# A Simple Example of "Quantum Darwinism": Redundant Information Storage in Many-Spin Environments

**Robin Blume-Kohout**[1] **and Wojciech H. Zurek**[2]

*As quantum information science approaches the goal of constructing quantum computers, understanding loss of information through decoherence becomes increasingly important. The information about a system that can be obtained from its environment can facilitate quantum control and error correction. Moreover, observers gain most of their information indirectly, by monitoring (primarily photon) environments of the "objects of interest." Exactly* how *this information is inscribed in the environment is essential for the emergence of "the classical" from the quantum substrate. In this paper, we examine how many-qubit (or many-spin) environments can store information about a single system. The information lost to the environment can be stored redundantly, or it can be encoded in entangled modes of the environment. We go on to show that randomly chosen states of the environment almost always encode the information so that an observer must capture a majority of the environment to deduce the system's state. Conversely, in the states produced by a typical decoherence process, information about a particular observable of the system is stored redundantly. This selective proliferation of "the fittest information" (known as Quantum Darwinism) plays a key role in choosing the preferred, effectively classical observables of macroscopic systems. The developing appreciation that the environment functions not just as a garbage dump, but as a communication channel, is extending our understanding of the environment's role in the quantum-classical transition beyond the traditional paradigm of decoherence.*

**KEY WORDS:** Decoherence; quantum Darwinism; quantum information; quantum-classical transition.

## 1. INTRODUCTION

In recent years, the field of *quantum information science* has grown tremendously. Quantum cryptography is rapidly becoming the domain of

---

[1] Inst. for Quantum information, Cal Tech;   e-mail: robin@blumekohout.com
[2] Theoretical Division, Los Alamos National Laboratory; e-mail:whz@lanl.gov

engineers more than of physicists; quantum optics and nanotechnology are making quantum systems more and more directly accessible; and quantum computers are seen not as a pipe dream, but as the next century's inevitable technology.

In this context, many scientific questions of fundamental interest have donned new hats as pressing technological issues. The transition from quantum to classical, a subject that used to be the field of philosophers and quantum foundationalists, is suddenly a critical area of research—where is it, and how does it happen? If we are to design quantum hardware, then one of the primary failure modes is precisely its transition to classical behavior. The process by which this transition occurs is decoherence, and decoherence is accordingly more interesting than ever today.

That quantum environments interact with quantum systems and, by doing so, decohere them, is by now widely appreciated. Until recently, models of decoherence usually (from a conceptual standpoint) stopped at that point; the challenge was to identify (a) what the environment was doing to the system, and (b) what was the resulting evolution of $\rho_S$, the state of the system. The environment obviously gained information about the system—effectively measuring it—but what happened to that information once it had been sucked away was largely irrelevant. The standard ansatz of decoherence analysis reflects this: from the state $|\Psi\rangle$ of the entire universe, one obtains the system's state $\rho_S$ by tracing out the environment $\mathcal{E}$. Once $\mathcal{E}$ is traced over, it vanishes from the scene.

Recently, the environment has started to become interesting again.[1–5] The most obvious feature that distinguishes the classical realm from the quantum is *objectivity*. Unknown classical states can be found out without being disturbed, so they are said to "exist objectively." This is, of course, not the case for quantum states. Thus, as it has often been emphasized by Asher Peres,[6,7] the nature of a quantum state is a much more elusive thing. In quantum systems (at least isolated ones) such states cannot be regarded as existing objectively. As was pointed out some time ago,[8] however, observers usually obtain information about a system by measuring its environment. Therefore, the manner in which the environment stores and transmits information is of great interest. In a sense, the environment is not just a wastebasket (where quantum information gets thrown away), but rather a communication channel through which observers find out about their universe. In this "*environment as a witness*" view of the transition from quantum to classical, objectivity is a consequence of the redundancy[2–5] with which information is inscribed in $\mathcal{E}$. The environment has also become interesting because a greater understanding of the entire decoherence process can facilitate the protection of fragile quantum information. Quantum control, for instance, is often based on "second

hand" information obtained from the environment. On the other hand, it's much easier to recover all of the lost quantum information when a single copy is stored in a local environment than when trillions of copies have been recorded throughout the universe.

In this paper, we skip almost completely over the traditional approaches to decoherence, and examine instead some very general properties of quantum states. We ignore the processes that produce states in which the environment has information about the system, analyzing instead the way in which that information is recorded in the environment. Our "toy" universe consists of a collection of qubits (or spin-$\frac{1}{2}$ particles), one of which is the system ($\mathcal{S}$). We begin by asking a very simple question: when the whole environment has got some information ($\mathcal{I}$) about the system, how much of that $\mathcal{I}$ can be found out by measuring a single qubit, or two qubits, or $m$ qubits of the environment? We ask this question about arbitrary states of the universe, and then we ask it about a particular sub-ensemble of states.

We conclude that our experience of the "real classical universe" is consistent with states that record certain information *redundantly*, at the expense of complementary information. Such states appear naturally in the process of decoherence. Redundantly recorded observables of the system are in effect objective; they can be independently deduced by many observers from the state of $\mathcal{E}$ that was responsible for decoherence.[2–5] Finally, we examine the concept of redundant information storage directly, and relate this quantitative measure of redundancy to our more qualitative tools for analyzing joint system–environment states.

## 2. LAYING OUT THE PLAYING FIELD

In any discussion of decoherence, we have to start by defining a system ($\mathcal{S}$) and an environment ($\mathcal{E}$). In order to go further and examine the redundancy of information storage, we also need to partition $\mathcal{E}$ into sub-environments $\mathcal{E}_1 \cdots \mathcal{E}_N$. For this entire discussion, then, our universe will be an $(N+1)$-qubit Hilbert space $\mathcal{H} = \mathcal{H}_\mathcal{S} \otimes \bigotimes_{i=1}^{N} \mathcal{H}_{\mathcal{E}_i}$, where both $\mathcal{H}_\mathcal{S}$ and all the $\mathcal{H}_{\mathcal{E}_i}$ are single qubits (or spin-$\frac{1}{2}$'s).

Our paradigm for understanding decoherence rests on the idea that the environment decoheres the system by measuring it, and that in measuring the system the subenvironments come to have information about the system. Of the several quantities we could choose to measure *how much* information a subenvironment $\mathcal{E}$ has about the system, we choose *quantum mutual information* (QMI), which is the amount of entropy produced by destroying correlations between $\mathcal{S}$ and $\mathcal{E}$:

$$\mathcal{I}_{\mathcal{S}:\mathcal{E}} \equiv H(\mathcal{S}) + H(\mathcal{E}) - H(\mathcal{S}\mathcal{E}). \tag{1}$$

While other appealing measures of information exist, [4] the QMI is mathematically elegant and generally easy to compute. We note here a few properties of $\mathcal{I}$ for future reference.

