Redundancy of einselected information in quantum Darwinism: The irrelevance of irrelevant environment bits

Michael Zwolak and Wojciech H. Zurek

Phys. Rev. A **95**, 030101 — Published 8 March 2017

DOI: 10.1103/PhysRevA.95.030101

Separating the wheat from the chaff: Redundancy and relevant information in quantum Darwinism

Michael Zwolak[1,*] and Wojciech H. Zurek[2]

[1]*Department of Physics, Oregon State University, Corvallis, OR 97331*
[2]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545*

The objective, classical world emerges from the underlying quantum substrate via the proliferation of redundant copies of selected information into the environment, which acts as a communication channel, transmitting that information to observers. These copies are independently accessible, allowing many observers to reach consensus about the state of a quantum system via its imprints in the environment. Quantum Darwinism recognizes that the redundancy of information is thus central to the emergence of objective reality in the quantum world. However, in addition to the "quantum system of interest", there are many other systems "of no interest" in the Universe that can imprint information on the common environment. There is therefore a danger that the information of interest will be diluted with irrelevant bits, suppressing the redundancy responsible for objectivity. Remarkably, we show that mixing of the "relevant" and "irrelevant" bits of information makes little quantitative difference to the redundancy of the information of interest. Thus, we demonstrate that it does not matter whether one separates the wheat (relevant information) from the (irrelevant) chaff: The large redundancy of the relevant information survives dilution, providing evidence of the objective, effectively classical world.

Amplification – already invoked by Bohr [1] – is the central process by which the underlying quantum substrate gives rise to the objective, classical world [2, 3]. Quantum Darwinism [4, 5] formalizes this notion into the concept of redundancy: When quantum systems are decohered [6–8], they transfer select information – information about their pointer states [9] – to their environment. Many observers can then infer the state of the system indirectly by intercepting some small fragment of the environment [10, 11]. In other words, this select information is redundant, as any small fragment will do, and is thus objective: Many observers can independently deduce the pointer state of the system and reach consensus about it.

In our Universe there are many fragments of the environment that have no or nearly no information about any given quantum system of interest at any given time. Thus, in what way can one then apply the quantum Darwinist considerations? To begin addressing this question, we consider a spin model introduced in Ref. [12] with two types of spins in the environment $\mathcal{E}$, ones that acquire perfect (classical) information about the system $\mathcal{S}$ and others that acquire no information. These are the "good" $\mathcal{E}_G$ and "bad" $\mathcal{E}_B$ environments, respectively. This can be represented by a state of the form

$$\left(\frac{1}{\sqrt{2}}|0\rangle_{\mathcal{S}}\overbrace{|0\cdots0\rangle}^{\mathcal{E}_G}+\frac{1}{\sqrt{2}}|1\rangle_{\mathcal{S}}\overbrace{|1\cdots1\rangle}^{\mathcal{E}_G}\right)\overbrace{|0\cdots0\rangle}^{\mathcal{E}_B}. \qquad (1)$$

Physically, this state is generated when a set $\mathcal{E}_G$ of "good" environment components – the "wheat" – each perfectly decohere the system in the z-basis and a set $\mathcal{E}_B$ – the "chaff" – do not interact at all with the system. In what follows, the results are not limited to "diametrically opposed" good and bad environments (i.e., ones like in Eq. (1)), but rather extend to the case with partial information in both the good and bad environments, as well as mixed states, states without permutational invariance (e.g., in the environment $\mathcal{E}_G$), and alternative measurements to extract the information [13].

