Everettian Formulation of the Second Law of Thermodynamics

Yu Feng*
Foundations of Quantum Mechanics Research Group
(Dated: September 1, 2021)

The second law of thermodynamics is traditionally interpreted as a coarse-grained result of classical mechanics. Recently its relation with quantum mechanical processes such as decoherence and measurement has been revealed in literature. In this paper we will formulate the second law and the associated time irreversibility following Everett's idea: systems entangled with an object getting to know the branch in which they live. Accounting for this self-locating knowledge, we get two forms of entropy: *objective entropy* measuring the uncertainty of the state of the object alone, and *subjective entropy* measuring the information carried by the self-locating knowledge. By showing that the summation of the two forms of entropy is a conserved and *perspective-free* quantity, we interpret the second law as a statement of irreversibility in knowledge acquisition. This essentially derives the thermodynamic arrow of time from the subjective arrow of time, and provides a unified explanation for varieties of the second law, as well as the past hypothesis.

I. INTRODUCTION

The second law of thermodynamics is traditionally derived from the classical phase-space trajectory described by the Liouville equation [1, 2]: the state of an isolated system tends to move towards coarse-grained macrostates with larger numbers of microscopic degrees of freedom [3]. This derivation of the second law, however, relies on the asymmetry of boundary conditions, for the underlying microscopic laws of physics are completely reversible. To explain the ubiquitous validity of the second law, one has to resort to the cosmological initial condition: a low-entropy universe at the beginning of time [4, 5]. This is, however, an *ad hoc* assumption (the past hypothesis) that is hard to further explain [6].

A different puzzle in modern physics is the quantum-mechanical measurement problem. When one measures a single superposed state, it randomly collapses to states with different outcomes. All we can predict are the probabilities of these outcomes. There is a long-lasting argument about the nature of this quantum state reduction (QSR). As Tegmark pointed out in [7], if QSR is a physical process, then it would unavoidably increase entropy over time. Given the current state of the universe, this suggests an initial state with even lower entropy. As a result, the situation of the past hypothesis can only get worse.

If instead we regard QSR as only apparent, and assume unitarity of the entire universe, then the fundamental laws of physics can be formulated with complete simplicity. This is the approach suggested by Everett decades ago [8]. He proposed that our experience is limited in the sense that we could only perceive a branch of the entire reality. When opening the box that contains Schrödinger's cat, the observer is entangled with the cat and its environment. This results in apparent branching such that the observer in different branches perceives different states of the cat.

* foundationsofQM@gmail.com

The essence of Everett's idea is threefold. First, the "world" we perceive is in fact only a "branch" of the entire reality. We will therefore define entire reality as "*global* reality", and reality in a specific branch as "*local* reality". Accordingly, we could also talk about *global* state and *local* state of an object. Note that in this paper the terms *global* and *local* are used specifically for quantum-mechanical branching, and have nothing to do with physical spacing. Secondly, there is nothing fundamentally different between the subject (the observer) and the object (the observed). This is in contrast to the von Neumann-Wigner interpretation, which claims that the consciousness of the observer is essential to QSR [9]. Thirdly, the *complete* information of an object is not only encoded in itself, but also in the systems with which it interacts. From the perspective of these systems, what really matters are the *local* states of the object in their own branches.

As an example, consider a simplified scenario where an observer and an electron are entangled in a *global* state:

$$(|\uparrow\rangle \otimes |\text{perceiving} \uparrow\rangle + |\downarrow\rangle \otimes |\text{perceiving} \downarrow\rangle)/\sqrt{2}\ . \quad (1)$$

Note that this *global* state is for the composite system consisting of both the object (i.e. the electron) and the observer. The *global* state of the object alone can only be described by a density matrix $(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2$. It represents a statistical ensemble, because some of the information is encoded in the entanglement between the object and the observer. Ignoring such information by excluding the observer makes it an *incomplete* description of the object. But even knowing the *global* state of the whole composite system, Eq. 1, is still *incomplete* from the observer's perspective.

The missing part is knowing in which branch the observer *actually* lives. This is called "self-locating knowledge". Without such knowledge, the observer cannot possibly know the *local* state from her own perspective. Suppose the observer is deprived of the ability of perception, and so might ask, "Between the two branches where I should perceive ↑ or ↓, in which one do I *actually* live?". This self-locating uncertainty is then cleared after perception returns. The observer forms the self-locating knowledge, "I am *actually* living in the branch in which I perceive $\uparrow$!". Now she is certain about the object state being a pure state $|\uparrow\rangle\langle\uparrow|$ rather than the mixed state $(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2$. The former state, namely the *local* state of the object, represents the *local* reality in her own branch, which is all that matters from the observer's perspective. Accordingly, we can translate concepts of the standard formulation of quantum mechanics to concepts related to the self location of the observer: "indeterminism" to "self-locating uncertainty", "probability" to "self-locating probability", and "collapse" to "acquisition of self-locating knowledge". This provides rooms for these concepts under Everettian quantum mechanics [10–14].

In the example above, the observer reduces the *global* state of the object, $(|\uparrow\rangle\langle\uparrow|+|\downarrow\rangle\langle\downarrow|)/2$, to the *local* state in her own branch, $|\uparrow\rangle\langle\uparrow|$, by making use of her self-locating knowledge. We will see later that this self-locating knowledge gives a second form of entropy, namely the *subjective entropy*. An entropic equality is then derived, which relates the second law to the acquisition of self-locating knowledge. This equality bridges between the *global* and *local* descriptions of an object. By accounting for the *subjective entropy*, we show that change in entropy is in fact a perspective change: either from the *global* description to the *local* description, or from one *local* description to another. Such perspective change is measured by flows between different forms of entropy. We will find that the formulation based on entropy transfer, instead of entropy change, can well explain both the second law and the past hypothesis.

The following sections are organized as follows. The second section explains why von Neumann entropy can be used when discussing the second law. It serves as a pre-requisition, as in the following sections we will use thermodynamic and von Neumann entropies interchangeably. The third and fourth sections examine a bipartite object-environment system, and explains the necessity of defining a *subjective entropy* with the help of the quantum Szilard engine. The fifth section introduces an external observer and examines the entropic equality under the tripartite object-environment-observer system. The sixth and seventh sections show the validity of the equality under arbitrary system partition and general positive-operator-valued measures (POVM). The eighth section discusses the physical meaning of the mathematical formulation and how it provides a unified explanation for the second law and the past hypothesis.

II. THERMODYNAMIC ENTROPY AND VON NEUMANN ENTROPY

There is a common confusion on the notion of von Neumann entropy and whether it is applicable to thermodynamics. Sometimes von Neumann entropy is referred as the "fine-grained" entropy, in contrast to the "coarse-grained" thermodynamic entropy [15]. This, however, regards von Neumann entropy as the entropy that measures the uncertainty of the *global* description, which is only one of its use cases. In fact, when discussing the fine-grained entropy, we regard ourselves as "ideal observers" that are not entangled with the object. Besides, the object does not interact with its environment, and is regarded as an isolated quantum system. The *global* state of this object then follows the Liouville-von Neumann equation which gives a constant von Neumann entropy [16]. In this case, no thermodynamic irreversibility arises.

In reality, the object interacts with its environment and the observer, and is no longer an isolated system. We may instead use von Neumann entropy to measure the uncertainty of its *local* description. This gives rise to thermodynamic irreversibility. In practice, we may reduce the *global* description to *local* descriptions with two operations: conditioning on the observation (which corresponds to quantum mechanical measurement) and partial-tracing over the environment (which corresponds to decoherence). We will see in the following section that the latter operation arises from the former by averaging over all branches. Both operations are quantum mechanical counterparts of the classical coarse-graining procedure.

Consider a simple example in which an isolated box contains two distinguishable particles labeled 1 and 2 (see Fig. 1). The box has two rooms labeled L (left) and R (right), separated by an opaque wall. The state of each particle lives in a two-dimensional Hilbert space spanned by orthonormal states $|L\rangle_i$ and $|R\rangle_i$ ($i = 1, 2$). We assume the box does not interact with the particles, and the wall denoted by W only interacts with the particle(s) on its left side. The state of W is therefore denoted by the particle(s) in the left room: $|n\rangle_W$ (none), $|1\rangle_W$, $|2\rangle_W$, $|b\rangle_W$ (both), or their superposition. An external observer with label O is watching the box from its left side. Due to the opaque wall, the observer can only see inside the left room. Her state is either $|e\rangle_O$ (seeing an empty room), $|f\rangle_O$ (seeing a non-empty room), or their superposition.

Now suppose the entire system is in a superposed state of $\big(a_1|L\rangle_1|L\rangle_2|b\rangle_W + a_2|L\rangle_1|R\rangle_2|1\rangle_W + a_3|R\rangle_1|L\rangle_2|2\rangle_W\big)|f\rangle_O + a_4|R\rangle_1|R\rangle_2|n\rangle_W|e\rangle_O$ where a_i ($i = 1, \cdots, 4$) are coefficients that satisfy the normalization condition $\sum_{i=1}^4 |a_i|^2 = 1$. If we would like to use the von Neumann entropy formula to calculate the thermodynamic entropy, we need to first condition on the observation so that the description is *local*. If the observer sees an empty room, she is in the branch in which the particles and the box are in the state $|R\rangle_1|R\rangle_2|n\rangle_W$. If instead she sees a non-empty room, then she is in the branch in which the particle-box subsystem is in a superposed state of (ignoring any normalization factor) $a_1|L\rangle_1|L\rangle_2|b\rangle_W + a_2|L\rangle_1|R\rangle_2|1\rangle_W + a_3|R\rangle_1|L\rangle_2|2\rangle_W$.

Following the conditioning, we need to partial-trace over the wall, for the entropy we consider measures the uncertainty of the particle states only. For example, in

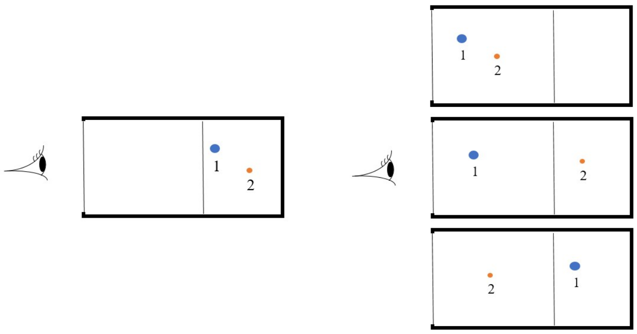

FIG. 1. A system consists of a box, an opaque wall, an observer, and two distinguishable particles labeled 1 and 2.

the branch where the observer sees a non-empty room, partial-tracing results in a mixed state of the particles, described by a density matrix $\rho = p_1|L\rangle_1\langle L|_1 \otimes |L\rangle_2\langle L|_2 + p_2|L\rangle_1\langle L|_1 \otimes |R\rangle_2\langle R|_2 + p_3|R\rangle_1\langle R|_1 \otimes |L\rangle_2\langle L|_2$ where the probabilities $p_i = |a_i|^2/\sum_{j=1}^3 |a_j|^2$.

