. Ford, Semiclassical gravity theory and quantum fluctuations, *Phys. Rev. D* **47**, 4510 (1993).
 - [5] A. Tilloy and L. Diósi, Sourcing semiclassical gravity from spontaneously localized quantum matter, *Phys. Rev. D* **93**, 024026 (2016).
 - [6] A. Tilloy, Binding quantum matter and space-time, without romanticism, *Found Phys* **48**, 1753 (2018).
 - [7] N. Gisin, Weinberg's non-linear quantum mechanics and superluminal communications, *Phys. Lett. A* **143**, 1 (1990).
 - [8] A. Kent, Nonlinearity without superluminality, *Phys. Rev. A* **72**, 012108 (2005).
 - [9] L. Masanes, T. D. Galley, and M. P. Müller, Response to the measurement postulates of quantum mechanics are not redundant, *Quantum* **9**, 1592 (2025).
 - [10] A. Kent, The measurement postulates of quantum mechanics are not redundant, *Quantum* **9**, 1749 (2025).
 - [11] B. C. Stacey, Contradictions or curiosities? On Kent's critique of the Masanes-Galley-Müller derivation of the quantum measurement postulates, [arXiv:2405.17733](https://arxiv.org/abs/2405.17733).
 - [12] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777 (1935).
 - [13] A. S. Wightman, La théorie quantique locale et la théorie quantique des champs, *Annales de l'institut Henri Poincaré. Section A., Physique Théorique* **1**, 403 (1964).
 - [14] H. Halvorson and M. Mueger, Algebraic quantum field theory, in *Handbook of the Philosophy of Physics*, edited by J. Butterfield and J. Earman (Kluwer, Amsterdam, 2006).
 - [15] S. Bose, A. Mazumdar, G. W. Morley, H. Ulbricht, M. Toroš, M. Paternostro, A. A. Geraci, P. F. Barker, M. S. Kim, and G. Milburn, Spin entanglement witness for quantum gravity, *Phys. Rev. Lett.* **119**, 240401 (2017).
 - [16] C. Marletto and V. Vedral, Gravitationally induced entanglement between two massive particles is sufficient evidence of quantum effects in gravity, *Phys. Rev. Lett.* **119**, 240402 (2017).
 - [17] R. Howl, V. Vedral, D. Naik, M. Christodoulou, C. Rovelli, and A. Iyer, Non-Gaussianity as a signature of a quantum theory of gravity, *PRX Quantum* **2**, 010325 (2021).
 - [18] A. Kent and D. Pitalúa-García, Testing the nonclassicality of spacetime: What can we learn from Bell-Bose *et al.* -Marletto-Vedral experiments? *Phys. Rev. D* **104**, 126030 (2021).
 - [19] A. Kent, Testing quantum gravity near measurement events, *Phys. Rev. D* **103**, 064038 (2021).
 - [20] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, NJ, 1955).
 - [21] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Amsterdam, 1993).
 - [22] E. Hänggi and S. Wehner, A violation of the uncertainty principle implies a violation of the second law of thermodynamics, *Nat. Commun.* **4**, 1670 (2013).
 - [23] B. Mielnik, Mobility of nonlinear systems, *J. Math. Phys.* **21**, 44 (1980).
 - [24] T. D. Galley, F. Giacomini, and J. H. Selby, Any consistent coupling between classical gravity and quantum matter is fundamentally irreversible, *Quantum* **7**, 1142 (2023).
 - [25] D. R. Terno, Classical-quantum hybrid models, [arXiv:2309.05014](https://arxiv.org/abs/2309.05014).
 - [26] S. Fedida and A. Kent, Mixture equivalence principles and postquantum theories of gravity, *Phys. Rev. D* **111**, 126016 (2025).
 - [27] L. Susskind, L. Thorlacius, and J. Uglum, The stretched horizon and black hole complementarity, *Phys. Rev. D* **48**, 3743 (1993).
 - [28] M. Krumm, H. Barnum, J. Barrett, and M. P. Müller, Thermodynamics and the structure of quantum theory, *New J. Phys.* **19**, 043025 (2017).
 - [29] I. Layton, and H. J. D. Miller, Restoring the second law to classical-quantum dynamics, [arXiv:2504.10587](https://arxiv.org/abs/2504.10587).
 - [30] K. R. W. Jones, Newtonian quantum gravity, *Aust. J. Phys.* **48**, 1055 (1995).
 - [31] R. Ruffini and S. Bonazzola, Systems of self-gravitating particles in general relativity and the concept of an equation of state, *Phys. Rev.* **187**, 1767 (1969).
 - [32] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Generalized no-broadcasting theorem, *Phys. Rev. Lett.* **99**, 240501 (2007).
 - [33] A. J. Short and S. Wehner, Entropy in general physical theories, *New J. Phys.* **12**, 033023 (2010).
 - [34] L. Szilard, Über die Entropieverminderung in einem thermodynamischen system bei Eingriffen intelligenter Wesen, *Z. Angew. Phys.* **53**, 840 (1929).
 - [35] D. Deutsch, Uncertainty in quantum measurements, *Phys. Rev. Lett.* **50**, 631 (1983).
 - [36] H. Maassen and J. B. M. Uffink, Generalized entropic uncertainty relations, *Phys. Rev. Lett.* **60**, 1103 (1988).
 - [37] J. Oppenheim and S. Wehner, The uncertainty principle determines the nonlocality of quantum mechanics, *Science* **330**, 1072 (2010).
 - [38] M. P. Müller, J. Oppenheim, and O. C. O. Dahlsten, The black hole information problem beyond quantum theory, *J. High Energy Phys.* **09** (2012) 116.