Pure decoherence – whether in a globally pure state or in mixed state – is the process by which the state structure in Eq. (1) arises. The Hamiltonian and initial states for pure decoherence by independent environment components [14, 15] are $\mathbf{H}=\mathbf{H}_{\mathcal{S}}+\hat{\Pi}_{\mathcal{S}}\sum_{k=1}^{\sharp\mathcal{E}}\Upsilon_k+\sum_{k=1}^{\sharp\mathcal{E}}\Omega_k$ with $\left[\hat{\Pi}_{\mathcal{S}},\mathbf{H}_{\mathcal{S}}\right]=0$ and $\rho(0)=\rho_{\mathcal{S}}(0)\otimes\left[\bigotimes_{k=1}^{\sharp\mathcal{E}}\rho_k(0)\right]$, where k specifies an environment subsystem. Under pure decoherence, no transitions occur between the pointer states $\hat{s}$ (the eigenstates of $\hat{\Pi}_{\mathcal{S}}$ [7, 9]). Up to unimportant local unitary rotations, the state, Eq. (1), develops via a pure decoherence process, e.g., one having a $g_k\sigma_{\mathcal{S}}^z\sigma_k^z$ interaction in the Hamiltonian with $g_k=1$ for $k\in\mathcal{E}_G$ and $g_k=0$ for $k\in\mathcal{E}_B$, and an initial state $(|0\rangle_{\mathcal{S}}+|1\rangle_{\mathcal{S}})\,|+\cdots+\rangle|0\cdots0\rangle/2$, where $|+\rangle$ is a σ^x eigenstate. These models – which include run-of-the-mill, everyday photon environments [2] – approximate the case where decoherence is strong compared to the natural dynamics of the system. Moreover, spin models of this type help elucidate the nature of redundancy in various settings [3, 16], which is what we will do here.

Intuitively, we know the redundancy of information in the state in Eq. (1): There are $^\sharp\mathcal{E}_G$ "good" bits, which are in a GHZ state, and thus perfectly classically correlated with the pointer observable of the system, and there are $^\sharp\mathcal{E}_B$ "bad" bits, which are in a product state, and thus not correlated at all with the system. Hence, the redundancy is just $^\sharp\mathcal{E}_G$. However, in a world where we are bombarded with good and bad bits alike, the question arises: What

*Electronic address: mpzwolak@gmail.com

is the typical fragment size that we need to intercept from the total environment, $\mathcal{E} = \mathcal{E}_G \otimes \mathcal{E}_B$, to get nearly complete information about $\mathcal{S}$? Likewise, in what way should we define redundancy, with respect to $\mathcal{E}$ or just $\mathcal{E}_G$?

To answer these questions, we examine the mutual information, which quantifies the correlations between the system and some fragment of the environment,

$$I(\mathcal{S} : \mathcal{F}) = H_{\mathcal{S}} + H_{\mathcal{F}} - H_{\mathcal{S}\mathcal{F}}. \tag{2}$$

This can be divided into classical (the Holevo quantity [17]) and quantum (the discord [18–20]) components [21],

$$I(\mathcal{S} : \mathcal{F}) = \chi(\Pi_{\mathcal{S}} : \mathcal{F}) + \mathcal{D}(\Pi_{\mathcal{S}} : \mathcal{F}), \tag{3}$$

where $\Pi_{\mathcal{S}}$ specifies a basis (or, more generally, a POVM) on $\mathcal{S}$. The Holevo quantity, $\chi(\Pi_{\mathcal{S}} : \mathcal{F})$, gives the maximum classical information available about $\Pi_{\mathcal{S}}$ in $\mathcal{F}$, while the quantum discord $\mathcal{D}(\Pi_{\mathcal{S}} : \mathcal{F})$ is what remains. The information most efficiently transmitted by the environment will be about the pointer basis, $\hat{\Pi}_{\mathcal{S}}$ [10, 21]. Thus, the Holevo quantity for that basis, $\chi\left(\hat{\Pi}_{\mathcal{S}} : \mathcal{F}\right)$, will be of interest.