The von Neumann entropy of the particles in this branch, after conditioning and partial-tracing, gives rise to the Gibbs entropy formula:

$$-\mathrm{tr}(\rho\log\rho) = -p_1\log p_1 - p_2\log p_2 - p_3\log p_3\ . \tag{2}$$

We may compare Eq. 2, which gives a positive entropy, with the other branch where the observer is in $|e\rangle_O$, in which the particles have zero entropy. According to the second law, there is a macroscopic tendency for the observer state to switch from $|e\rangle_O$ to $|f\rangle_O$, but not the other way around.

To see that Eq. 2 indeed gives the thermodynamic entropy, we let the particles interact with the box, assuming it is a thermal reservoir at a constant inverse temperature β. For simplicity, we also assume each particle has only two energy states, either 0 or $\varepsilon > 0$, and the wall acts as a selectively permeable membrane such that particle(s) are on its left side if and only if they have zero energy. Due to the symmetry between the two particles, the probabilities in the density matrix follow $p_2 = p_3$. We define $p := p_2 = p_3$ and $q := p_1 = 1 - 2p$. Assuming the whole system is in thermal equilibrium, we have $q = p\exp(\beta\varepsilon)$ according to the Boltzmann distribution. We may then derive the fundamental thermodynamic relation from the internal energy formula $U = 2p\varepsilon$ and the entropy formula (Eq. 2) $S = -2p\log p - q\log q$:

$$\begin{aligned} dS &= -(\log q + 1)dq - 2(\log p + 1)dp \\ &= \log(p/q)d(-2p) = 2\beta\varepsilon dp = \beta dU\ . \end{aligned} \tag{3}$$

This shows that S as given by Eq. 2 is indeed the thermodynamic entropy.

In general, consider an arbitrary system that is decomposed into an object, an external observer, and the rest as the environment. Suppose the whole tripartite system is in a pure state $|\Phi\rangle \in H \otimes H_e \otimes H_s$ where H, H_e, and H_s denote the respective Hilbert spaces of the object (e.g. particles), the environment (e.g. wall/box), and the observer. These Hilbert spaces (assumed being finite-dimensional) have dimensions of n, n_e, and n_s, respectively. With the help of Schmidt decomposition [17], $|\Phi\rangle$ can be expressed by (assuming $nn_e \geq n_s$ as the environment usually has a much larger number of degrees of freedom compared to the object and the observer):

$$|\Phi\rangle = \sum_{i=1}^{n_s} a_i|\phi_i\rangle \otimes |s_i\rangle\ , \tag{4}$$

where $|s_i\rangle$ $(i = 1, 2, \cdots, n_s)$ form an orthonormal basis of H_s, and the corresponding $|\phi_i\rangle$ are orthonormal states in $H \otimes H_e$. We may regard them as states that the observer can perceive. This is consistent with quantum Darwinism, which claims that pointer states correspond to those that can leave numerous copies in the environment and are therefore determined by the unique Schmidt decomposition [18].

In each of the n_s branches, the observer perceives observation outcomes predefined by the corresponding observer state $|s_i\rangle$. We may condition on the observation outcomes, such that the object-environment subsystem is in the corresponding *local* state $|\phi_i\rangle$. This state may be Schmidt decomposed further (assuming $n_e \geq n$):

$$|\phi_i\rangle = \sum_{j=1}^{n} c_j^{(i)}|\psi_j^{(i)}\rangle \otimes |\psi_j^{(i)}\rangle_e\ , \tag{5}$$

where $|\psi_j^{(i)}\rangle \in H$ and $|\psi_j^{(i)}\rangle_e \in H_e$ $(j = 1, \cdots, n)$ are orthonormal states of the object and the environment, respectively. Note that they are, in general, different across different branches. Partial-tracing $|\phi_i\rangle\langle\phi_i|$ over the environment gives the *local* object state

$$\rho^{(i)} = \sum_{j=1}^{n} p_j^{(i)}|\psi_j^{(i)}\rangle\langle\psi_j^{(i)}|\ , \tag{6}$$

where $p_j^{(i)} = |c_j^{(i)}|^2$ is the conditional probability of the object being in a microstate $|\psi_j^{(i)}\rangle$, as perceived by the observer in $|s_i\rangle$. The von Neumann entropy formula then gives rise to the thermodynamic entropy:

$$-\mathrm{tr}\left(\rho^{(i)}\log\rho^{(i)}\right) = -\sum_{j=1}^{n} p_j^{(i)}\log p_j^{(i)}\ . \tag{7}$$

III. QUANTUM SZILARD ENGINE

In the previous section, we have demonstrated that the density matrix after conditioning and partial-tracing gives rise to the ordinary thermodynamic entropy via the von Neumann entropy formula. Now we would like to define a new form of entropy that is consistent with Everettian quantum mechanics. As illustrated in [7, 19], thermodynamic entropy changes in opposite directions when an object interacts with its environment (entropy increases) or with an observer (entropy on average decreases). Such perspective dependence makes the application of entropic descriptions less ubiquitous. Suppose we divide a system into two, an object and the rest part. It is preferable to use physical descriptions that are *perspective-free*, irrespective of the second part being either a sentient observer (or a subject) or some physical environment without sentience.

The preference of using a *perspective-free* description is consistent with Everettian quantum mechanics, which treats objects and subjects on an equal footing. The object and subject may be entangled in a way such that there is a one-to-one correspondence between definite object states and subject states, e.g. $|\uparrow\rangle$ corresponding to $|\text{perceiving}\uparrow\rangle$ and $|\downarrow\rangle$ corresponding to $|\text{perceiving}\downarrow\rangle$. In this sense, the object-subject pair undergoes apparent branching such that the subject in each branch perceives a definite object state, a "relative state" with respect to a particular subject state as formulated in Everett's original paper [8]. Equivalently, we could also claim that this subject state is a "relative state" with respect to the corresponding object state. The *perspective-free* descriptions, e.g. the density matrix $(|\uparrow\rangle\langle\uparrow|\otimes|\text{perceiving}\uparrow\rangle\langle\text{perceiving}\uparrow|+|\downarrow\rangle\langle\downarrow|\otimes|\text{perceiving}\downarrow\rangle\langle\text{perceiving}\downarrow|)/2$, are arguably symmetric if we switch the roles of the object and the subject. In accordance with this feature of Everett's interpretation, *physical* descriptions should be equally applicable to any physical system, regardless of whether it contains a sentient observer. We will therefore consider the bipartite system as if it is composed of an object and the rest part being a sentient observer. The results should hold as well for physical systems without sentience.

When the object is entangled with the sentient observer, the observer experiences a *local* reality that differs from the *global* reality. It suggests that only knowing the *global* reality is *insufficient*, for the observer also knows in which branch she *actually* lives (referred as her self-locating knowledge) when recognizing the *local* reality. One way to interpret this is the self-locating knowledge effectively "localizes" the *global* reality to the *local* reality. Note that it is not only *insufficient* but also *unnecessary* to know the *global* reality [19]. To know the object state in the observer's own branch, the way in which different branches combine (e.g. phase factor of each branch) needs not come into the picture. Accordingly, a *complete* description of the object should consist of both a coarse-grained description of the object (obtained by partial-tracing over the rest of the system) and the self-locating knowledge acquired by the observer. We will show that the two parts give rise to two corresponding forms of entropy.

Before defining the two forms of entropy, we may further illustrate the self-locating knowledge with the help of the quantum Szilard engine, a quantum-mechanical realization of Maxwell's demon [20, 21]. The quantum Szilard engine may be illustrated by the isolated box shown in Fig. 1 (now assume the wall separating the left and the right rooms is movable). The observer perceives either an empty or a non-empty left room. Provided that the left room is empty, she will attach a load to the wall and let the right room expand. This extracts work from the engine (see Fig. 2 for an illustration of the process). Essentially, the observer utilizes her self-locating knowledge to extract work, in analogy with Maxwell's demon. This shows that self-locating knowledge is *physical* in the sense that it exerts impact on the physical world.

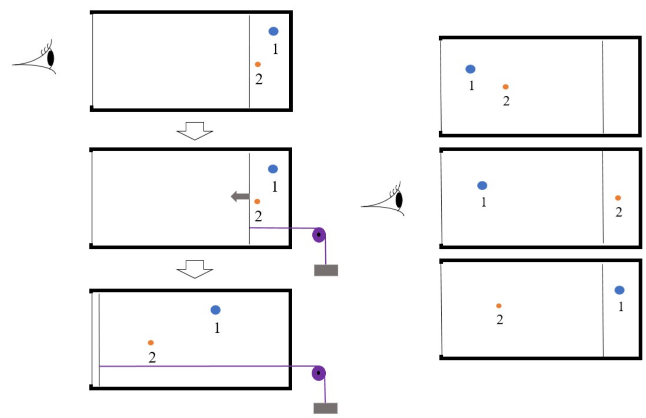

FIG. 2. Illustrative graph of the quantum Szilard engine.

We will analyze the quantum Szilard engine illustrated in Fig. 2. Suppose the box is a thermal reservoir at inverse temperature β, and particles on the left and the right sides of the wall are at the respective energy levels of 0 and $\varepsilon \gg 1/\beta$. Note that the two uneven energy levels may be achieved by an uneven partition, such that particles on the right side are initially confined in an extremely small room. The observer finds herself perceiving either an empty left room with probability $p \approx e^{-2\beta\varepsilon}$, or a non-empty room with probability $1-p \approx 1$. Such self-locating knowledge carries a Shannon information entropy of approximately $-p\ln p \approx 2\beta\varepsilon e^{-2\beta\varepsilon}$. In a single cycle, the quantum Szilard engine can, on average, extract $p\cdot 2\varepsilon \approx 2\epsilon e^{-2\beta\varepsilon}$ amount of work. This equals the Shannon entropy of the self-locating knowledge, multiplying a temperature coefficient $1/\beta$ (see [20] for a completely general analysis).