- $\mathcal{I}_{\mathcal{S}:\mathcal{E}} = 0$ if and only if $\rho_{\mathcal{S}\mathcal{E}} = \rho_{\mathcal{S}} \otimes \rho_{\mathcal{E}}$ (i.e., the sub-systems are uncorrelated).

- When $\rho_{\mathcal{S}\mathcal{E}}$ is a pure state, $\mathcal{I}_{\mathcal{S}:\mathcal{E}} = 2H(\mathcal{S})$, since $H(\mathcal{E}) = H(\mathcal{S})$ and $H(\mathcal{S}\mathcal{E}) = 0$.

- When $\rho_{\mathcal{S}\mathcal{E}}$ is a pure state, and $\mathcal{E}$ is partitioned into $\mathcal{E}_1$ and $\mathcal{E}_2$, $\mathcal{I}_{\mathcal{S}:\mathcal{E}_1} + \mathcal{I}_{\mathcal{S}:\mathcal{E}_2} = \mathcal{I}_{\mathcal{S}:\mathcal{E}}$. This follows simply from expanding $\mathcal{I}$ using Eq. (1), and the symmetry of bipartite entanglement for pure states. Note that this formula is *not* true for mixed $\rho_{\mathcal{S}\mathcal{E}}$, or for divisions of $\mathcal{E}$ into more than two fragments!

- For pure states of $\mathcal{S}\mathcal{E}$, $\mathcal{I}_{\mathcal{S}:\mathcal{E}}$ is also an entanglement measure. This connects our work with an existing body of work on multipartite entanglement, particularly in randomly or dynamically produced states, (see, e.g., Refs. 9 and 10), and on the "entangling power" of dynamical maps (see e.g., Refs. 11, 12 and Refs. 24–35 in Ref. 11). In this paper, however, we intentionally avoid the complicated field of entanglement by treating mutual information only as a measure of correlation.

Within this framework, the overall question we examine is: *How do individual subenvironments' information (about $\mathcal{S}$) compare to the whole environment's information?* We'll assume that no particular subenvironment is "special"; rather, we'll examine the information possessed by a typical subenvironment or collection thereof. It should be noted that this is *not* the same as assuming that all subenvironments are identical; we allow them to have more or less information about the system, but we obtain a statistical "typical subenvironment" by averaging. This assumption reduces our overall question to a more definite one: *How does the mutual information $\mathcal{I}$ between the system and a typical collection of m subenvironments $\mathcal{E}_{\{m\}}$ depend on m?*

Our primary tool for answering this question is the *partial information plot* or PIP (see Fig. 1), which plots $\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}}$ against $m$. Since there are many fragments $\mathcal{E}_{\{m\}}$ that contain $m$ individual environments, the partial information plot is not a curve, but a scatter-plot or histogram. To obtain a smooth curve, we average over all $\mathcal{E}_{\{m\}}$ to obtain $\overline{\mathcal{I}}(m)$. In later sections, we'll discuss the advantages and disadvantages of this averaging scheme (as opposed to, e.g., taking the average of $m$ over all $\mathcal{E}_{\{m\}}$ that
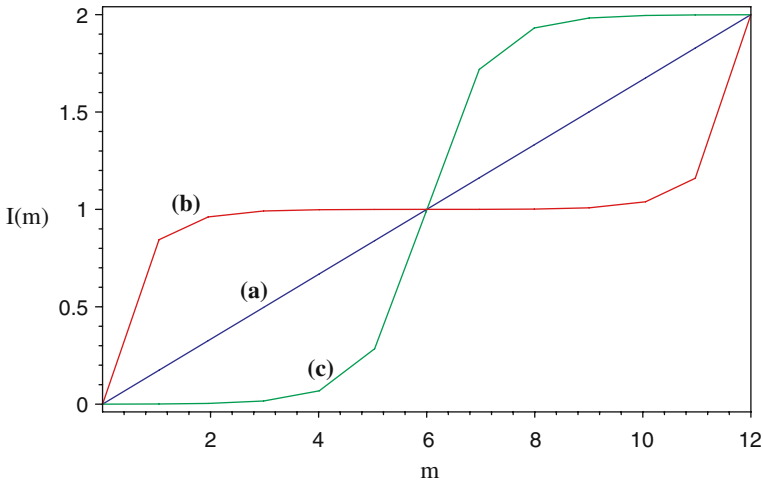
Fig. 1. The three basic profiles for partial information plots ($\mathcal{I}$ vs. $m$). Curve (a) shows the behavior of *independent* subenvironments; curve (b) shows *redundantly* stored information; and curve (c) illustrates the case where information is *encoded* in multiple subenvironments.

provide information $\mathcal{I}$). Before examining specific PIP's, we prove a simple but important lemma.

**Lemma 1.** When $\rho_{S\mathcal{E}}$ is a pure state, $\overline{\mathcal{I}}(m)$ is antisymmetric around $m = (N/2)$.

*Proof.* As noted previously, when $\mathcal{E}$ is partioned into $\mathcal{E}_{\{m\}}$ and $\mathcal{E}_{\{N-m\}}$, then $\mathcal{I}_{S:\mathcal{E}_{\{m\}}} + \mathcal{I}_{S:\mathcal{E}_{\{N-m\}}} = \mathcal{I}_{S:\mathcal{E}}$. We obtain $\overline{\mathcal{I}}(m)$ by averaging $\mathcal{I}_{S:\mathcal{E}_{\{m\}}}$ over all fragments of size $m$, and for each $\mathcal{E}_{\{m\}}$ there exists a unique complement $\mathcal{E}_{\{N-m\}}$. Thus

$$\overline{\mathcal{I}}(m) + \overline{\mathcal{I}}(N-m) = \binom{N}{m}^{-1} \sum_{\{m\}} \mathcal{I}_{S:\mathcal{E}_{\{m\}}}$$

$$+ \binom{N}{N-m}^{-1} \sum_{\{N-m\}} \mathcal{I}_{S:\mathcal{E}_{\{N-m\}}}$$

$$= \binom{N}{m} \sum_{\{m\}} \left( \mathcal{I}_{S:\mathcal{E}_{\{m\}}} + \mathcal{I}_{S:\mathcal{E}_{\{N-m\}}} \right)$$

$$= \binom{N}{m} \sum_{\{m\}} \mathcal{I}_{S:\mathcal{E}}$$

$$= \mathcal{I}_{S:\mathcal{E}}.$$

Since $\overline{\mathcal{I}}(m) + \overline{\mathcal{I}}(N - m) = \mathcal{I}_{\mathcal{S}:\mathcal{E}}$, $\overline{\mathcal{I}}(m)$ must be antisymmetric around $m = (N/2)$; further, $\overline{\mathcal{I}}\left(\frac{N}{2}\right) = (1/2)\mathcal{I}_{\mathcal{S}:\mathcal{E}} = H(\mathcal{S})$.                    □

From Lemma 1 and the obvious condition that $\overline{\mathcal{I}}(m)$ must be nondecreasing, the PIP can take three basic shapes (Fig. 1). The simplest profile (a) is one where $I(m) \propto m$–that is, each subenvironment provides unique and independent information, so that the total information gained is simply proportional to the number of subenvironment examined. Alternatively (b), information may be *redundantly* stored; $\overline{\mathcal{I}}(m)$ increases rapidly at first, then plateaus at $\overline{\mathcal{I}} \sim H(\mathcal{S})$. Finally (c), information about the system may be *encoded*, so that $\overline{\mathcal{I}}(m)$ remains close to 0, then increases rapidly around $m \sim (N/2)$. As we shall see, each of these cases is realizeable and physically relevant.