To define redundant information, we need to know how much (what portion) of the "classical information of interest" is contained in a typical fragment $\mathcal{F}$ (of size $^{\sharp}\mathcal{F}_{\delta} < {}^{\sharp}\mathcal{E}$) of the total environment, $\mathcal{E}$. Thus, we seek $^{\sharp}\mathcal{F}_{\delta}$, the size of the fragments that contain all but the information deficit δ of the classical information,

$$\langle \chi\left(\hat{\Pi}_{\mathcal{S}} : \mathcal{F}\right)\rangle_{^{\sharp}\mathcal{F}_{\delta}} \simeq (1-\delta) H_{\mathcal{S}}. \tag{4}$$

Above $\langle\cdot\rangle_{^{\sharp}\mathcal{F}_{\delta}}$ is the average over fragments of size $^{\sharp}\mathcal{F}_{\delta}$ and $H_{\mathcal{S}} = H(\hat{\Pi}_{\mathcal{S}})$, i.e., the entropy of the pointer observable – the missing information about $\mathcal{S}$. Observers do not require (and usually cannot get) all the missing information. The information deficit, δ, is the amount of information that observers are prepared to forgo. The averaging can be done over the relevant environment, $\mathcal{E}_G$, or the total environment, $\mathcal{E}$. Alternatively, one can maximize the number of distinct fragments that give nearly complete information about $\hat{\Pi}_{\mathcal{S}}$, i.e., $(1-\delta) H_{\mathcal{S}}$.

Each of these approaches may yield different results. We will show that – remarkably – all these procedures give the same value for the redundancy,

$$R_{\delta} = \frac{^{\sharp}\mathcal{E}}{^{\sharp}\mathcal{F}_{\delta}}, \tag{5}$$

up to an insignificant scaling factor, where $^{\sharp}\mathcal{F}_{\delta}$ is taken from Eq. (4) or from the maximization procedure.

The Holevo quantity will approach $H_{\mathcal{S}}$ according to [2, 3]

$$\chi\left(\hat{\Pi}_{\mathcal{S}} : \mathcal{F}\right) \sim H_{\mathcal{S}} - H(P_e), \tag{6}$$

where P_e, a function of $\mathcal{F}$, is the error probability for distinguishing the conditional states of the fragment. The latter are given by $\langle \hat{s}|\rho_{\mathcal{S}\mathcal{F}}|\hat{s}\rangle / p_{\hat{s}}$, where $|\hat{s}\rangle$ is a pointer state, $p_{\hat{s}}$ is the probability that state occurs, and $\rho_{\mathcal{S}\mathcal{F}}$ is the reduced state of the system and fragment. In the case of the state in Eq. (1), the conditional states are just $|0\cdots0\rangle_{\mathcal{E}_G}|0\cdots0\rangle_{\mathcal{E}_B}$ and $|1\cdots1\rangle_{\mathcal{E}_G}|0\cdots0\rangle_{\mathcal{E}_B}$.

The asymptotic behavior of the error probability is given by

$$P_e \sim \exp\left[-\bar{\xi}_{QCB}{}^{\sharp}\mathcal{F}\right], \tag{7}$$

where the exponent

$$\bar{\xi}_{QCB} = -\ln\langle \mathrm{tr}\left[\rho_{k|1}^{c}\,\rho_{k|2}^{1-c}\right]\rangle_{k\in\mathcal{E}} \tag{8}$$

is the "typical" Chernoff information [2], which generalizes the quantum Chernoff bound (QCB) [22–25] to sources of quantum states that are not independent and identically distributed (i.i.d.). The error probability also depends on the $p_{\hat{s}}$, but only in a prefactor to the exponential, and thus it does not play a role as $^{\sharp}\mathcal{F}$ becomes large. The optimal value of c ($0 \le c \le 1$) is the one that maximizes $\bar{\xi}_{QCB}$. The latter is not an easy task for non-i.i.d states. However, for spin systems undergoing pure decoherence (of which the state in Eq. (1) is an example) the value of c is $1/2$ [3] (the value of c can also be found for certain classes of photon environments [2]).