Note that the sentient observer is not essential for extracting work. All we need is an automated device that attaches a load to the wall if and only if the wall is in the state $|n\rangle_W$ (i.e. none of the particles are on its left side). This again implies that self-locating knowledge is *physical*, in the sense that it does not require sentience. It is therefore reasonable to talk about self-locating knowledge "acquired by" the environment of an object, even if it is not a sentient observer. In the following section, we will use the two terms "object" and "environment" to denote the two parts of a bipartite system, regardless of their properties related to sentience.

IV. BIPARTITE SYSTEM

In the previous section, we have demonstrated that the *complete* description of an object consists of a coarse-grained description of the object alone, plus the self-locating knowledge acquired by its environment. We may assign one entropy to measure the uncertainty of the former, namely the *objective entropy*, and assign another to measure the information carried by the latter, namely the *subjective entropy*.

We have demonstrated in Sec. II that the von Neumann entropy gives rise to the coarse-grained thermodynamic entropy (see Eq. 2). This implies that the *objective entropy*, which measures the uncertainty of the coarse-grained object state, takes the form of the von Neumann entropy formula

$$S_{\text{objective}} = -\text{tr}\left(\rho \log \rho\right) \ , \tag{8}$$

where the density matrix ρ represents the coarse-grained object state, obtained by partial-tracing the state of the entire system over the environment.

In the previous example of the quantum Szilard engine (Fig. 2), we have shown that the average work performed by the engine is proportional to the Shannon entropy of the self-locating knowledge. It therefore makes sense to express the *subjective entropy* by the (negative) Shannon entropy of the self-locating knowledge

$$S_{\text{subjective}} = \sum_k p_k \log p_k \ , \tag{9}$$

where p_k denotes the probability of the k-th branch and the summation is over all branches. Note that Eq. 9 only applies when the *local* state of the object is pure in each branch. In Sec. VI, we will generalize it to cases with mixed *local* states by replacing the Shannon entropy with the quantum relative entropy [22].

We may illustrate the *objective* and *subjective entropies* with a simple example. Consider an object (e.g. an electron) initially in a pure state, represented by a state vector $(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ or equivalently a density matrix $(|\uparrow\rangle\langle\uparrow| + |\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2$. In the process of decoherence, the object interacts with its environment ϵ and branches. Together they form an entangled state $(|\uparrow, \epsilon_\uparrow\rangle + |\downarrow, \epsilon_\downarrow\rangle)/\sqrt{2}$. By partial-tracing over ϵ, we end up with a density matrix with off-diagonal elements removed: $(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|)/2$. The *objective entropy* given by Eq. 9 increases from 0 before decoherence to 1 bit after decoherence, This gives an appearance that decoherence causes 1 bit of information loss.

But actually the 1 bit of information is now encoded in the self-locating knowledge instead of being lost forever. After decoherence, both branches (i.e. $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$) have 50% probabilities. According to Eq. 9, the *subjective entropy* decreases from 0 to -1 bit, compensating the increase in the *objective entropy* exactly. Therefore, rather than claiming entropy increases as a result of decoherence, we may instead state that entropy transfers between two alternative forms.

To understand the physical meaning of the statement, consider the environment asking "In which branch am I living?" The answer to this question gives the self-locating knowledge. Before recognizing the answer, the best estimation of the object state is the indeterministic *global* state, 50% $|\uparrow\rangle$ and 50% $|\downarrow\rangle$. After recognizing the self-locating knowledge, however, the environment is capable of reducing this indeterministic *global* state to a deterministic *local* state: either 100% $|\uparrow\rangle$ in the branch of $|\epsilon_\uparrow\rangle$, or 100% $|\downarrow\rangle$ in the branch of $|\epsilon_\downarrow\rangle$. This means that the self-locating knowledge removes any uncertainty of the *global* state, and therefore carries exactly 1 bit of information, which may be utilized by a quantum Szilard engine to extract work.

In this example, the sum of the *objective* and *subjective entropies* remains unchanged under decoherence. Accordingly, we define the *perspective-free entropy* of an object as the sum of the two entropies:

$$S = S_{\text{objective}} + S_{\text{subjective}} \ . \tag{10}$$

S is *perspective-free* in the sense that it is invariant under perspective changes. To illustrate this point, we will define the *apparent entropy* as the entropy perceived by the observer, essentially equivalent to the concept of thermodynamic entropy. In the example above, the *apparent entropy* equals the *objective entropy*, which increases as a result of decoherence. Despite the fact that the *subjective entropy* decreases, it is not part of the *apparent entropy*. This is because the *subjective entropy* measures the self-locating knowledge acquired by the environment. Any "ideal observer" *outside* the bipartite system is incapable of accessing it to reduce the uncertainty of the *global* object state.

On the other hand, in the process of measurement the *apparent entropy* decreases on average [19]. This is because the observer is now *inside* the system, thus no longer an "ideal observer". She acquires self-locating knowledge during the measurement (this will be discussed in detail in the following section). As a result, the *apparent entropy* not only includes the *objective entropy*, but also a decreasing part of the *subjective entropy*. We may interpret the apparent difference between decoherence and measurement as a perspective difference: between a perspective *outside* the system and a perspective *inside* the system. Using the *perspective-free entropy* makes the description independent of any specific perspective, or in some sense the subject-object division. This implies the *perspective-free entropy* is a conserved quantity.

To show that *perspective-free entropy* is indeed conserved under decoherence, we consider a general case in which there exists an orthonormal basis $\{|o_i\rangle\}$ of the object, such that the object-environment interaction follows the unitary transformation

$$|o_i\rangle|\epsilon^*\rangle \rightarrow |o_i\rangle|\epsilon_i\rangle \ , \tag{11}$$

where $|\epsilon^*\rangle$ is the state of the environment prior to the interaction, $|o_i\rangle$ is a basis state of the object, and $|\epsilon_i\rangle$ is the corresponding final state of the environment. We assume perfect decoherence in the sense that the final states of the environment are orthogonal, i.e. $\langle\epsilon_i|\epsilon_j\rangle = 0$ for all i and j such that $i \neq j$. We define $\{|\epsilon_i\rangle\}$ as the "pointer basis" of the environment. The name comes from quantum Darwinism, which claims that pointer states are those that can be cloned to the environment without being changed [18]. This means that pointer states need to satisfy Eq. 11.

We will analyze decoherence using the language of density matrix. In the following context, we use the function $S(\cdot)$ to denote either von Neumann or Shannon entropy: if it takes a density matrix as its argument, then it denotes von Neumann entropy; if it takes a set of probabilities as its argument, then it denotes Shannon entropy. Now consider the case in which the object is originally in a pure state ρ, such that the von Neumann entropy $S(\rho) = 0$. According to Eqs. 8 and 9, $S_{\text{objective}} = S_{\text{subjective}} = 0$.

After decoherence, the bipartite system branches. Each branch, defined by a "pointer basis" state $|\epsilon_k\rangle$, corresponds to a density matrix $\rho^{(k)} = |o_k\rangle\langle o_k|$ according to Eq. 11. In other words, conditioning to $|\epsilon_k\rangle$ gives the *local* object state $\rho^{(k)}$. The coarse-grained density matrix defined by partial-tracing, $\tilde{\rho} := \mathrm{tr}_\epsilon \left(\sum_k p_k|o_k\rangle\langle o_k| \otimes |\epsilon_k\rangle\langle\epsilon_k|\right)$, equals the average of $\rho^{(k)}$ over all branches, i.e. $\tilde{\rho} = \mathbb{E}[\rho^{(k)}]$ (this illustrates how decoherence removes off-diagonal elements, for each $\rho^{(k)}$ only contains a diagonal element). Note that the probability p_k is given by $\langle o_k|\rho|o_k\rangle$.

Due to the purity of the *local* state $\rho^{(k)}$ the von Neumann entropy $S(\tilde{\rho}) = -\mathrm{tr}(\tilde{\rho}\log\tilde{\rho})$ equals the classical Shannon entropy $S(\{p_k\}) = -\sum_k p_k \log p_k$. According to Eqs. 8 and 9, $S_{\text{objective}}$ and $S_{\text{subjective}}$ compensates each other exactly. This means that the self-locating knowledge removes any uncertainty in the coarse-grained description of the object.

In summary, under perfect decoherence $S_{\text{objective}}$ increases from 0 to $S(\{p_k\})$ while $S_{\text{subjective}}$ decreases from 0 to $-S(\{p_k\})$. Their summation, defined as the *perspective-free entropy*, is conserved (see Tab. I). Decoherence therefore leads to a transfer (rather than a net increase) of entropy between two forms, interpreted as a process in which the environment acquires self-locating knowledge (rather than a process in which the information gets lost forever).

TABLE I. Entropies of an object in a bipartite system (assuming perfect decoherence). The *perspective-free entropy* is conserved under decoherence.

Entropy	Objective	Subjective	Perspective-free
Before decoherence	0	0	0
After decoherence	$S(\{p_k\})$	$-S(\{p_k\})$	0

V. TRIPARTITE SYSTEM

By dividing a system into an object and the remaining part (referred as the "environment" in the previous section), we have shown that *perspective-free entropy* is conserved under decoherence. However, to describe a realistic measurement process, one needs to further divide this remaining part to leave some room for an observer. That is, the system we consider should include explicitly both the environment and the observer, thus an object-environment-observer tripartite system. We assume the entire system is always in pure states and consider a measurement conducted by the observer.

Before the measurement, the object being measured is in a mixed state, represented by a density matrix $\rho = \sum_i |o_i\rangle p_i \langle o_i|$. We further assume that the object-environment subsystem is in a pure state. It can then be considered as the bipartite system discussed in the previous section, in which the *objective entropy* $S(\rho)$ is offset by a *subjective entropy* $-S(\rho)$ measuring the self-locating knowledge acquired by the environment.

After the measurement, the observer is entangled with the object, and acquires self-locating knowledge which determines the *local* state of the object. In entropic terms, the *apparent entropy* perceived by the observer now incorporates a non-zero part of the *subjective entropy*, and therefore decreases after the measurement [19]. To explain this entropy decrease, we need to divide the *subjective entropy* into two parts, one of which measures the self-locating knowledge acquired by the observer and the other measures the self-locating knowledge of the environment. Eq. 10 is therefore modified to

$$S = S_{\text{objective}} + S_{\text{subjective}}^{(\text{observer})} + S_{\text{subjective}}^{(\text{environment})} \ , \tag{12}$$

where the second and third terms on the right-hand side denote the *subjective entropy* of the observer and the *subjective entropy* of the environment, respectively. Their summation equals the total *subjective entropy* $S_{\text{subjective}}$ as given in Eq. 10.