## 3. INFORMATION STORAGE FOR RANDOMLY CHOSEN STATES

In order to have a baseline to compare other ensembles of states against, we start by examining the typical PIP for the ensemble of *all* states (the uniform ensemble). To perform this average requires a measure; we use the unique invariant measure induced by the Haar measure over the unitary group $U(2^{N+1})$. This simplifies the process, since averaging over all possible fragments $\mathcal{E}_{\{m\}}$ is neatly subsumed in the average over all states.

In fact, while the calculation can be done numerically, there exists an analytic solution. Page[13] conjectured a formula for the mean entropy (averaged over Haar measure) of a subsystem, which was later proven by Sen [14] and others.[15,16] Page's formula for the mean entropy $\overline{H}(m, n)$ of an $m$-dimensional subsystem of an $mn$-dimensional system (where $m \leqslant n$) is

$$\overline{H}(m, n) = \sum_{k=n+1}^{mn} \frac{1}{k} - \frac{m - 1}{2n}. \tag{2}$$

We use this formula to calculate the mean entropy $\overline{H}_{k;N}$ of a $k$-qubit ($2^k$-dimensional) subsystem of an $N$-qubit ($2^N$-dimensional) universe, obtaining an expression in terms of the digamma function $\Psi(n)$:

$$\overline{H}_{k;N} = \Psi(2^N + 1) - \Psi(2^{(N-k)} + 1) - 2^{k-N-1}(2^k - 1). \tag{3}$$

The average mutual information between a 1-qubit system and an $m$-qubit fragment, all within an $N + 1$-qubit universe, is

$$\overline{\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}}} = \overline{H}_{1;N+1} + \overline{H}_{m;N+1} - \overline{H}_{m+1;N+1}. \tag{4}$$
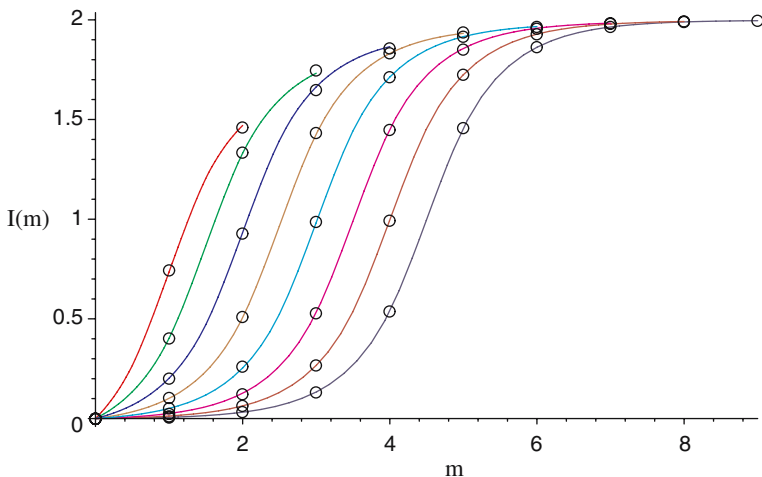
Fig. 2.   Partial information plots (both $\mathcal{I}$ and $m$ are measured in bits), averaged over *all* states, for environments of 2–9 qubits. The circles represent numerical computations, while the solid lines represent theory (analytically continued to non-integer $m$). The plots indicate that information about the system is encoded in entangled states of the environment; nearly $N/2$ individual subenvironments must be captured to yield more than a pittance of information. At $m = N/2$, where $\overline{\mathcal{I}}(m)$ climbs most rapidly, its slope is almost exactly 1.

PIPs computed from Eq. (4) are shown in Fig. 2, along with brute-force numerical calculations which agree extremely well with Eq. (4). The plots demonstrate that for even a modestly large ($N \simeq 9$) environment, the vast majority of randomly chosen states are "encoding states", which encode the information about the system so that it cannot be accessed without capturing at least $\sim N/2$ subenvironments. Such states display no redundancy whatsoever; when three or more observers attempt to divvy up the environment so that each has the same amount of information, each observer learns nothing.

Since the single-qubit system's entropy is limited to 1 bit, $\mathcal{I}_{\mathcal{S}:\mathcal{E}} \leqslant 2$. Nonetheless, for environments larger than about 5 qubits, almost every state achieves the bound. An additional calculation also confirms the indication in Figs. 1 and 2, that the slope of $\overline{\mathcal{I}}(m)$ at $m = N/2$ is almost exactly 1. Thus, for large $N$, almost the entire 2 bits of available information are gained between $m = (N/2) - 1$ and $m = (N/2) + 1$. As we shall see in the next section, these results don't imply that information is encoded this way in *every* state; rather, the vast majority of states are "encoding states."

## 4. INFORMATION STORAGE FOR PRODUCT ENVIRONMENTS

The results of the previous section indicate that the vast majority of states are non-redundant encoding states, in which information is recorded so that multiple observers can never independently measure the same system. Our everyday experience, however, contradicts this result. By intercepting a tiny fraction of the photons (or phonons) which scatter off (or are emitted by) another human being, we obtain a great deal of information about their appearance (or speech); moreover, hundreds or thousands of other human beings obtain essentially the same information! We are thus led to conjecture [2,3] that the ensemble of system–environment states produced by natural processes is substantially different from the ensemble of all states.

Physical processes that produce decoherence typically have a tensor–product structure. The system interacts independently with each subenvironment, and the subenvironments interact only weakly with each other. We thus consider interaction Hamiltonians of the form $H(\mathcal{S}, \mathcal{E}) = \sum_i H(\mathcal{S}, \mathcal{E}_i)$. However, if the initial states of the environment are uniformly distributed over the Haar measure of $\mathcal{H}_\mathcal{E}$, then the action of *any* set of unitaries will yield a uniformly distributed set of states. In order to get a ensemble which is not uniform, we consider only initial product states: $|\Psi_{\mathcal{SE}}\rangle = |\psi_\mathcal{S}\rangle \otimes \bigotimes_i |\psi_{\mathcal{E}_i}\rangle$. Under these assumptions, the "universe" will evolve into "singly-branching" states of the form:

$$|\Psi_{\mathcal{SE}}\rangle = \alpha \, |0\rangle_\mathcal{S} \bigotimes_i |\psi_i\rangle_{\mathcal{E}_i} + \beta \, |1\rangle_\mathcal{S} \bigotimes_i |\psi'_i\rangle_{\mathcal{E}_i}, \tag{5}$$

where $|0\rangle$ and $|1\rangle$ are used, without loss of generality, to indicate the pointer basis of the system.[3,7] This model is close to the one discussed (using a different approach, and asking different questions) in Ref. 4. Subtly different ensembles of such states can be constructed, depending on the measure used for $|\psi\rangle_\mathcal{S}$ and $|\psi\rangle_{\mathcal{E}_i}$, but the most important feature of these states is the *structure* of correlations. The system is strongly entangled with the overall environment, and potentially with some of the subenvironments, but the environments are never entangled with each other. Strong classical correlations, however, typically exist between all subsystems. All these correlations are mediated through $\mathcal{S}$.