This maps the understanding of classical information communicated by $\mathcal{F}$ (and $\mathcal{E}$) into a problem of understanding the distinguishability of conditional states, $\rho_{k|\hat{s}}$ on individual environment components k (i.e., for the case here, individual environment spins). Equations (4), (6), and (7) allow us to estimate the redundancy [2, 3], Eq. (5), as

$$R_{\delta} \simeq {}^{\sharp}\mathcal{E}\frac{\bar{\xi}_{QCB}}{\ln 1/\delta}. \tag{9}$$

Indeed, this shows that macroscopic redundancy is unavoidable for pure decoherence – except for states of measure zero (i.e., completely mixed initial environment states or ones that commute with the Hamiltonian), redundancy is always present [2, 3, 14, 15]. The worst scenario for the dilution of information is when $^{\sharp}\mathcal{E}_G$ is small and fixed, while $^{\sharp}\mathcal{E}_B$ is taken to be larger and larger. Taking $^{\sharp}\mathcal{E}_B \gg {}^{\sharp}\mathcal{E}_G$, $^{\sharp}\mathcal{F}$, Eq. (8) is simple,

$$\begin{aligned}\bar{\xi}_{QCB} &= -\ln\left[\frac{^{\sharp}\mathcal{E}_B\cdot 1 + {}^{\sharp}\mathcal{E}_G\cdot 0}{^{\sharp}\mathcal{E}}\right]\\ &= -\ln\left[\frac{^{\sharp}\mathcal{E} - {}^{\sharp}\mathcal{E}_G}{^{\sharp}\mathcal{E}}\right] \simeq \frac{^{\sharp}\mathcal{E}_G}{^{\sharp}\mathcal{E}}.\end{aligned} \tag{10}$$

Thus, the redundancy is

$$R_{\delta} \simeq \frac{^{\sharp}\mathcal{E}_G}{\ln 1/\delta}. \tag{11}$$

We see that the calculation requires that $^{\sharp}\mathcal{E}_G \gtrsim \ln(1/\delta)$, as redundancy can not be less than one. If $^{\sharp}\mathcal{E}_G < \ln(1/\delta)$, it means that the scenario is outside of the realm of validity of the QCB calculation, as an observer needs essentially the whole environment to approach $(1-\delta)\,H_{\mathcal{S}}$ bits of information about $\mathcal{S}$, if they can acquire that amount of information at all.

We note that one can also *exactly* solve for the average Holevo quantity or mutual information, which yields the same result: Given a fragment $\mathcal{F}$, $\chi\left(\hat{\Pi}_{\mathcal{S}}:\mathcal{F}\right)$ will be one if a good bit is intercepted and zero otherwise (similarly for the quantum mutual information, $I\left(\mathcal{S}:\mathcal{F}\right)$, unless all the good bits are intercepted, but this happens with negligible probability). The probability that the observer will intercept $\sharp\mathcal{F}$ bad bits – and have zero information about the system – is

$$P_B=\frac{\sharp\mathcal{E}_B}{\sharp\mathcal{E}}\cdot\frac{\sharp\mathcal{E}_B-1}{\sharp\mathcal{E}-1}\cdots\frac{\sharp\mathcal{E}_B-(\sharp\mathcal{F}-1)}{\sharp\mathcal{E}-(\sharp\mathcal{F}-1)},\qquad(12)$$

with $\sharp\mathcal{E}=\sharp\mathcal{E}_G+\sharp\mathcal{E}_B$. The probability to intercept at least one good bit is $1-P_B$, giving

$$\left\langle\chi\left(\hat{\Pi}_{\mathcal{S}}:\mathcal{F}\right)\right\rangle_{\sharp\mathcal{F}}=H_{\mathcal{S}}\left(1-P_B\right)=1-\frac{\sharp\mathcal{E}_B!\left(\sharp\mathcal{E}-\sharp\mathcal{F}\right)!}{(\sharp\mathcal{E}_B-\sharp\mathcal{F})!\sharp\mathcal{E}!},\qquad(13)$$

with $H_{\mathcal{S}}=1$. Redundancy requires that

$$\delta\simeq\frac{\sharp\mathcal{E}_B!\left(\sharp\mathcal{E}-\sharp\mathcal{F}_\delta\right)!}{(\sharp\mathcal{E}_B-\sharp\mathcal{F}_\delta)!\sharp\mathcal{E}!}\qquad(14)$$