Similar to Eq. 11, the process of measurement follows an unitary transformation

$$|\tilde{o}_k\rangle|s^*\rangle \rightarrow |\tilde{o}_k\rangle|s_k\rangle \ , \tag{13}$$

where $\{|\tilde{o}_k\rangle\}$ is a basis of the object. $|s^*\rangle$ and $|s_k\rangle$ are observer states before and after the measurement. We assume $|s_k\rangle$ is a state that can be directly perceived. This essentially means that the corresponding object state $|\tilde{o}_k\rangle$ is a pointer state. We therefore define $\{|s_k\rangle\}$ as the "pointer basis" of the observer. We will keep this assumption throughout the rest of this section, and relax it in the following sections.

In practice, the partition of the system is somewhat arbitrary. We may regard it as a factorization of the entire Hilbert space: $H_{\text{total}} = H \otimes H_e \otimes H_s$. Its logarithmic dimension $\log n_{\text{total}}$ is defined as the number of total degrees of freedom. Among these degrees of freedom we find out $\log n$ (assuming $n \ll n_{\text{total}}$) degrees of freedom which are responsible for the observer's direct perception. These degrees of freedom are used to construct the observer Hilbert space H_s, spanned by the "pointer basis" $|s_k\rangle$ $(k = 1, \cdots, n)$. We then find out the n corresponding states $|\tilde{o}_k\rangle$ $(k = 1, \cdots, n)$ according to Eq. 13, which span the object Hilbert space H. The remaining $\log n_{\text{total}} - 2\log n$ degrees of freedom are recognized as the environment degrees of freedom for constructing the environment Hilbert space H_e.

Under such system partition, there is a one-to-one correspondence between the n observer states, $|s_k\rangle$, and the n object states, $|\tilde{o}_k\rangle$, where $k = 1, \cdots, n$. We will see that after the measurement, the observer in one of these "pointer basis" states, $|s_k\rangle$, is certain that the object is in the corresponding pure state, $|\tilde{o}_k\rangle$. Therefore, the measurement results in the *apparent entropy* of the object decreasing from $S(\rho)$ to 0.

Before the measurement, the object-environment subsystem is in a pure state, which can be written as $\sum_i a_i|\epsilon_i\rangle|o_i\rangle$. This gives the object state $\rho = \sum_i |o_i\rangle p_i \langle o_i|$ where $p_i = |a_i|^2$. It is noted that the branching of the object state is determined by the "pointer basis", $\{|\epsilon_i\rangle\}$, of the environment (plus the unitary transformation of the object-environment interaction, Eq. 11). After the measurement, the object-environment subsystem gets entangled with the observer. We may write the unitary transformation of the entire system with the help of Eq. 13:

$$\sum_i a_i|\epsilon_i\rangle|o_i\rangle|s^*\rangle = \sum_{i,k} a_i|\epsilon_i\rangle|\tilde{o}_k\rangle\langle\tilde{o}_k|o_i\rangle|s^*\rangle \tag{14}$$

$$\to \sum_k \left(\sum_i a_i\langle\tilde{o}_k|o_i\rangle|\epsilon_i\rangle\right)|\tilde{o}_k\rangle|s_k\rangle \,.$$

According to Eq. 14, the observer living in the k-th branch, defined by the observer state $|s_k\rangle$, perceives a pure state $|\tilde{o}_k\rangle$ and concludes that the object has a zero *apparent entropy*. In the language of density matrix, the *local* state of the object may be expressed by $\rho^{(k)} = |\tilde{o}_k\rangle\langle\tilde{o}_k|$.

The self-locating probability of the observer being in the k-th branch is given by

$$\tilde{p}_k = \sum_i |a_i|^2|\langle\tilde{o}_k|o_i\rangle|^2 = \langle\tilde{o}_k|\,\rho\,|\tilde{o}_k\rangle \,. \tag{15}$$

Partial-tracing Eq. 14 over both the observer and the environment gives the coarse-grained *global* state $\tilde{\rho} := \sum_k |\tilde{o}_k\rangle\tilde{p}_k\langle\tilde{o}_k|$. According to Eq. 8, the *objective entropy* changes from $S(\rho)$ before the measurement to $S(\tilde{\rho}) = S(\{\tilde{p}_k\})$ after the measurement.

Before the measurement, the *subjective entropy* of the observer is zero, as she acquires no self-locating knowledge. After the measurement, the *subjective entropy* of the observer decreases to $\sum_k \tilde{p}_k \log \tilde{p}_k = -S(\{\tilde{p}_k\})$ according to Eq. 9. This decrease compensates the increase in the *objective entropy* more than enough. The *apparent entropy*, which equals the sum of the *objective entropy* and the *subjective entropy* of the observer, drops to zero, suggesting the observer acquires sufficient self-locating knowledge to remove any uncertainty in the *global* description of the object. See Tab. II for the entropies before and after the measurement.

In the process of measurement, the branching of the object state switches from being determined by the "pointer basis" of the environment, $\{|\epsilon_i\rangle\}$, to being determined by the "pointer basis" of the observer, $\{|s_k\rangle\}$. Knowing the observer state $|s_k\rangle$ is now sufficient to reduce the *global* mixed state $\tilde{\rho}$ to a *local* pure state $\rho^{(k)}$. Further acquiring the self-locating knowledge of the environment would not deepen the observer's understanding of the object. In entropic terms, the *subjective entropy* of the environment increases from $-S(\rho)$ back to 0. This change is in the opposite direction compared to the *subjective entropy* of the observer, which decreases from 0 to $-S(\tilde{\rho})$. One may interpret this as the observer competing against the environment in order to acquire more self-locating knowledge (or equivalently to lower her *subjective entropy*). As a result, the previous branches defined by the environment are completely overridden, illustrated in Fig. 3(a), by the new branches.

TABLE II. Entropies of an object in a tripartite system (assuming measurement leads to a final state $|s_k\rangle$ which belongs to the "pointer basis" of the observer). The *perspective-free entropy* is conserved under measurement. Note that the *apparent entropy* from the observer's perspective accounts for both the *objective entropy* and the *subjective entropy* of the observer.

Entropy	Objective	Subjective (observer)	Apparent = Objective + Subjective (observer)	Subjective (environment)	Perspective-free
Before measurement	$S(\rho)$	0	$S(\rho)$	$-S(\rho)$	0
After measurement	$S(\{\tilde{p}_k\})$	$-S(\{\tilde{p}_k\})$	0	0	0

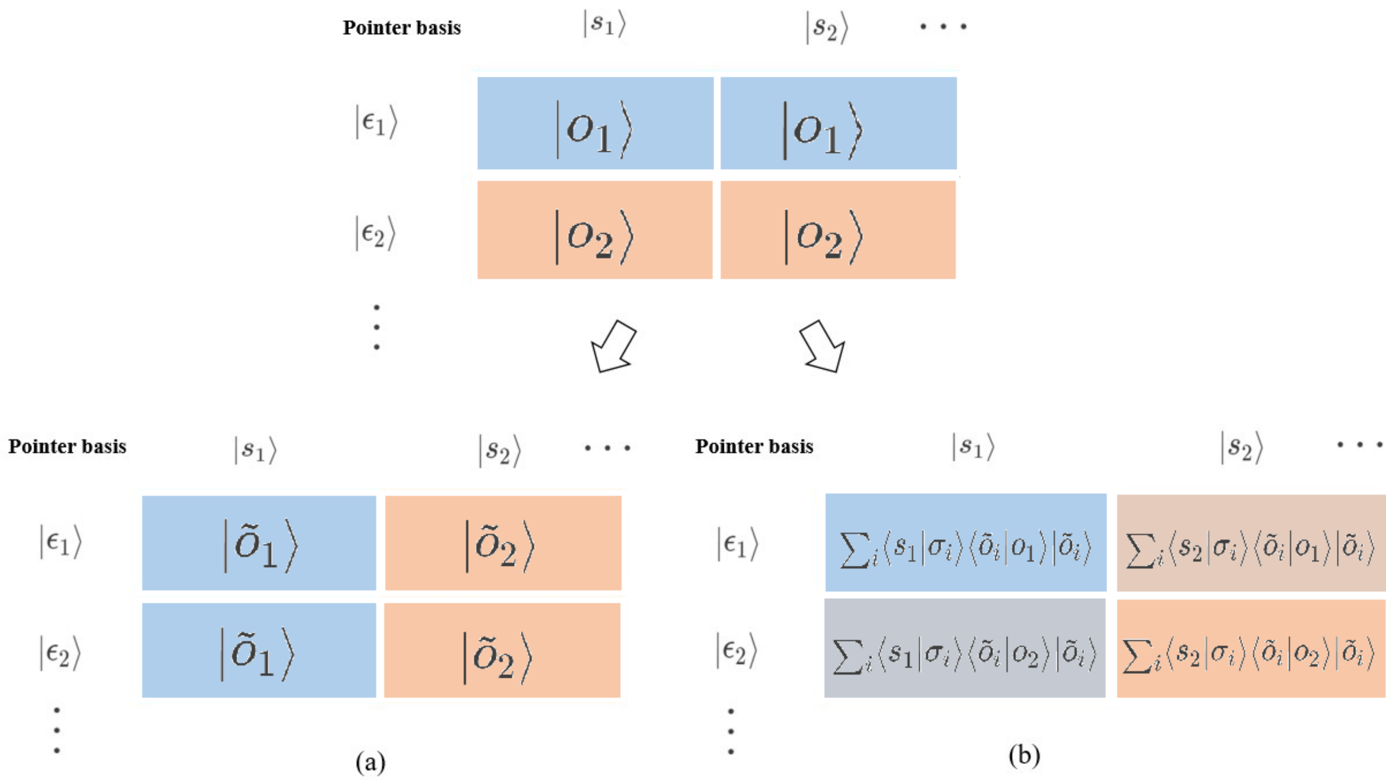

FIG. 3. (a) The previous branches, determined by the "pointer basis" of the environment, are completely overridden by the new branches, determined by the "pointer basis" of the observer, in a process described by Eq. 14. (b) The previous branches are partially overridden by the new branches in a process described by Eq. 16. The row indices are "pointer basis" states, $|\epsilon_i\rangle$, of the environment, the column indices are "pointer basis" states, $|s_k\rangle$, of the observer, and the table cells give the corresponding *local* states of the object.