Such singly-branching states turn out to be quite amenable to mutual information calculations. Because of the correlation structure, every relevant reduced density matrix is at most rank-2—that is, a state of a virtual qubit. In order to obtain mutual informations, we compute the density matrices for (a) a reduced universe consisting of $\mathcal{S}$ and a set

$\mathcal{E}_m = \{i_1 \ldots i_m\}$ of environments; (b) just $\mathcal{S}$; and (c) just $\mathcal{E}_m$. Using the definition $\gamma_i = \langle \psi_i | \psi_i' \rangle$, we obtain:

$$\rho_{\mathcal{S}\mathcal{E}_m} = \begin{pmatrix} |\alpha|^2 & \alpha^* \beta \left( \prod_{i \notin \mathcal{E}_m} \gamma_i \right) \\ \alpha \beta^* \left( \prod_{i \notin \mathcal{E}_m} \gamma_i^* \right) & |\beta|^2 \end{pmatrix}, \tag{6}$$

$$\rho_{\mathcal{S}} = \begin{pmatrix} |\alpha|^2 & \alpha^* \beta \left( \prod_i \gamma_i \right) \\ \alpha \beta^* \left( \prod_i \right) & |\beta|^2 \end{pmatrix}, \tag{7}$$

$$\rho_{\mathcal{E}_m} = \begin{pmatrix} |\alpha|^2 & \alpha^* \beta \left( \prod_{i \in \mathcal{E}_m} \gamma_i \right) \\ \alpha \beta^* \left( \prod_{i \in \mathcal{E}_m} \gamma_i^* \right) & |\beta|^2 \end{pmatrix}. \tag{8}$$

$$\tag{9}$$

We characterize each density matrix of the form

$$\rho = \begin{pmatrix} x & \sqrt{x(1-x)}\gamma \\ \sqrt{x(1-x)}\gamma^* & 1-x \end{pmatrix} \tag{10}$$

by a *base purity* $P_0 = x^2 + (1-x)^2$ and a *additive decoherence factor* $d = -\log\left(|\gamma|^2\right)$. The idea is that $P_0$ represents the susceptibility of $\rho$ to decoherence in the diagonal basis, while $d$ represents the amount of decoherence ($d = 0$ indicates a pure state, while total decoherence is indicated by $d = \infty$). Conveniently, it's also appropriate to think of $d$ as representing *distinguishability*, since it measures how nearly orthogonal $|\psi_i\rangle$ and $|\psi_i'\rangle$ are. In terms of $P_0$ and $d$, the Von Neumann entropy $H = -\mathrm{Tr}(\rho \log \rho)$ is:

$$H(P_0, d) = \ln 2 - \frac{1}{2}\left[(1+z)\ln(1+z) + (1-z)\ln(1-z)\right],$$
$$\text{where} \quad z = \sqrt{1 - 2(1-P_0)\left(1 - e^{-d}\right)}. \tag{11}$$

The key advantage to this treatment is that the $d$-factors are additive. If we define $d_i = -\log|\gamma_i|^2$, then the decoherence factors for $\mathcal{S}\mathcal{E}_m$, $\mathcal{E}_m$, and $\mathcal{S}$ are

$$d_{\mathcal{S}\mathcal{E}_m} = \sum_{i \notin \mathcal{E}_m} d_i, \tag{12}$$

$$d_{\mathcal{E}_m} = \sum_{i \in \mathcal{E}_m} d_i, \tag{13}$$

$$d_{\mathcal{S}\mathcal{E}_m} = \sum_{\text{all } i} d_i$$
$$= d_{\mathcal{S}\mathcal{E}_m} + d_{\mathcal{E}_m}, \tag{14}$$

while the base purities are all the same. Thus, in order to compute the mutual information between $\mathcal{S}$ and an fragment $\mathcal{E}_m$, we need only know $P_0$, $d_{\mathcal{S}}$ (a measure of the information lost to the *full* environment), and $d_{\mathcal{E}_m}$ (a measure of the amount of information held by $\mathcal{E}_m$). Given these, the mutual information is simply

$$\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}} = H(P_0, d_{\mathcal{S}}) + H(P_0, d_{\mathcal{E}_m}) - H(P_0, d_{\mathcal{S}} - d_{\mathcal{E}_m}), \tag{15}$$

and in cases where $P_0$ and $d_{\mathcal{S}}$ are fixed, we can simply consider $\mathcal{I}(d_{\mathcal{E}_m})$. For the rest of this paper, we'll assume that $P_0 = 1/2$—that is, that the initial state has one full bit of information to be measured. Our investigations indicate that most of our results are almost identical for other values of $P_0$, however.

The simplest singly–branching state is one where the states $|\psi_i\rangle$ and $|\psi_i'\rangle$ are equally distinguishable for each $i$; that is, $d_i = d_0$, $\forall\ i$. In this case, $\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}} = \mathcal{I}(md_0)$, and no averaging over distinct subenvironments is necessary. In Fig. 3, we plot such PIPs for various values of $d_{\mathcal{S}}$. Note that the only free parameter is the total decoherence $d_{\mathcal{S}}$. For small $d_{\mathcal{S}}$, the total mutual information between system and fragment is substantially less than $2H(\mathcal{S})$, and $\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}}$ increases almost linearly with $m$. As $d_{\mathcal{S}}$ increases, $\mathcal{I}_{\mathcal{S}:\mathcal{E}}$ approaches $2H(\mathcal{S})$, and the $\mathcal{I}(m)$ curve becomes more sharply curved, indicating increased redundancy. Perfect redundancy—where each subenvironment has all the classical information that the entire environment has—is achieved only for a GHZ-type state with $d_{\mathcal{S}} = \infty$.