and we can use Stirling's approximation to get the limiting forms of this expression, $\ln\delta\simeq-\sharp\mathcal{F}_\delta\sharp\mathcal{E}_G/\sharp\mathcal{E}_B$. This yields the redundancy

$$R_\delta=\frac{\sharp\mathcal{E}}{\sharp\mathcal{F}_\delta}\simeq\frac{\sharp\mathcal{E}_B+\sharp\mathcal{E}_G}{\frac{\sharp\mathcal{E}_B}{\sharp\mathcal{E}_G}\ln 1/\delta}\simeq\frac{\sharp\mathcal{E}_G}{\ln 1/\delta},\qquad(15)$$

in agreement with the QCB result. While this exact calculation is simple, the QCB calculation is even simpler still and is easily extended to many other cases (see, e.g., Ref. [3]).

If instead of the usual definition of redundancy (Eqs. (4) and (5)), one defined redundancy as the maximum number of disjoint fragments for which $\chi\left(\hat{\Pi}_{\mathcal{S}}:\mathcal{F}\right)\simeq(1-\delta)\,H_{\mathcal{S}}$, the result would be $R_\delta=\sharp\mathcal{E}_G$. For this "perfect" good-bad state, Eq. (1), this is true for any $\delta\neq1$ [27]. By the same token, if one averaged Eq. (4) only over the environment $\mathcal{E}_G$, then $\sharp\mathcal{F}_\delta=1$ and $R_\delta=\sharp\mathcal{E}_G$: Only one spin from $\mathcal{E}_G$ is necessary to acquire the requisite information (we note for the perfect GHZ state of the good bits with the system, the QCB calculation is not valid when limiting to the relevant environment, as $\bar{\xi}_{QCB}$ diverges, reflecting that $R_\delta=\sharp\mathcal{E}_G$ for any δ). Indeed, these two latter approaches agree with our intuition about the state in Eq. (1), as there are just $\sharp\mathcal{E}_G$ copies of the pointer state information.

These definitions lead to $R_\delta=\sharp\mathcal{E}_G/\ln(1/\delta)$ and $R_\delta=\sharp\mathcal{E}_G$ and are thus not equivalent. However, they only differ by a factor of $\ln 1/\delta$ (and this result is unaffected by the choice of p_0, p_1, the probabilities for $|0\rangle_{\mathcal{S}}$ and $|1\rangle_{\mathcal{S}}$ in the initial state of the system) [28]. For any reasonable δ, the two definitions are practically the same [29]. The reason for the correspondence between the definitions is that – when an observer intercepts both relevant and irrelevant bits – the probability not to get any good bit drops exponentially with the size of the fragment $\sharp\mathcal{F}$. This only weakly affects the ability of the observer to capture a good bit and deduce the state of the system. This is remarkable, as it says that when states of the form in Eq. (1) arise we need not worry about distinguishing between parts of the larger environment – parts that interact with the system and parts that either do not or only weakly interact – for quantifying the redundancy of information. In other words, we need not worry about separating the wheat from the chaff.

This example can be extended to the case where the good and bad spins are not diametrically opposed, i.e., not perfectly good or bad. For instance, one can consider $\sharp\mathcal{E}_G$ good spins that contribute $|\gamma_G|^2$ to the decoherence factor and $\sharp\mathcal{E}_B$ bad spins that contribute $|\gamma_B|^2$. In this case, the QCB gives immediately

$$R_\delta\simeq\frac{\sharp\mathcal{E}\ln\left[\frac{\sharp\mathcal{E}_B}{\sharp\mathcal{E}}\left|\gamma_B\right|^2+\frac{\sharp\mathcal{E}_G}{\sharp\mathcal{E}}\left|\gamma_G\right|^2\right]}{\ln\delta},\qquad(16)$$

where we have made use of the relationship between decoherence and information in pure states, $\mathrm{tr}\rho_{k|1}^{c}\,\rho_{k|2}^{1-c}=|\gamma_k|^2$ [3].