As an implication, we may regard decoherence and measurement as two reciprocal processes: decoherence from the perspective of the observer is measurement from the perspective of the environment, and vice versa. In entropic terms, decoherence and measurement transfer *subjective entropy* between the two components, the observer and the environment, in opposite directions. From the perspective of the observer, she can utilize her self-locating knowledge to determine the *local* object state, thus perceiving an *apparent entropy* that includes both the *objective entropy* and the *subjective entropy* of the observer (see the black solid circle in Fig. 4).

In Fig. 4, since measurement transfers *subjective entropy* out (from the observer to the environment), the *apparent entropy* decreases. On the other hand, in the process of decoherence, *subjective entropy* flows in the opposite direction. This is illustrated in the previous section, where conditioning to an environment state $|\epsilon_i\rangle$ results in a pure object state $|o_i\rangle$ after decoherence. This implies the negative *subjective entropy* of the environment fully offsets the positive *objective entropy*, leaving a zero *subjective entropy* of the observer. As a result, decoherence transfers *subjective entropy* from the environment to the observer, thus increasing the *apparent entropy*.

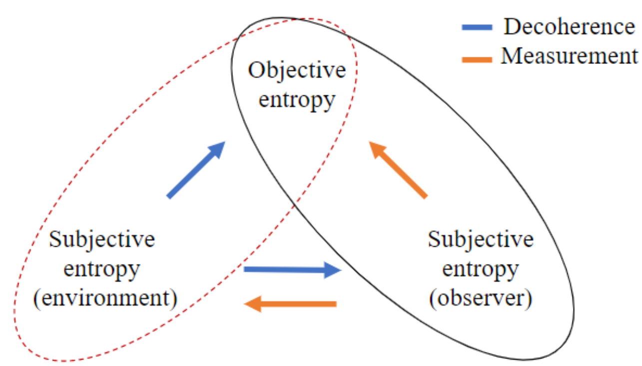

FIG. 4. Transfers between different forms of entropy in the processes of decoherence and measurement. The black solid circle indicates the *apparent entropy* from the perspective of the observer. The red dotted circle indicates the *apparent entropy* from the perspective of the environment.

We may change the definition of the *apparent entropy* by switching our perspective. If our perspective is instead from the environment side, then the *apparent entropy* includes the *subjective entropy* of the environment instead of the observer. Measurement, which transfers *subjective entropy* from the observer to the environment, now increases the *apparent entropy.* Decoherence, which transfers *subjective entropy* in the opposite direction, now decreases the *apparent entropy.* This provides an entropic characterization of the reciprocity between decoherence and measurement. We will therefore only focus on the process of measurement when discussing the general results in Sec. VI and VII. For the process of decoherence, the same results apply if we switch the roles of the environment and the observer.

VI. GENERAL RESULT

In this section, we will relax the assumption that measurement leads to a final state which belongs to the "pointer basis" of the observer as illustrated by Eq. 13. Instead we consider a general measurement which follows the unitary transformation $|\tilde{o}_k\rangle|s^*\rangle \rightarrow |\tilde{o}_k\rangle|\sigma_k\rangle$, where the final state, $|\sigma_k\rangle$, is now different from the "pointer basis" state, $|s_k\rangle$, that can be directly perceived by the observer. This general case requires us to modify Eq. 14 to

$$\begin{aligned}\sum_i a_i|\epsilon_i\rangle|o_i\rangle|s^*\rangle &\rightarrow \sum_{i,j} a_i\langle\tilde{o}_j|o_i\rangle|\epsilon_i\rangle|\tilde{o}_j\rangle|\sigma_j\rangle \\ &= \sum_{i,j,k} a_i\langle\tilde{o}_j|o_i\rangle\langle s_k|\sigma_j\rangle|\epsilon_i\rangle|\tilde{o}_j\rangle|s_k\rangle \ .\end{aligned} \tag{16}$$

After the measurement, in the k-th branch (where the observer finds herself perceiving $|s_k\rangle$), the *local* state of the object-environment subsystem is an entangled state $\sum_{i,j} a_i\langle\tilde{o}_j|o_i\rangle\langle s_k|\sigma_j\rangle|\epsilon_i\rangle|\tilde{o}_j\rangle$ (ignoring any normalization factor).

By partial-tracing over the environment, we find that in the k-th branch, the object is generally in a mixed state, represented by the density matrix

$$\begin{aligned}\rho^{(k)} =& \frac{1}{\tilde{p}_k}\sum_{ij_1j_2}|a_i|^2\langle\tilde{o}_{j_1}|o_i\rangle\langle o_i|\tilde{o}_{j_2}\rangle\langle s_k|\sigma_{j_1}\rangle\langle\sigma_{j_2}|s_k\rangle|\tilde{o}_{j_1}\rangle\langle\tilde{o}_{j_2}| \\ =& \sum_{j_1j_2}|\tilde{o}_{j_1}\rangle\frac{\langle\tilde{o}_{j_1}|\rho|\tilde{o}_{j_2}\rangle\langle s_k|\sigma_{j_1}\rangle\langle\sigma_{j_2}|s_k\rangle}{\tilde{p}_k}\langle\tilde{o}_{j_2}| \ ,\end{aligned} \tag{17}$$

where $\rho = \sum_i |a_i|^2|o_i\rangle\langle o_i|$ is the density matrix of the object before the measurement. The normalization factor $\tilde{p}_k$ is given by the (self-locating) probability of the k-th branch: $\tilde{p}_k = \sum_j\langle\tilde{o}_j|\rho|\tilde{o}_j\rangle|\langle s_k|\sigma_j\rangle|^2$. Eq. 17 is called the quantum-mechanical Bayes' theorem [7].

It is noted that in this general case, the observer can no longer use her self-locating knowledge to reduce the object state to a pure state. In other words, the previous branches, determined by the "pointer basis" of the environment $\{|\epsilon_i\rangle\}$, are now only partially overridden by the new branches, determined by the "pointer basis" of the observer $\{|s_k\rangle\}$. This is different from the special result derived in the previous section. The difference is visualised in Fig. 3. The result in the previous section, illustrated in Fig. 3(a), associates different environment states with completely identical object states. Knowing which observer state is perceived (e.g. $|s_1\rangle$) is sufficient to determine a pure state of the object (e.g. $|\tilde{o}_1\rangle$). Therefore, the self-locating knowledge of the environment is redundant, measured by a zero *subjective entropy* of the environment (see Table II).

In contrast, in this general case different environment states are no longer associated with identical object states, as illustrated in Fig. 3(b). Consider the branch in which the observer perceives $|s_k\rangle$. The *local* state of the object is now a statistical ensemble, represented by the density matrix $\rho^{(k)}$. This mixed state could be "purified" by further utilizing the self-locating knowledge of the environment. For instance, by knowing that the environment is *actually* in $|\epsilon_i\rangle$, one may reduce the mixed state $\rho^{(k)}$ to a pure state $\sum_j a_i\langle\tilde{o}_j|o_i\rangle\langle s_k|\sigma_j\rangle|\tilde{o}_j\rangle$. We can see that the self-locating knowledge of the environment is no longer redundant, measured by a negative *subjective entropy* of the environment.

The quantum-mechanical Bayes' theorem, Eq. 17, is a special case of the positive-operator-valued measure (POVM). In POVM, the measurement outcomes take values from a set of positive semi-definite Hermitian operators $\{F_k\}$, where $\sum_k F_k$ equals the identity operator, and the probability of the k-th outcome is given by $\tilde{p}_k = \mathrm{tr}(\rho F_k)$ [22]. These operators take the form $F_k = \sum_j|\tilde{o}_j\rangle|\langle s_k|\sigma_j\rangle|^2\langle\tilde{o}_j|$ in the special case of the quantum-mechanical Bayes' theorem. In the remaining part of the section, we will demonstrate the conservation of *perspective-free entropy* under general POVM.

In fact, each positive semi-definite Hermitian F_k has a decomposition $F_k = M_k^\dagger M_k$ such that the post-measurement object state is given by $\rho^{(k)} = M_k\rho M_k^\dagger\,/\,\tilde{p}_k$. Note that this immediately gives the quantum-mechanical Bayes' theorem, Eq. 17, if we assign $M_k = \sum_j|\tilde{o}_j\rangle\langle s_k|\sigma_j\rangle\langle\tilde{o}_j|$. Now suppose the observer ignores any self-locating knowledge (i.e. an "ideal observer"). Then she could only describe the object by the coarse-grained density matrix $\tilde{\rho} := \sum_k\tilde{p}_k\rho^{(k)}$. Accordingly, the *objective entropy* changes from $S(\rho)$ to $S(\tilde{\rho})$ under the measurement.

A realistic observer, however, acquires self-locating knowledge after the measurement, which can be utilized to determine the *local* object state. In the k-th branch, the information gain of recognizing the self-locating knowledge is measured by her *subjective entropy* decreasing from 0 before the measurement to $-S(\rho^{(k)}||\,\tilde{\rho})$ after the measurement. Here the quantum-mechanical version of relative entropy takes the form of [22]

$$S\left(\rho^{(k)}\,||\,\tilde{\rho}\right) = -\mathrm{tr}\left(\rho^{(k)}\log\tilde{\rho}\right) - S\left(\rho^{(k)}\right) \ . \tag{18}$$

The non-negativity of the quantum relative entropy follows from Klein's inequality [23]. This guarantees the *subjective entropy* $-S(\rho^{(k)}||\,\tilde{\rho})$ to be non-positive, consistent with the fact that self-locating knowledge always carries an information gain.

From the information-theoretic point of view, the use of the quantum relative entropy illustrates an information updating procedure. In the absence of self-locating knowledge, the observer can at best estimate the object state by the coarse-grained density matrix $\tilde{\rho}$ (which is predictable using the object state before the measurement). After realizing she *actually* lives in the k-th branch, the observer immediately updates her estimate to $\rho^{(k)}$. Note that this information update needs not to be realized by a physical process. In fact, after the measurement the observer *experiences* only a *local* reality, which implicitly involves updating the object state to the *local* state.

The information gain inherent to this updating procedure is measured by the relative entropy $S(\rho^{(k)}\,||\,\tilde{\rho})$. It therefore provides a proper measure of the *subjective entropy*. Alternatively, one may interpret it by the quantum Szilard engine: the average work performed by the engine is given by the relative entropy of the distributions between forward and backward processes [20, 24]. Suppose after each operation the engine is decohered from the observer and its *local* state is identical to the *global* state. Then the engine has *local* states $\rho^{(k)}$ and $\tilde{\rho}$ before and after the operation, thus capable of extracting $S(\rho^{(k)}\,||\,\tilde{\rho})\,/\,\beta$ amount of work.