Of course, we can imagine a state in which the system is fully decohered ($d_{\mathcal{S}} = \infty$), but not all subenvironments possess perfect information. More generally, the $d_i$ may be unequal, so that some subenvironments have more information than others. In this case, to obtain a valid PIP we must actually perform an average over subenvironments. Given a distribution function $f_1(d)$ for the $d_i$, we can obtain a distribution function $f_m(d)$ for the total decoherence of $m$ randomly chosen subenvironments. The PIP is then given by
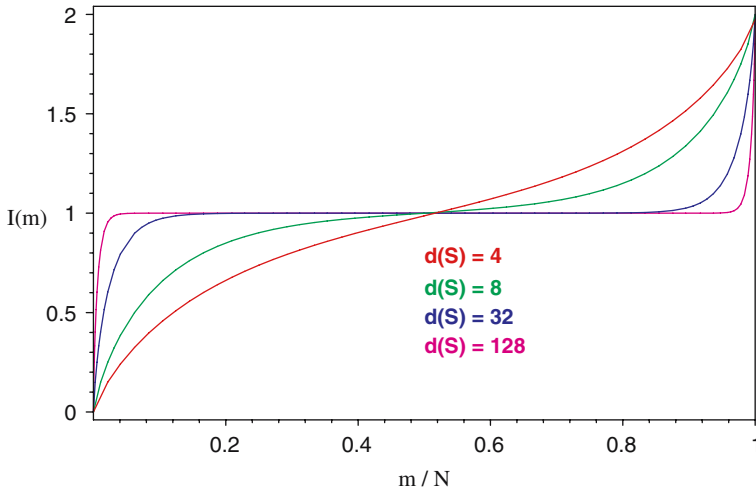
Fig. 3. Partial information plots ($\mathcal{I}(m)$) for unimodal distributions of environment $d$-factors. $\mathcal{I}(m)$ is plotted against the fraction $m/N$ of subenvironments captured. As the total $d_{\mathcal{S}}$ increases from 2 to 8 to 32 to 128, the plots become more sharply inflected near $m = 0$ and $m = N$, while the central plateau at $\mathcal{I} = 1$ bit grows flatter. In the limit as $d_{\mathcal{S}} \longrightarrow \infty$, $\mathcal{I}(m)$ jumps immediately to 1 bit at $m = 0^+$, then jumps again to 2 bits at $m = N^-$.

$$\overline{\mathcal{I}}(m) = \int f_m(d)\mathcal{I}(d)\mathrm{d}d. \qquad (16)$$

In general, both obtaining $f_m(d)$ and computing the given integral are nontrivial. We therefore present two sample cases.

## Case 1: A Single State

Of particular interest is the case where the state $|\Psi\rangle$ is known and has a finite number $N$ of subenvironments. In this case, the $d_i$ form a finite collection, and so both $f_1(d)$ and $f_m(d)$ are discrete probability distributions. A general technique for large $N$ is to treat the process of subenvironment selection as a random walk: $f_1(d)$ is characterized by a mean value $\overline{d}$ and a width $\Delta d$. The Central Limit Theorem then predicts that as $m$ gets large, regardless of what $f_0(d)$ actually is, $f_m(d)$ is very well approximated by a Gaussian distribution. In a conventional random walk, $\overline{d}_m$ is simply $m\overline{d}$, and $\Delta d_m = \sqrt{m}\Delta d$. Our example is slightly different, in that there exists a finite collection of $d_i$, and when $m = N$, each of them must have been picked. A good analogy is a random walk which *must*,

after $N$ steps, arrive at a known endpoint. Thus, while $\bar{d}_m$ still increases linearly with $m$, the width of $f_m(d)$ increases to a maximum at $m = N/2$, then falls to 0 again at $m = N$. The precise result is

$$\bar{d}_m = m\bar{d}, \tag{17}$$

$$\Delta d_m = \sqrt{m\left(1 - \frac{m-1}{N-1}\right)}\Delta d. \tag{18}$$

Using this ansatz to approximate $f_m(d)$ is quite effective when $N$ is large; unfortunately, integrating Eq. (16) is then analytically intractable, so its usefulness is restricted to numerical computations.

A simple example of this sort of analysis which *is* tractable occurs when there are only two kinds of subenvironments "good" subenvironments with $d = d_0$, and "useless" subenvironments with $d = 0$. Such an environment can be characterized by three quantities: the total number $N$ of subenvironments; the number $n_u$ of useful subenvironments, and the distinguishability $d_0$ of each useful subenvironment. When $m$ subenvironments are randomly selected, the probability of choosing $m_u \in [\max(0, m + n_u - N) \ldots \min(m, n_u)]$ useful ones can be computed with elementary combinatorics:

$$P_m(m_u) = \frac{\binom{n_u}{m_u}\binom{N-n_u}{m-m_u}}{\binom{N}{m}}, \tag{19}$$

$$f_m(d) = \sum_{m_u} P_m(m_u)\delta(d - m_u d_0). \tag{20}$$

We integrate Eq. (16) to obtain

$$\overline{\mathcal{I}}(m) = \sum_{k=\max(0, m+n_u-N)}^{\min(m, n_u)} \frac{\binom{n_u}{k}\binom{N-n_u}{m-k}}{\binom{N}{m}}\mathcal{I}(kd_0). \tag{21}$$

Figure 4 compares exact $\overline{\mathcal{I}}(m)$ plots with the naïve approximation $\overline{\mathcal{I}}(m) = \mathcal{I}(n_u d_0/N)$, which achieves the same $d_{\mathcal{S}}$ with indistinguishable environments. In all cases, the bimodal distribution produces a smoother PIP, which rises less quickly and has less of a "plateau" in the middle. This is a simple consequence of the averaging procedure, which effectively convolves $\mathcal{I}(d)$ with a Gaussian, and thus produces a version of the original curve which is greatly smoothed around $m = N/2$, and relatively less smoothed near the ends.
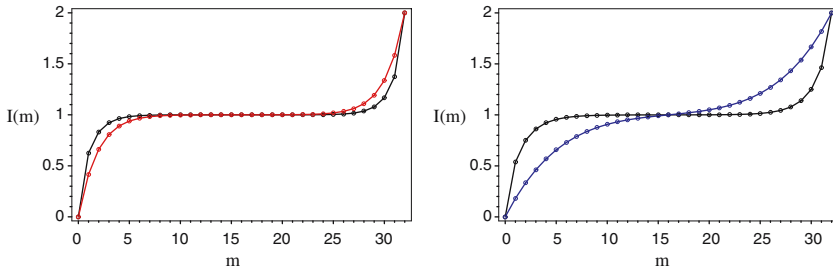
Fig. 4.   Partial information plots ($\mathcal{I}(m)$) for bimodal distributions of environment $d$-factors. The red curve shows $\overline{\mathcal{I}}(m)$ for an environment where half of the $N = 32$ subenvironments have $d = 1.5$ and the other half have $d = 0$. The blue curve shows $\overline{\mathcal{I}}(m)$ when six of them have $d = 3$ and the rest are useless ($d = 0$). The corresponding black curves demonstrate what happens when the equivalent amounts of $d_S$ are spread evenly over all 32 subenvironments—respectively, $d_i = (3/4)$ and $d_i = (9/16)$. In both cases, there is a smoothing effect, but the smoothing is most pronounced when $d_S$ is concentrated in just a few subenvironments.