As with the "perfect" good-bad state, this calculation can be done in an alternative manner. To find the averaged Holevo quantity, one makes use of the equality $\chi\left(\hat{\Pi}_{\mathcal{S}}:\mathcal{F}\right)=H\left(\frac{1+|\gamma_G|^{\sharp\mathcal{F}-\sharp\mathcal{F}_B}|\gamma_B|^{\sharp\mathcal{F}_B}}{2}\right)$ for pure states or its more general form for $p_0\neq p_1$ [15]. Expanding the binary entropy $H\left(x\right)$ for x near $1/2$ shows that one just needs to find the average decoherence factor. The latter for fragments of size $\sharp\mathcal{F}$ is given by the sum of $\sharp\mathcal{F}_B$ from 0 to $\sharp\mathcal{F}$ of $|\gamma_G|^{\sharp\mathcal{F}-\sharp\mathcal{F}_B}|\gamma_B|^{\sharp\mathcal{F}_B}$ times the probability

$$\frac{\sharp\mathcal{F}!}{\sharp\mathcal{F}_B!\left(\sharp\mathcal{F}-\sharp\mathcal{F}_B\right)}\frac{\sharp\mathcal{E}_B!}{(\sharp\mathcal{E}_B-\sharp\mathcal{F}_B)!}\frac{\sharp\mathcal{E}_G!}{(\sharp\mathcal{E}_G-\sharp\mathcal{F}_G)!}\frac{\left(\sharp\mathcal{E}-\sharp\mathcal{F}\right)!}{\sharp\mathcal{E}!}\qquad(17)$$

of intercepting $\sharp\mathcal{F}_B$ bad spins and $\sharp\mathcal{F}_G$ good spins in the fragment. Stirling's approximation can be used on the three last factors in the probability and the sum performed. This yields the QCB result, Eq. (16), at much greater expense.

For simplicity, we now assume that the bad spins give $|\gamma_B|^2=1$, i.e., a worst case where those bits never interacted with the system (and thus have no information). Expanding the QCB result, Eq. (16), one finds,

$$R_\delta\simeq\frac{\sharp\mathcal{E}_G\left(1-\left|\gamma_G\right|^2\right)}{\ln 1/\delta}.\qquad(18)$$

Maximizing the number of fragments that give $(1-\delta)\,H_{\mathcal{S}}$ bits of information is asymptotically equivalent to a QCB calculation (for permutationally invariant states) that simply ignores the $\sharp\mathcal{E}_B$ bad spins:

$$R_\delta \simeq \frac{\sharp\mathcal{E}_G \ln|\gamma_G|^2}{\ln\delta}. \tag{19}$$

These results, Eqs. (18) and (19), for the redundancy look different and it is not readily apparent that the same conclusions hold as in the "perfect" good-bad model, as the difference depends on the contribution to the decoherence factor from the good spins. The ratio between the two results, Eqs. (18) and (19), is

$$\frac{\left(1-|\gamma_G|^2\right)}{\ln 1/|\gamma_G|^2}. \tag{20}$$

As we will now show, the smallest this ratio can be is $\sim 1/\ln 1/\delta$, just like the perfect good-bad model above.

One first has to note that Eq. (20) is monotonically increasing to 1 as $|\gamma_G|^2$ increases to 1 (giving a redundancy of zero and making the definitions of redundancy the same in this limit). This can be proven by taking its derivative and applying the inequality in footnote [25] of Ref. [2] to show the derivate is always positive. The smallest value of Eq. (20) can then be found by taking the smallest value of $|\gamma_G|^2$ allowed by a consistent application of the QCB. There are two calculations of redundancy – one with good and bad spins, and one with only good spins. The former, Eq. (18), allows for all $|\gamma_G|^2$ (indeed, this just recovers Eq. (11)). The latter, Eq. (19), however, requires $\delta \le |\gamma_G|^2$ (i.e., Eq. (19) can not be larger than $\sharp\mathcal{E}_G$ no matter how small $|\gamma_G|^2$ is, or, in other words, $\sharp\mathcal{F}_\delta$ can not be less than 1). When