The *subjective entropy* averaged over all branches is given by

$$\begin{aligned} S_{\text{subjective}}^{(\text{observer})} &= -\,\mathbb{E}\left[S\left(\rho^{(k)}\,||\,\tilde{\rho}\,\right)\right] \\ &= \operatorname{tr}\left[\mathbb{E}\left(\rho^{(k)}\right)\log\tilde{\rho}\right] + \mathbb{E}\left[S\left(\rho^{(k)}\right)\right] \\ &= -\,S\left(\tilde{\rho}\right) + \mathbb{E}\left[S\left(\rho^{(k)}\right)\right] \ . \end{aligned} \tag{19}$$

It is noted that the *subjective entropy* formula, Eq. 9, used in previous sections is a special case of Eq. 19. It holds only when $\rho^{(k)}$ is pure for all k. In such cases, the second term on the right-hand side of Eq. 19 is zero, while the first term equals the Shannon entropy, i.e. $S(\tilde{\rho}) = S(\{\tilde{p}_k\})$. This immediately leads to Eq. 9.

Now let's discuss the *subjective entropy* of the environment after the measurement. By applying the projection operator $|s_k\rangle\langle s_k|$ to the *global* state of the entire system, we may find out the *local* state of the object-environment subsystem in the k-th branch. Since the entire system is in a pure state, the projected state is also pure, in contrast to the mixed state $\rho^{(k)}$ of the object alone. This suggests that the self-locating knowledge of the environment can be used to "purify" the object state $\rho^{(k)}$. By calculating the information gain (measured by the quantum relative entropy) of updating $\rho^{(k)}$ to a pure state, one immediately concludes that the *subjective entropy* of the environment equals $-S(\rho^{(k)})$ in the k-th branch. The average *subjective entropy* of the environment is then given by $-\,\mathbb{E}[S(\rho^{(k)})]$.

Table III summarizes the *objective* and *subjective entropies* involved before and after the measurement. We may verify our results by evaluating the *apparent entropy*, which equals the sum of the *objective entropy* and the *subjective entropy* of the observer:

$$\begin{aligned} S_{\text{objective}} + S_{\text{subjective}}^{(\text{observer})} &= S\left(\tilde{\rho}\right) - S\left(\tilde{\rho}\right) + \mathbb{E}\left[S\left(\rho^{(k)}\right)\right] \\ &= \mathbb{E}\left[S\left(\rho^{(k)}\right)\right] \ . \end{aligned} \tag{20}$$

The calculated *apparent entropy* equals exactly the entropy of the *local* state, $\rho^{(k)}$, as perceived by the observer, averaged over all branches indexed by k. This equality validates the values of the *objective* and *subjective entropies* listed in Table III.

In summary, when performing a general POVM, the observer acquires self-locating knowledge. This is measured by her *subjective entropy* decreasing from 0 to $-S(\rho^{(k)}\,||\,\tilde{\rho})$ in the k-th branch. On average, the decrease in *subjective entropy* compensates the increase in *objective entropy* more than enough. As a result, the observer is expected to perceive a decreasing *apparent entropy* [19]. For the environment component of the system, its *subjective entropy* increases from $-S(\rho)$ to $-\,\mathbb{E}[S(\rho^{(k)})]$. This may be interpreted as the measurement resulting in new branches, which makes the previous branches less relevant as illustrated in Fig. 3(b). In other words, the measurement decorrelates the object states and the environment states, recording an average increase in the *subjective entropy* of the environment.

VII. WIGNER'S MANY FRIENDS

We have shown that the *perspective-free entropy* is conserved under general POVM and in both bipartite and tripartite systems. In this section, we will generalize this result for systems with an arbitrary number of components. This generalization is inspired by the famous Wigner's friend thought experiment [25].

Here we consider a group of Wigner's friends labeled by their degrees of separation. For example, a first order friend is a friend of Wigner, and a second order friend is a friend of Wigner's friend, etc. Now suppose Wigner is watching all these friends performing a measurement in a nested laboratory (see Fig. 5). In this laboratory, the

TABLE III. Entropies of an object in a tripartite system (assuming general POVM). The *perspective-free entropy* is again conserved in this general case. Note that the standard version of the second law assumes that the observer is agnostic about self-locating knowledge and only accounts for the coarse-grained *objective entropy*, which increases from $S(\rho)$ to $S(\tilde{\rho})$. If the observer is aware of the self-locating knowledge, then the *apparent entropy* also needs to account for her *subjective entropy*. As a result, a measurement performed by the observer leads to the *apparent entropy* decreasing on average, from $S(\rho)$ to $\mathbb{E}[S(\rho^{(k)})]$.

Entropy	Objective	Subjective (observer)	Apparent = Objective + Subjective (observer)	Subjective (environment)	Perspective-free
Before measurement	$S(\rho)$	0	$S(\rho)$	$-S(\rho)$	0
After measurement	$S(\tilde{\rho})$	$-\mathbb{E}\left[S\left(\rho^{(k)} \parallel \tilde{\rho}\right)\right]$	$\mathbb{E}\left[S(\rho^{(k)})\right]$	$-\mathbb{E}\left[S(\rho^{(k)})\right]$	0

i-th order friends get together to watch the j-th order ($j = i+1, i+2, \cdots, N$) friends collectively performing the measurement. Such setup is iterative until the innermost part of the laboratory, in which the N-th order friends directly watch an apparatus (as part of the environment) measuring a spin state (as the object).

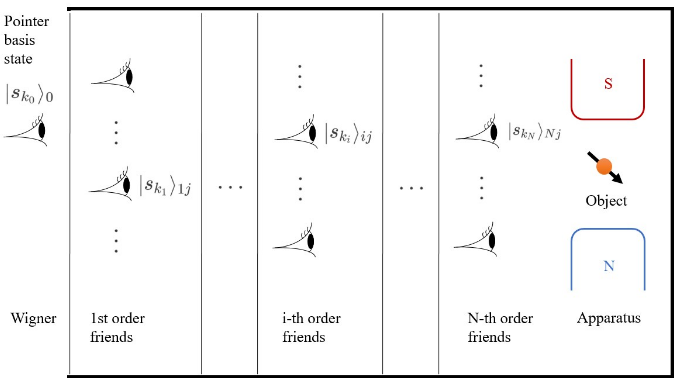

FIG. 5. Illustrative graph of Wigner's many friends.

In this thought experiment, the master observer, Wigner, and the object together form a subsystem that is in a mixed state after the measurement. Under general POVM, Wigner has a self-locating probability p_{k_0} of living in the k_0-th branch, which corresponds to one of his "pointer basis" states $|s_{k_0}\rangle_0$. In this branch, the *local* state of the object is represented by a density matrix $\rho^{(k_0)}$. This contributes a term

$$p_{k_0}\rho^{(k_0)} \otimes |s_{k_0}\rangle_0\langle s_{k_0}|_0 \tag{21}$$

to the density matrix of the object-Wigner subsystem. Partial-tracing over Wigner gives the coarse-grained object state $\tilde{\rho} = \sum_{k_0} p_{k_0}\rho^{(k_0)}$, where the summation is over all values of k_0, i.e. the indices of Wigner's "pointer basis" states.

Now if we include in this subsystem the first order friends indexed by j, we get finer branches determined by the product "pointer basis" states between Wigner and his first order friends. Each of these product states, denoted by $|s_{k_0}\rangle_0 \otimes (\otimes_j |s_{k_1}\rangle_{1j})$, corresponds to an object state $\rho^{(k_0k_1)}$ with a probability $p_{k_0k_1}$. This contributes

$$p_{k_0k_1}\rho^{(k_0k_1)} \otimes |s_{k_0}\rangle_0\langle s_{k_0}|_0 \otimes \left(\otimes_j\, |s_{k_1}\rangle_{1j}\langle s_{k_1}|_{1j}\right) \tag{22}$$

to the density matrix of the subsystem consisting of the object, Wigner, and his first order friends. After partial-tracing this density matrix over the first order friends, one should obtain the density matrix of the object-Wigner subsystem, Eq. 21. This condition imposes

$$\rho^{(k_0)} = \sum_{k_1} \frac{p_{k_0k_1}\rho^{(k_0k_1)}}{p_{k_0}} \quad \text{and} \quad p_{k_0} = \sum_{k_1} p_{k_0k_1}\ . \tag{23}$$

In general, the object state perceived by Wigner and his first i-th order friends collectively, who are in the product state of $|s_{k_0}\rangle_0 \otimes (\otimes_j |s_{k_1}\rangle_{1j}) \otimes \cdots \otimes (\otimes_j |s_{k_i}\rangle_{ij})$, can be expressed by a recursive relation

$$\rho^{(k_0k_1\cdots k_i)} = \sum_{k_{i+1}} \frac{p_{k_0k_1\cdots k_ik_{i+1}}\rho^{(k_0k_1\cdots k_ik_{i+1})}}{p_{k_0k_1\cdots k_i}} \tag{24}$$

for $i = 1, 2, \cdots, N-1$, where the probability

$$p_{k_0k_1\cdots k_i} = \sum_{k_{i+1}} p_{k_0k_1\cdots k_ik_{i+1}}\ . \tag{25}$$

Therefore, we can start from the object states $\rho^{(k_0k_1\cdots k_N)}$ and iterate backwards to get the density matrix $\rho^{(k_0k_1\cdots k_i)}$ for an arbitrary i.

We will describe this thought experiment using entropic terms. First, Wigner in his "pointer basis" state $|s_{k_0}\rangle_0$ can utilize his self-locating knowledge to reduce the *global* object state $\tilde{\rho}$ to a *local* state $\rho^{(k_0)}$. On average, this gives a *subjective entropy* of

$$\begin{aligned} S^{(\text{Wigner})}_{\text{subjective}} &= -\mathop{\mathbb{E}}_{k_0}\left[S\left(\rho^{(k_0)} \parallel \tilde{\rho}\right)\right] \\ &= -S\left(\tilde{\rho}\right) + \mathop{\mathbb{E}}_{k_0}\left[S\left(\rho^{(k_0)}\right)\right]\ . \end{aligned} \tag{26}$$

Secondly, suppose Wigner and his first $(i-1)$-th order friends find themselves in a product state $|s_{k_0}\rangle_0 \otimes (\otimes_j |s_{k_1}\rangle_{1j}) \otimes \cdots \otimes (\otimes_j |s_{k_{i-1}}\rangle_{(i-1)j})$. Together they will describe the object using the density matrix $\rho^{(k_0\cdots k_{i-1})}$.