## Case 2: An Ensemble of States

In Sec.3, we obtained an average PIP for randomly chosen states. For comparison, we'll now obtain an average PIP for all the singly–branching states in Eq. (5). We select the $|\psi_i\rangle$ and $|\psi_i'\rangle$ randomly from the Haar measure for a single qubit, which is uniform over the Bloch sphere. A simple calculation shows that $\gamma_i = |\langle\psi_i|\psi_i'\rangle|^2$ is uniformly distributed between 0 and 1, which yields a perfect Poisson distribution for $d_i$ and $d_m$:

$$f_0(d) = e^{-d}, \tag{22}$$

$$f_m(d) = \frac{d^{m-1}e^{-d}}{(m-1)!}. \tag{23}$$

Thus armed, we proceed as in Sec. 3, splitting $\mathcal{I}_{\mathcal{S}:\mathcal{E}_{\{m\}}}$ into its three component entropies $H(\mathcal{S})$, $H(\mathcal{E}_{\{m\}})$, and $H(\mathcal{S}\mathcal{E}_{\{m\}})$, and averaging them independently. Computing the average entropy of a qubit that has been decohered by $m$ subenvironments turns out to involve a horrendous integral which yields infinite sums of Riemann zeta functions $\zeta(j)$. In the interest of brevity, we provide only the result here.

$$
\begin{aligned}
\overline{H}_m &= \int f_m(x)\mathcal{I}(x)dx \\
&= \frac{1}{(m-1)!}\int_0^\infty dx\ x^{m-1}e^{-x} \\
&\quad \times \left[\ln(2) - \frac{\left(1+e^{-(x/2)}\right)\ln\left(1+e^{-(x/2)}\right) + \left(1-e^{-(x/2)}\right)\ln\left(1-e^{-(x/2)}\right)}{2}\right]
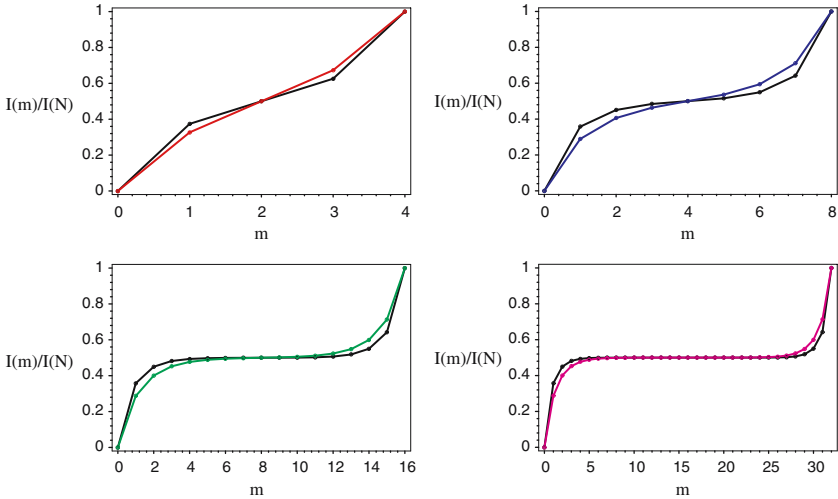\end{aligned}
$$

Fig. 5. Partial information plots $(\overline{\mathcal{I}}(m))$ averaged over the ensemble of all singly–branching states. The four plots represent different numbers $N$ of subenvironments: the red, blue, green, and purple plots correspond, respectively to $N = 4, 8, 16, 32$. The associated curves in black show the behavior of unimodal distributions with the same average $\overline{d}_m = m$. While the ensemble averages are (as expected) less inflected than the unimodal plots, redundant information storage is still quite evident. The ensemble averages are less smoothed than those obtained from the bimodal distribution (Fig. 4), since ensemble average yields a relatively sharply-peaked Poisson distribution for $f_m(d)$. NOTE: all plots have been rescaled on the vertical axis by $\overline{\overline{\mathcal{I}}}(N)$ for ease of comparison.

$$= \left(\frac{3^m - 2^m}{3^m}\right)(\ln(2) - 1) + \frac{m}{2} - \frac{1}{2}\sum_{j=2}^{m}\left(\frac{3^{m+1-j} - 2^{m+1-j}}{3^{m+1-j}}\right)\zeta(j). \quad (24)$$

Figure 5 shows averaged singly–branching PIPs for various $N$. These plots contrast sharply with those in Fig. 2; instead of hovering near $\mathcal{I} = 0$, then shooting up to $\mathcal{I} = \mathcal{I}_{\mathcal{S}:\mathcal{E}}$ at $m = N/2$, they rise rapidly to $\mathcal{I} \simeq H(\mathcal{S})$, then plateau until nearly $m = N$. For the typical singly–branching state, a small fraction of the environment yields a great deal of information about the system. The typical random state, on the other hand, encodes the information so that even a large minority of the subenvironments yields nothing. Fig. 5 also shows that, while typical singly-branching states are slightly less generous with their information than the simple unimodal states discussed earlier, the difference is relatively small. We conclude, then, that singly–branching states *as a class* display a much higher level of redundancy in information storage than typical unstructured states.

## 5. REDUNDANCY

The notion of redundancy has been mentioned, but so far not addressed in sufficient detail. Redundancy was introduced in this context in Refs. 2 and 3, and developed into a more precise tool in Ref. 4, where the information about $\mathcal{S}$ was considered $R$-fold redundant if and only if the environment could be partitioned into $R$ fragments, each of which had nearly full information ("nearly full" meaning that it could supply a fraction $1 - \delta$ of the total information). Reference 4 defines such "$\delta$-redundancy" using Shannon mutual information between observables of the system $\mathcal{S}$ and subsystems of the environment. We follow the idea of asking about "nearly all" information, but depart from Ref. 4 by using quantum mutual information as our metric (thus choosing the other option from the two possibilities considered in Refs. 2 and 3).

Using the quantum mutual information forces us to change our definition of "nearly full information." The metric of information used in Ref. 4 measure *classical* mutual information—information about a single observable. Since it's possible for each of many fragments to have the same amount of information about an observable, "nearly full" in the context of Ref. 4 means $\mathcal{I}_{\text{fragment}} \geqslant (1 - \delta)\mathcal{I}_{\text{total}}$. The quantum mutual information, by contrast, measures information about all possible observables—if qubits A and B are entangled, they can have *two* bits of quantum mutual information, instead of just one. By this metric, the most information about $\mathcal{S}$ that can be shared among many environments is $\mathcal{I}_{\text{total}}/2$ (see Fig. 1), so we redefine "nearly full information" as $\mathcal{I}_{\text{fragment}} \geqslant (1/2)(1 - \delta)\mathcal{I}_{\text{total}}$.