$$\frac{\ln|\gamma_G|^2}{\ln\delta} = 1, \tag{21}$$

or $|\gamma_G|^2 = \delta$, each good spin holds a sufficient record to immediately put the information to within δ of the plateau upon receipt of one spin [30]. When $|\gamma_G|^2 < \delta$, the redundancy no longer depends on $|\gamma_G|^2$, it is just $R_\delta = \sharp\mathcal{E}_G$ [31]. Thus, the QCB result, Eq. (19), is valid only when $|\gamma_G|^2 \ge \delta$. When this is the case, the ratio of the two computed redundancies is Eq. (20). This ratio is minimal when $|\gamma_G|^2 = \delta$, giving

$$\frac{(1-\delta)}{\ln 1/\delta}. \tag{22}$$

and is also unaffected by the choice of p_0 (p_1). Since we want δ to be small, this ratio is essentially $1/\ln 1/\delta$. Thus, just as with the "perfect" good-bad model, the two definitions differ only by an insignificant factor.

In both cases, we see that – when examining symmetric (permutationally invariant) "good" environments, $\mathcal{E}_G$ – there is no difference between redundancy defined as the maximization and redundancy defined with the averaging in Eq. (4) over only the good environment. To extend this calculation to mixed and/or non-permutationally invariant states is a straightforward matter. One only has to note the Chernoff Information is no longer directly related to the decoherence factor, $|\gamma_G|^2$, but rather takes on a different form (compare Eq. (16) and Eq. (19) in Ref. [3]), but otherwise the calculation is formally identical. Thus, the correspondence between definitions of redundancy – i.e., a difference of at most $\sim \ln 1/\delta$ – is a general feature of pure decoherence.

This example not only shows the ease of computation using the QCB, it also helps us understand the definition of redundancy itself. If redundancy was defined by maximizing the number of copies of information, rather than taking an average over all fragments of a given size, then one would obtain a different value. However, both definitions give $R_\delta \propto \sharp\mathcal{E}_G$, where the proportionality is different only by a factor of $\ln(1/\delta)$ – i.e., a factor that is only weakly dependent on δ. This is the case even when taking an essentially arbitrarily large number of bad bits in the environment. The reason for such a close correspondence between definitions is that, as the observer intercepts a larger and larger fragment, the probability of not receiving a good bit decreases exponentially with fragment size. The observer is thus likely to always to receive a good bit. The definition of redundancy in Eqs. (4) and (5) applied to the total environment $\mathcal{E}$, therefore, gives reasonable estimates – lower than the maximal redundancy but different only by an insignificant factor – for the number of records proliferated into the environment. Thus, the emergence of the classical, objective reality is unavoidable [2, 10, 26] and there is no need to separate the wheat from the chaff to perceive objective states of the systems of interest.

Acknowledgments

We would like to thank Marek Rams for helpful discussions and the Center for Integrated Quantum Science and Technology (IQST) and the University of Ulm, where part of this work was carried out. This research was supported in part by the U.S. Department of Energy through the LANL/LDRD Program and by the Foundational Questions Institute Grant No. 2015-144057 on "Physics of What Happens."

[1] N. Bohr, *Atomic Physics and Human Knowledge* (Wiley, New York, 1958) p. 83.

[2] M. Zwolak, C. J. Riedel, and W. H. Zurek, Phys. Rev.