Now if they could further access the self-locating knowledge acquired by any of the i-th order friends, they would immediately update the object state from $\rho^{(k_0\cdots k_{i-1})}$ to $\rho^{(k_0\cdots k_{i-1}k_i)}$. Following the information-theoretic arguments in the previous section, the *subjective entropy* of the i-th order friends is given by the quantum relative entropy $-S(\rho^{(k_0\cdots k_{i-1}k_i)} \,||\, \rho^{(k_0\cdots k_{i-1})})$. Using the probabilities given by Eq. 24, we average it over all values of k_i to get a result similar to Eq. 26:

$$S^{(i\text{-th})}_{\text{subjective}} = \mathop{\mathbb{E}}_{k_i}\left[S\left(\rho^{(k_0\cdots k_{i-1}k_i)}\right)\right] - S\left(\rho^{(k_0\cdots k_{i-1})}\right) \ . \tag{27}$$

Finally, Wigner and all his friends together are in a product state $|s_{k_0k_1\cdots k_N}\rangle := |s_{k_0}\rangle_0 \otimes (\otimes_j|s_{k_1}\rangle_{1j}) \otimes \cdots \otimes (\otimes_j|s_{k_N}\rangle_{Nj})$. Collectively they describe the object by a *local* state $\rho^{(k_0\cdots k_N)}$. If in addition we account for the self-locating knowledge of the environment (e.g. the measuring apparatus), then we could further update $\rho^{(k_0\cdots k_N)}$ to $\rho^{(k_0\cdots k_N k_s)}$, where $|\epsilon_{k_s}\rangle$ denotes a "pointer basis" state of the environment. Since we have now acquired all self-locating knowledge about the object, we expect the new *local* state $\rho^{(k_0\cdots k_N k_s)}$ to be pure. This pure state can be obtained simply by projecting the state vector of the entire system to the product state $|s_{k_0k_1\cdots k_N}\rangle \otimes |\epsilon_{k_s}\rangle$.

The self-locating knowledge of the environment contributes an information gain $S(\rho^{(k_0\cdots k_N k_s)} \,||\, \rho^{(k_0\cdots k_N)})$. Averaging over all values of k_s gives the *subjective entropy* of the environment:

$$\begin{aligned} S^{(\text{environment})}_{\text{subjective}} &= -\mathop{\mathbb{E}}_{k_s}\left[S\left(\rho^{(k_0\cdots k_N k_s)} \,||\, \rho^{(k_0\cdots k_N)}\right)\right] \\ &= \mathop{\mathbb{E}}_{k_s}\left[\operatorname{tr}\left(\rho^{(k_0\cdots k_N k_s)} \log \rho^{(k_0\cdots k_N)}\right)\right] \\ &= \operatorname{tr}\left[\mathop{\mathbb{E}}_{k_s}\left(\rho^{(k_0\cdots k_N k_s)}\right) \log \rho^{(k_0\cdots k_N)}\right] \\ &= -S\left(\rho^{(k_0\cdots k_N)}\right) \ , \end{aligned} \tag{28}$$

where the second equality holds due to the purity of $\rho^{(k_0\cdots k_N k_s)}$. The last equality follows from $\rho^{(k_0\cdots k_N)} = \mathbb{E}_{k_s}[\rho^{(k_0\cdots k_N k_s)}]$, which can be derived in the same way as Eq. 24.

The total *subjective entropy*, accounting for Wigner and all his friends as well as the environment, is given by the following after averaging over all indices:

$$\begin{aligned} S_{\text{subjective}} &= S^{(\text{Wigner})}_{\text{subjective}} + \sum_{i=1}^{N} S^{(i\text{-th})}_{\text{subjective}} + S^{(\text{environment})}_{\text{subjective}} \\ &= \mathbb{E}\left\{\left[S\left(\rho^{(k_0)}\right) - S\left(\tilde{\rho}\right)\right] + \sum_{i=1}^{N}\left[S\left(\rho^{(k_0\cdots k_{i-1}k_i)}\right) - S\left(\rho^{(k_0\cdots k_{i-1})}\right)\right] - S\left(\rho^{(k_0\cdots k_N)}\right)\right\} = -S\left(\tilde{\rho}\right) \ . \end{aligned} \tag{29}$$

Recall that before the measurement, the *objective* and *subjective entropies* are $S(\rho)$ and $-S(\rho)$, respectively (see Table III). After the measurement, the coarse-grained object state $\tilde{\rho}$ gives rise to an *objective entropy* of $S(\tilde{\rho})$, and the total *subjective entropy* is $-S(\tilde{\rho})$ as shown in Eq. 29. The *perspective-free entropy* defined in Eq. 10 is unchanged before and after the measurement. It is noted that this thought experiment uses "friend" only as a metaphor. In reality, the system partition needs not to involve sentient beings. All it requires is a general factorization of the entire Hilbert space. Therefore, the thought experiment of Wigner's many friends illustrates the conservation of the *perspective-free entropy* in a quite general setting.

VIII. SECOND LAW AND PAST HYPOTHESIS

A main result from the discussions above is the conservation of the *perspective-free entropy* S, defined as the sum of the *objective* and *subjective entropies*:

$$\Delta S = 0 \ \Leftrightarrow \ \Delta S_{\text{objective}} = -\Delta S_{\text{subjective}} \ . \tag{30}$$

This equation is valid under general system partition and POVM. For most practical applications, we split the system into three components, namely the object, the environment, and an external observer. In this case, the total *subjective entropy* may be divided into the *subjective entropy* of the environment and the *subjective entropy* of the observer. This gives a special form of Eq. 30:

$$\Delta S_{\text{objective}} + \Delta S^{(\text{observer})}_{\text{subjective}} = -\Delta S^{(\text{environment})}_{\text{subjective}} \ . \tag{31}$$

From the perspective of the observer, the change in *apparent entropy* is given by the left-hand side (LHS) of Eq. 31. This includes the *subjective entropy* of the observer, thus taking into account the fact that the observer determines the *local* state of the object by utilizing her self-locating knowledge. The right-hand side (RHS) of Eq. 31 measures the decrease in the *subjective entropy* of the environment. With the help of Eq. 31, despite decoherence and measurement appearing to follow two separate laws as illustrated in [7, 19], they are indeed two special cases of an unified law:

$$\Delta S_{\text{subjective}} \leq 0 \text{ when interacting with the object.} \quad (32)$$

Eq. 32 has a clear information-theoretic interpretation, for one always acquires self-locating knowledge by interacting with the object. If it is the observer who interacts with the object, Eq. 32 applies to the observer, decreasing the LHS of Eq. 31. As a result, the *apparent entropy* decreases after the object-observer interaction, corresponding to the process of measurement. If, on the other hand, it is the environment that interacts with the object, then Eq. 32 applies to the environment, increasing the RHS of Eq. 31. This results in an entropy increase after the object-environment interaction, or the process of decoherence.

To make a detailed analysis of these changes in *subjective entropies*, we first consider a measurement conducted by the external observer. After the measurement the observer becomes aware of the branch she inhabits. This decreases her *subjective entropy* from 0 to $-S\left(\rho^{(k)} \,\middle\|\, \sum_l p_l \rho^{(l)}\right)$ in the k-th branch, where p_k and $\rho^{(k)}$ are the self-locating probability and the *local* state of the object, respectively. This negative *subjective entropy* measures the information gain of updating the *global* state $\sum_l p_l \rho^{(l)}$ to the *local* state $\rho^{(k)}$. We can see that Eq. 32 holds for interactions between an object and an external observer.

The same logic applies when it is the environment that interacts with the object. In the process of decoherence, *subjective entropy* of the environment changes to $-S\left(\rho^{(k)} \,\middle\|\, \sum_l p_l \rho^{(l)}\right)$ in the k-th branch. In the case that the object-environment composite is in a pure state, the object state $\rho^{(k)}$, obtained via projection, is also pure. Using the purity of $\rho^{(k)}$ we get

$$\begin{aligned} S_{\text{subjective}}^{(\text{environment})} &= \mathbb{E}\left[-S\left(\rho^{(k)} \,\middle\|\, \sum_l p_l \rho^{(l)}\right)\right] \\ &= -S\left(\sum_l p_l \rho^{(l)}\right) \leq -S(\rho) \ , \end{aligned} \quad (33)$$

where ρ is the initial state of the object before decoherence, and the inequality is proved in [7]. Note that initially the *objective entropy* and the *subjective entropy* of the environment are $S(\rho)$ and $-S(\rho)$, respectively. Eq. 33 shows that the *subjective entropy* decreases on average, validating Eq. 32 for interactions between an object and its environment.

The most common version of the second law may be regarded as a special case of Eq. 32, in which an isolated system is divided into subsystems (which may be achieved by an arbitrary factorization of the Hilbert space). These subsystems interact with each other, and therefore have decreasing *subjective entropies* according to Eq. 32. This reflects the fact that they acquire self-locating knowledge about each other (recall that self-locating knowledge is completely physical and is realized by quantum entanglement). An equivalent claim is that they get more correlated with each other. To compensate the decrease in *subjective entropies*, the *objective entropy* increases. Given that we (as "ideal observers") do not access any self-locating knowledge acquired by these subsystems, we can only perceive an *apparent entropy* that equals the increasing *objective entropy.* This explains the entropy-increasing dynamics as claimed by the thermodynamic second law.

One might argue that this thermodynamic irreversibility originates from the assumption that the subsystems are initially uncorrelated. Suppose, on the other hand, these subsystems are initially highly correlated. Then they would probably get less correlated over time which reverses our normal arrow of time, for the fundamental laws of physics are time reversible. Therefore, it does seem necessary to postulate an initial cosmological condition with an extremely low entropy or correlation (the past hypothesis).