We note, however, that if two fragments each have equal information about *different* parts of the system (this is only possible with systems larger than a single qubit), then each has $(1/2)\mathcal{I}_{\mathcal{SE}}$ mutual information with the system, yet there is no redundancy at all. In fact, in such a case it's not possible to tell for certain whether the stored information is one-fold redundant (i.e., not redundant at all), or two-fold redundant. In general, if $R$ fragments have $(1/2)\mathcal{I}_{\mathcal{SE}}$ information, then the actual redundancy of the information could be anywhere between $R - 1$ and $R$. For this reason, we take the lower bound, and define redundancy as follows:

**Definition 1.** When the full environment $\mathcal{E}$ has $\mathcal{I}_{\mathcal{SE}}$ mutual information with the system, and can be partitioned into at least $R_\delta + 1$ fragments $\mathcal{E}_i$ such that each $\mathcal{E}_i$ has at least $\frac{(1-\delta)}{2}\mathcal{I}_{\mathcal{SE}}$ mutual information with the system, then the information is stored with at least $R_\delta$-fold redundancy, and we say that $R_\delta$ is the redundancy (with an information deficit $\delta$).

## 6. IMPLICATIONS

The main implication of the preceding analysis is clear and simple: arbitrary states of a many-part environment *encode* information in entangled modes of the parts, whereas singly–branching states of the type likely to be produced by decoherence store the information *redundantly*. However, a number of more subtle implications deserve attention.

### 6.1. Redundancy and $\overline{\mathcal{I}}(m)$

The relationship between redundancy and the $\overline{\mathcal{I}}(m)$ plot requires some elucidation. If we consider Definition 1, it's apparent that $R_\delta$ can be calculated for singly–branching states with very little trouble. For a particular $d_\mathcal{S}$, let $d_r$ be the minimum value of $d_{\mathcal{E}_{\{m\}}}$ such that $\mathcal{I}(d_\mathcal{S}, d_{\mathcal{E}_{\{m\}}}) = \frac{(1-\epsilon)}{2}\mathcal{I}_{\mathcal{S}\mathcal{E}}$. Then any fragment $\mathcal{E}_{\{m\}}$ with $d_{\mathcal{E}_{\{m\}}} \geqslant d_r$ has at least $\frac{(1-\epsilon)}{2}\mathcal{I}_{\mathcal{S}\mathcal{E}}$ mutual information with the system. Assuming infinite divisibility of the environment, the number of such fragments that can be created by a partition of $\mathcal{E}$ is precisely $d_\mathcal{E}/d_r$, so the redundancy can be calculated exactly as

$$R = \frac{d_\mathcal{E}}{d_r} - 1. \tag{25}$$

The assumption of infinite divisibility is not really justified, but (a) optimizing the partition exactly is a very tedious combinatoric problem with no closed-form solution, and (b) the $R$ obtained this way is not only a strict upper bound on the exact $R$, but also more useful for some purposes. For instance, when $d_\mathcal{E} = 10$, $d_r = 1$, and each of 11 environments has $d_i = (10/11)$, the strict definition of $R$ would force us to partition $\mathcal{E}$ into five fragments, each of which has almost twice the required $d$. The assumption of infinite divisibility essentially allows us to partition $\mathcal{E}$ into 10 fragments containing 1.1 subenvironments each, which (while physically absurd) better reflects the fact that the information is much more than 4-fold redundant.

The simplicity of this analysis begs the question: why examine $\overline{\mathcal{I}}(m)$ in the first place? Not only does $\overline{\mathcal{I}}(m)$ seem to be irrelevant to redundancy, but in some cases (see the small-$N$ plots in Fig. 5) it can be actively misleading! The answer is that $\overline{\mathcal{I}}(m)$ can be calculated and analyzed for a wide range of states, whereas the redundancy calculation just given is only simple for the sub-ensemble of singly–branching states. Additionally, $\overline{\mathcal{I}}(m)$ provides a good intuition about the qualititative nature of a given

state or ensemble of states; is information generally encoded in entangled states, stored redundantly, or independently distributed among subenvironments?

## 6.2. Perfectly Redundant and Perfectly Encoding States

As noted previously, the primary intuition that we obtain from the $\overline{\mathcal{I}}(m)$ plots is that most states are "coding" states, but an important sub-ensemble of states are "redundant" states. We are naturally led to ask whether "perfect" examples of each type of state exist—that is, a state that encodes information more redundantly than any other state, or a state that hides the encoded information better than any other state.

The answer is somewhat surprising: whereas perfectly redundant states exist for any $N$, perfect coding states apparently exist only for certain $N$. The perfectly redundant states are easy to understand; they are the generalized GHZ (and GHZ-like) states of the form:

$$|\Psi_{\mathcal{SE}}\rangle = \alpha \, |0\rangle_{\mathcal{S}} \bigotimes_i |0\rangle_{\mathcal{E}_i} + \beta \, |1\rangle_{\mathcal{S}} \bigotimes_i |1\rangle_{\mathcal{E}_i}. \tag{26}$$

A true GHZ state is invariant under interchange of any two sub-systems; however, since mutual information is invariant under local unitaries, we only require that the states $|0\rangle_{\mathcal{E}_i}$ and $|1\rangle_{\mathcal{E}_i}$ be orthogonal. Clearly, such states exist for all $N$. Any fragment with $0 < m < N$ has exactly $H(\mathcal{S})$ information, but only by capturing the entire environment ($m = N$) can we obtain the full $\mathcal{I} = 2H(\mathcal{S})$. Thus, the information is stored with $N$-fold redundancy. Note that this sort of state is slightly irregular when analyzed as in Section 4, since $d_i = \infty$.

A perfect coding state, on the other hand, would be one where $\overline{\mathcal{I}}(m) = 0$ for any $m < N/2$, and $\overline{\mathcal{I}}(m) = \mathcal{I}_{\mathcal{SE}}$ for $m > N/2$. An equivalent condition is the existence of two orthogonal states of $N$ qubits, each of which is maximally entangled under all possible bipartite divisions. If such pairs of states exist, then the system states $|0\rangle$ and $|1\rangle$ can be correlated with them to produce the perfect coding state. It is known (as detailed in Ref. 11) that such states only exist for $N = 2, 3, 5, 6$, and possibly for $N = 7$ (for $N = 6$, only a single state exists.[17]). Thus, while for large $N$ almost every state is an excellent coding state, perfect examples seem not to exist except for $N = 2, 3, 5, (7?)$!