Lett. **112**, 140406 (2014).
[3] M. Zwolak, C. J. Riedel, and W. H. Zurek, Sci. Rep. **6**, 25277 (2016).
[4] W. H. Zurek, Nat. Phys. **5**, 181 (2009).
[5] W. H. Zurek, Phys. Today **67**, 44 (2014).
[6] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer-Verlag, Berlin, 2003).
[7] W. H. Zurek, Rev. Mod. Phys. **75**, 715 (2003).
[8] M. Schlosshauer, *Decoherence and the Quantum-to-Classical Transition* (Springer-Verlag, Berlin, 2008).
[9] W. H. Zurek, Phys. Rev. D **24**, 1516 (1981).
[10] H. Ollivier, D. Poulin, and W. H. Zurek, Phys. Rev. Lett. **93**, 220401 (2004).
[11] R. Blume-Kohout and W. H. Zurek, Phys. Rev. A **73**, 062310 (2006).
[12] R. Blume-Kohout and W. H. Zurek, Found. Phys. **35**, 1857 (2005).
[13] Kincaid et al., to be published (2017).
[14] M. Zwolak, H. T. Quan, and W. H. Zurek, Phys. Rev. Lett. **103**, 110402 (2009).
[15] M. Zwolak, H. T. Quan, and W. H. Zurek, Phys. Rev. A **81**, 062110 (2010).
[16] C. J. Riedel, W. H. Zurek, and M. Zwolak, New J. Phys. **14**, 083010 (2012).
[17] A. S. Holevo, Probl. Peredachi Inf. **9**, 3 (1973).
[18] W. H. Zurek, Ann. Phys. (Leipzig) **9**, 855 (2000).
[19] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001).
[20] L. Henderson and V. Vedral, J. Phys. A: Math. Gen. **34**, 6899 (2001).
[21] M. Zwolak and W. H. Zurek, Sci. Rep. **3**, 1729 (2013).
[22] K. M. R. Audenaert, J. Calsamiglia, R. Munoz-Tapia, E. Bagan, L. Masanes, A. Acin, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007).
[23] K. Audenaert, M. Nussbaum, A. Szkoła, and F. Verstraete, Commun. Math. Phys. **279**, 251 (2008).
[24] M. Nussbaum and A. Szkoła, Ann. Stat. **37**, 1040 (2009).
[25] K. Li, Ann. Stat. **44**, 1661 (2016).
[26] F. G. S. L. Brandao, M. Piani, and P. Horodecki, Nat. Commun. **6**, 7908 (2015).
[27] When $\delta = 1$, $R_\delta = {}^\sharp\mathcal{E}$, as every component of the environment contains at least *zero* information.
[28] We note that Eq. (9) is a lower bound to the redundancy, but it was proven in Ref. [2] that it is within a factor of 2 of the exact result. For all cases where exact results can be found, such as the case here, Eq. (9) is in agreement and is thus not an estimate. However, since the redundancy computed from the maximization procedure is larger than Eq. (9), this means that we are finding a worst case difference. Similarly, we expanded the logarithm in Eq. (10). This approximation decreases the value of the redundancy. Without it, the difference between the definitions of redundancy would be closer.
[29] For $\delta = \exp(-X)$, the definitions differ only by a factor of X. Thus, for a very small δ, e.g., $\delta = \exp(-10)$ or $\exp(-20)$, the definitions differ by only a factor of, e.g., 10 or 20. In the opposite regime, when δ approaches, but is not equal to, 1 (i.e., requiring that the records hold very little information), the two definitions become exactly equivalent. This is a consequence of the fact that we must have $\bar{\xi}_{QCB}/\ln(1/\delta) < 1$ or else the QCB calculation cannot be consistently applied, as we will momentarily discuss. This means that when δ approaches 1, the redundancy plateaus at a maximum value – the size of the good environment, ${}^\sharp\mathcal{E}_G$, in this case. When $\delta = 1$, we are requiring that every fragment have no information, which is an uninteresting case (and the correspondence between definitions meaningless).
[30] This ignores finite size effects, which are nevertheless order 1.
[31] In this regime, $|\gamma_G|^2 < \delta$, one has to perform computations in an alternative way, but the conclusion is the same: The redundancy using the maximum is $R_\delta = {}^\sharp\mathcal{E}_G$ and the redundancy from the other definition does not change (the presence of $\mathcal{E}_B$ "regularizes" the QCB, so it does not matter if $|\gamma_G|^2 < \delta$). Thus, their minimum ratio is still Eq. (22).