It is indeed true that the thermodynamic irreversibility is derived assuming a low-entropy (or equivalently weakly correlated) initial condition. However, this assumption is regarded as an *observer effect* rather than an *ad hoc* postulation. According to Eq. 32, every observation we make results in a decrease in the *subjective entropy* and hence the LHS of Eq. 31. This suggests that even though the *global* state can have high entropy, the entropy of the *local* state stays at a low level. In fact, the former ignores any self-locating knowledge and hence excludes any *subjective entropy*, while the latter includes a negative *subjective entropy* measuring the self-locating knowledge acquired via observation. This explains why we *actually* live in a low-entropy world, even though high-entropy states are much more likely. Since Eq. 32 explains both the entropy-increasing dynamics and the low-entropy observation, it provides a complete explanation for the past hypothesis: an extremely low-entropy initial condition given by extrapolating the current low-entropy *local* reality backwards in time.

We will explain in detail why the *local* states that we observe are weakly correlated (or equivalently why our *local* reality has low entropy). Consider the case in Sec. V in which the object and the environment are initially in a highly correlated state $\sum_i a_i |\epsilon_i\rangle|o_i\rangle$. According to Eq. 14, after the measurement, the state of the object-environment subsystem, as perceived by the observer in the "pointer basis" state $|s_k\rangle$, is instead a simply separable state: $(\sum_i a_i \langle\tilde{o}_k|o_i\rangle|\epsilon_i\rangle) \otimes |\tilde{o}_k\rangle$. This completely decorrelates the object and the environment, measured by a sufficient decrease in the *subjective entropy* of the observer such that the LHS of Eq. 31 drops to zero. As a result, the observer actually perceives a zero-entropy state in which the object and the environment are completely uncorrelated.

As discussed in Sec. VI, a more general result follows from Eq. 16, in which the observer in the "pointer basis" state $|s_k\rangle$ perceives the object-environment subsystem being in a partially decorrelated state $\sum_{i,j} a_i \langle\tilde{o}_j|o_i\rangle\langle s_k|\sigma_j\rangle|\epsilon_i\rangle|\tilde{o}_j\rangle$. We use the entanglement entropy to check if it is indeed a less correlated state than the initial state $\sum_i a_i|\epsilon_i\rangle|o_i\rangle$. By partial-tracing the two states over the environment, we find the corresponding density matrices of the object to be $\rho^{(k)}$ and ρ as given in Eq. 17. Therefore, the entanglement entropy is $S(\rho)$ before the measurement and $S(\rho^{(k)})$ after the measurement (in the k-th branch). The decreasing entanglement entropy, $\mathbb{E}\left[S\left(\rho^{(k)}\right)\right] \leq S(\rho)$ as proved in [19], demonstrates that measurement indeed leads to a less correlated state. In entropic terms, this is captured by an increasing *subjective entropy* of the environment. According to Eq. 31, the observer perceives a *local* reality with lower entropy.

As illustrated above, the second law is now fully explained by Eq. 32 instead of relying on some unexplained initial condition. We therefore claim that Eq. 32 provides a complete and unified explanation of the second law. One implication of Eq. 32 is that the thermodynamic irreversibility may be regarded as a *subjective* fact, and therefore a result of the subject-object division. This does not imply anything non-physical or anything that is free to change at will. It simply means the irreversibility is related to *subjective experience*. In fact, even though the fundamental laws of physics are completely reversible, we still *experience* irreversibility from a particular perspective.

The second law, formulated by Eq. 32, thus reflects a *subjective* or *perceptual* irreversibility. This provides an alternative solution to Loschmidt's paradox [26]: the thermodynamic arrow of time derives from the subjective arrow of time (while most literature work derives the latter from the former [27–29]). On the nature of time, its reversibility illustrated in laws of physics is therefore not contradictory to our everyday *experience* about its directionality.

We noticed a similar explanation of the thermodynamic second law proposed by Maccone [30]. It suggests that entropy-increasing and entropy-decreasing processes occur equally likely while only the former can be *recorded*. This approach does not distinguish explicitly between the *global* reality and the *local* reality. Instead it relies on the fact that recording or memorizing mechanism requires increase in correlation. It therefore concludes that the second law is a mere tautology, as the memorizing mechanism already presumes an arrow of time pointing towards increasing correlation.

In contrast, the explanation proposed in the present work distinguishes between the *global* reality and the *local* reality. It suggests that while the *global* reality has no directionality, the *local* reality is indeed directional. Therefore, one is incapable of not only *recording*, but also *experiencing* entropy-decreasing processes. In the *local* reality, entropy-increasing and entropy-decreasing processes are not equally likely. A vase breaking into pieces is way more likely than these pieces forming a brand new vase.

One implication of Eq. 32 is that measurement may decrease the *apparent entropy*. This fact, which is regarded as a variation of the second law, is not covered by Maccone's approach. As a result, Maccone's approach does not explain why we see a low-entropy world while high-entropy states are overwhelming. It is thus incapable of providing a complete solution to Loschmidt's paradox [31]. On the other hand, the present work provides a plausible solution. It suggests a low-entropy *local* state despite the high-entropy *global* state. This is due to the acquisition of self-locating knowledge, measured by a negative *subjective entropy* of the observer, which largely offsets the positive *objective entropy*.

As a final comment, we would like to compare the low-entropy initial singularity with the high-entropy black hole singularity. The entropy of the initial singularity, extrapolated from the current entropy of the universe, appears to be extremely low compared to the black hole entropy [32, 33]. As explained above, this low entropy applies to the *local* reality. Since we are *inside* the system, it accounts for the self-locating knowledge acquired via our perception.

In contrast, a black hole singularity is protected by a horizon, which separates us from the black hole interior. As a result, we are *outside* the system, incapable of acquiring self-locating knowledge about the black hole interior. In other words, the black hole degrees of freedom can be regarded as something completely uncorrelated with an external observer. From our perspective, they form an isolated quantum system which evolves unitarily, consistent with the "central dogma" [15]. Even if some part of the black hole interior is entangled with Hawking radiation, an external observer does not access the self-locating knowledge acquired by the radiation. Therefore, the *subjective entropy* of the observer is zero. The net result is a coarse-grained black hole entropy equaling one fourth of its horizon area, which saturates the Bekenstein bound [34].

The author would like to thank Cynthia Stewart, David Pearce, and Shu Lin for their very helpful advices and thorough proofread.

[1] J. Liouville, Note sur la théorie de la variation des constantes arbitraires, Journal de mathématiques pures et appliquées **3**, 342 (1838).

[2] J. W. Gibbs, *Elementary principles in statistical mechanics* (Courier Corporation, 2014).

[3] L. Boltzmann, The second law of thermodynamics, in *Theoretical physics and philosophical problems* (Springer, 1974) pp. 13–32.

[4] S. Carroll, *The big picture: on the origins of life, meaning, and the universe itself* (Penguin, 2017).

[5] P. C. Davies, Inflation and time asymmetry in the universe, Nature **301**, 398 (1983).

[6] S. M. Carroll and H. Tam, Unitary evolution and cosmological fine-tuning, arXiv preprint arXiv:1007.1417 (2010).

[7] M. Tegmark, How unitary cosmology generalizes thermodynamics and solves the inflationary entropy problem, Physical Review D **85**, 123517 (2012).

[8] H. Everett III, "Relative state" formulation of quantum mechanics, Reviews of Modern Physics **29**, 454 (1957).

[9] E. P. Wigner, Symmetries and reflections: Scientific essays, American Journal of Physics **35**, 1169 (1967).

[10] S. Saunders and D. Wallace, Branching and uncertainty, The British Journal for the Philosophy of Science **59**, 293 (2008).

[11] L. Vaidman, Probability in the many-worlds interpretation of quantum mechanics, in *Probability in physics* (Springer, 2012) pp. 299–311.

[12] S. M. Carroll and C. T. Sebens, Many worlds, the Born rule, and self-locating uncertainty, in *Quantum theory: A two-time success story* (Springer, 2014) pp. 157–169.

[13] C. T. Sebens and S. M. Carroll, Self-locating uncertainty and the origin of probability in everettian quantum mechanics, The British Journal for the Philosophy of Science **69**, 25 (2018).

[14] K. J. McQueen and L. Vaidman, In defence of the self-location uncertainty account of probability in the many-worlds interpretation, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **66**, 14 (2019).

[15] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, The entropy of Hawking radiation, arXiv preprint arXiv:2006.06872 (2020).

[16] H. Breuer and F. Petruccione, *The theory of open quantum systems* (Oxford University Press, 2002).

[17] A. Pathak, *Elements of quantum computation and quantum communication* (Taylor & Francis, 2013).

[18] W. H. Zurek, Quantum Darwinism, Nature Physics **5**, 181 (2009).

[19] H. Gharibyan and M. Tegmark, Sharpening the second law of thermodynamics with the quantum Bayes theorem, Physical Review E **90**, 032125 (2014).

[20] S. W. Kim, T. Sagawa, S. De Liberato, and M. Ueda, Quantum Szilard engine, Physical Review Letters **106**, 070401 (2011).

[21] S. Lloyd, Quantum-mechanical Maxwell's demon, Physical Review A **56**, 3374 (1997).

[22] M. A. Nielsen and I. Chuang, Quantum computation and quantum information (2002).

[23] M. B. Ruskai, Inequalities for quantum entropy: A review with conditions for equality, Journal of Mathematical Physics **43**, 4358 (2002).

[24] R. Kawai, J. M. Parrondo, and C. Van den Broeck, Dissipation: The phase-space perspective, Physical Review Letters **98**, 080602 (2007).

[25] E. P. Wigner, Remarks on the mind-body question, in *I. J. Good (ed.), The Scientist Speculates* (London: Heinemann, 1961).

[26] J. Loschmidt, Uber das warmegleichgewicht eines systems von korpern mit rucksicht auf die schwere, Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss. Classe **73**, 128 (1876).

[27] L. Mlodinow and T. A. Brun, Relation between the psychological and thermodynamic arrows of time, Physical Review E **89**, 052102 (2014).

[28] H. Price, *Time's Arrow and Archimede's Point* (Oxford University Press, 1996).

[29] H. D. Zeh, *The Physical Basis of The Direction of Time* (Springer, 2001).

[30] L. Maccone, Quantum solution to the arrow-of-time dilemma, Physical Review Letters **103**, 080401 (2009).

[31] L. Maccone, The thermodynamic arrow-of-time and quantum mechanics, Electronic Notes in Theoretical Computer Science **270**, 75 (2011).

[32] J. D. Bekenstein, Black holes and the second law, Lettere al Nuovo Cimento (1971-1985) **4**, 737 (1972).

[33] S. W. Hawking, Particle creation by black holes, Communications in Mathematical Physics **43**, 199 (1975).

[34] J. D. Bekenstein, Universal upper bound on the entropy-to-energy ratio for bounded systems, Physical Review D **23**, 287 (1981).