## 6.3. Why $\overline{\mathcal{I}}(m)$ can be Misleading

As noted in Secs. 6 and 6.1 plots of $\overline{\mathcal{I}}(m)$ can sometimes be confusing or misleading. An excellent example of this is the case of a singly–branching state with a simple bimodal distribution of $d$: out of $N$ subenvironments, $n_u$ have $d = \infty$ and $N - n_u$ have $d = 0$. This is a GHZ state with $n_u$ subenvironments, onto which $N - n_u$ totally useless ancillas have been tacked. As such, we note immediately that the information is stored with exactly $n_u$-fold redundancy, since any partition of the $N$ subenvironments into $n_u$ fragments such that each fragment contains one useful subenvironment will satisfy the condition. Naively, we might expect that the corresponding $\overline{\mathcal{I}}(m)$ plot would reflect this fact, and appear like a scaled version of a GHZ-state plot.

However, as shown in Fig. 6, diluting the environment in this way has a more dramatic effect. Compared to the scaled GHZ-state plot, the $\overline{\mathcal{I}}(m)$ curves (shown for varying $N$, with $n_u$ held fixed) have been smoothed. For
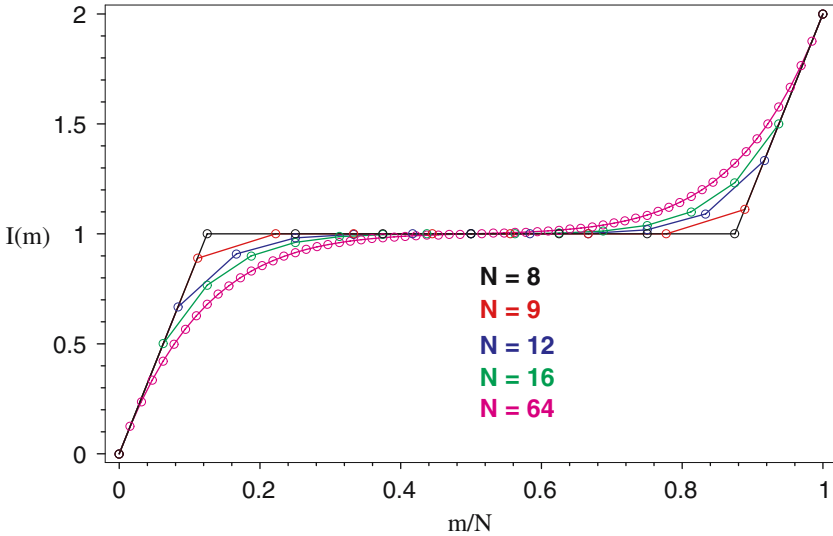


Fig. 6. The effect of diluting the environment. $\overline{\mathcal{I}}(m)$ is plotted for subenvironments containing $N = 8, 9, 12, 16, 64$ qubits; in each case, $n_u = 8$ of the qubit subenvironments are "useful" ($d = \infty$), while the other $N - 8$ are useless ($d = 0$). All five environments are the same from a redundancy perspective, since the extra $d = 0$ qubits are irrelevant. The $\overline{\mathcal{I}}(m)$ curves, however, are disproportionately distorted by the subenvironments which contain no useful information. This dilution effect demonstrates the hazards of inferring redundancy properties from $\overline{\mathcal{I}}(m)$.

sufficiently large $N$, the plots appear to indicate very little redundancy—yet by construction, each environment has exactly the same redundancy.

The explanation is inherent in the well-known fact that $\overline{f(x)} \neq f(\overline{x})$. If $N = 8$ and $n_u = 4$, then on average only $m = 2$ subenvironments must be captured in order to obtain one useful subenvironment (so that $\mathcal{I} = H(\mathcal{S})$). However, the probabilities of capturing (respectively) 0 and 2 useful ones do not cancel each other, since an additional useful subenvironment adds nothing to $\mathcal{I}$, but failure to capture any useful ones reduces $\mathcal{I}$ to 0. Thus, the unlikely event of getting no information at all is disproportionately reflected in the average over possible fragments. The diluted GHZ state is the most dramatic example of this phenomenon, but it occurs in every $\overline{\overline{\mathcal{I}}}(m)$ plot except the unimodal distribution discussed before. Thus, while $\overline{\mathcal{I}}(m)$ is very useful for qualitative analysis of redundancy, it's not quantitatively reliable.

## 7. CONCLUSIONS

This paper introduces a program of analyzing *how* information about a decohered system is stored in its environment, along with some of the first insights into typical results. While we motivate our discussion by discussing decoherence, our results are not specific to any model of decoherence, or even to decoherence processes in general. However, the most important conclusion that we've presented is that *most* states store information nonredundantly, while a certain subclass of states store information redundantly. The evidence of the world around us indicates unambiguously that information about macroscopic objects is highly redundant in their environments, so we conclude that (a) natural processes do *not* produce randomly chosen states, and (b) natural processes *may* favor states of the singly–branching type.

However, this is still only a conjecture [2,3] Such states are favored by the ultra-simplified models of quantum measurement used for toy models, but in reality both measurement and decoherence are far more complex. We do not believe that all physical processes lead exclusively to states of the form given in Eq. (5), but perhaps when decoherence dominates, one obtains some similar ensemble with the same properties. The obvious next step is to analyze actual models of decoherence and quantum measurement, identify the ensembles of states produced by them, and evaluate the redundancy of information storage in those states. We expect that this program of research will sharpen our understanding both of decoherence processes and of the classicality which they induce in the universe.

# REFERENCES

1. W. H. Zurek, *Philos. Transact. R. Soc.* Series A **356**, 1793 (1998).
2. W. H. Zurek, *Annalen der Physik* **9**, 855 (2000).
3. W. H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).
4. H. Ollivier, D. Poulin, and W. H. Zurek, *Phys. Rev. Lett.* **93**, 220401 (2004).
5. W. H. Zurek, arxiv.org/quant-ph/0308163 (2003).
6. A. Peres, *Quantum Theory: Concepts and Methods, vol. 72 of Fundamental Theories of Physics* (Kluwer Academic Publishers, Dordrecht, Boston; 1995).
7. C. A. Fuchs and A. Peres, *Phys. Today* **53**, 70 (2000).
8. W. H. Zurek, *Prog. Theoret. Phys*. **89**, 281 (1993).
9. V. M. Kendon, K. Nemoto, and W. J. Munro, *J. Mod. Opt*. **49**, 1709 (2002).
10. V. M. Kendon, K. Zyczkowski, and W. J. Munro, *Phys. Rev*. A **66**, 62310 (2002).
11. A. J. Scott, *Phys. Rev. A* **69**, 052330 (2004).
12. A. J. Scott and C. M. Caves, *J. Phys. A* **36**, 9553 (2003).
13. D. N. Page, *Phys. Rev. Lett*. **71**, 1291 (1993).
14. S. Sen, *Phys. Rev. Lett*. **77**, 1 (1996).
15. S. Foong and S. Kanno, *Phys. Rev. Lett*. **72**, 1148 (1994).
16. J. Sanchez-Ruiz, *Phys. Rev. E* **52**, 5653 (1995).
17. A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, *IEEE Trans. Inf. Theory* **44**, 1369 (1998).