Submitted to Proceedings of the ITP Conference on Quantum Coherence and Decoherence

Information-theoretic approach to quantum error correction and reversible measurement

M. A. Nielsen,*(1,2) Carlton M. Caves,†(1,2) Benjamin Schumacher,(3) and Howard Barnum(1,2)

(1) *Center for Advanced Studies, Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131-1156*

(2) *Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106-4030*

(3) *Department of Physics, Kenyon College, Gambier, OH 43022*

(June 24, 1997)

Abstract

Quantum operations provide a general description of the state changes allowed by quantum mechanics. The reversal of quantum operations is important for quantum error-correcting codes, teleportation, and reversing quantum measurements. We derive information-theoretic conditions and equivalent algebraic conditions that are necessary and sufficient for a general quantum operation to be reversible. We analyze the thermodynamic cost of error correction and show that error correction can be regarded as a kind of "Maxwell demon," for which there is an entropy cost associated with information obtained from measurements performed during error correction. A prescription for thermodynamically efficient error correction is given.

PACS numbers: 03.65.Bz

*Electronic address: mnielsen@tangelo.phys.unm.edu

†Electronic address: caves@tangelo.phys.unm.edu

I. INTRODUCTION

Quantum operations arise naturally in the study of noisy quantum channels, quantum computation, quantum cryptography, quantum measurements, and quantum teleportation. In each of these applications it is of interest to learn when a quantum operation can be *reversed*. This paper gives a simple, physically meaningful set of necessary and sufficient conditions for determining when a general quantum operation can be reversed, thereby unifying and extending earlier work [1–3]. The picture we develop applies equally well to the reversal of quantum measurements, as in the processes described by Mabuchi and Zoller [4], and to the protection of quantum states against decoherence, as described in the literature on quantum error correction (see, for example, [5] for references). Finally, teleportation [6] can be understood as the reversal of a quantum operation, as was shown in [3], and the results obtained here are being applied in further work on characterizing schemes for teleportation.

The paper is organized as follows. In Sec. II we review the formalism of quantum operations and its application to the theory of generalized measurements, and we define the notion of a reversible operation. Section III introduces information-theoretic measures associated with a quantum operation, which are used throughout the remainder of the paper. These measures, *entanglement fidelity* and *entropy exchange*, were introduced earlier by one of us [1] for the special case of deterministic operations; their definitions and properties are extended to general operations in Sec. III. Section IV A states and proves an information-theoretic characterization of reversibility for general quantum operations, which is then given some simple applications in Sec. IV B. Section V presents an alternative algebraic description of when a quantum operation can be reversed. Though many of the results in Sec. V are already known, we provide new proofs, and the constructions used in these proofs are important in Sec. VI, where we give a thermodynamic analysis of error correction and show that schemes for performing perfect error correction can be done in a thermodynamically efficient way. Section VII offers concluding remarks.

II. QUANTUM OPERATIONS

A. Definition and characterizations

A simple example of a state change in quantum mechanics is the unitary evolution experienced by a closed quantum system. The final state of the system is related to the initial state by a unitary transformation U,

$$\rho \to \mathcal{E}(\rho) = U\rho U^\dagger \; . \tag{2.1}$$

Unitary evolution is, however, not the most general type of state change possible in quantum mechanics. Other state changes, not describable by unitary transformations, arise when a quantum system is coupled to an environment or when a measurement is performed on the system.

How does one describe the most general possible quantum-mechanical dynamics that takes input states to output states? The answer to this question is provided by the formalism of "quantum operations." This formalism is described in detail by Kraus [7] and is given a short, but quite informative review in an appendix to [1]. In this formalism the input state is connected to the output state by the state change

$$\rho \to \frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))} , \tag{2.2}$$

which is determined by a *quantum operation* $\mathcal{E}$. The quantum operation is a linear, trace-decreasing, completely positive map. *Trace decreasing* means that $\mathrm{tr}(\mathcal{E}(\rho)) \leq 1$ for all normalized density operators ρ. *Complete positivity* means that in addition to preserving the positivity of density operators, the map preserves the positivity of all purifications of density operators. The trace in the denominator is included in order to maintain the normalization condition $\mathrm{tr}(\rho) = 1$.

The most general form for a completely positive map $\mathcal{E}$ can be shown to be [7,8]

$$\mathcal{E}(\rho) = \sum_j A_j \rho A_j^\dagger . \tag{2.3}$$

The system operators A_j, which must satisfy

$$\sum_j A_j^\dagger A_j \leq I \tag{2.4}$$

in order that $\mathcal{E}$ be trace decreasing, completely specify the quantum operation. We call Eq. (2.3) an *operator-sum decomposition* of the operation, and we refer to the operators A_j as *decomposition operators.*

The operator-sum decomposition for a quantum operation is not unique, in that another set of decomposition operators $\{A_j\}$ can give rise to the same operation. For example, the operation on a spin-$\frac{1}{2}$ system defined by

$$\mathcal{E}(\rho) = \frac{I}{\sqrt{2}} \rho \frac{I}{\sqrt{2}} + \frac{\sigma_z}{\sqrt{2}} \rho \frac{\sigma_z}{\sqrt{2}} \tag{2.5}$$

can also be written in the form

$$\mathcal{E}(\rho) = \frac{I+\sigma_z}{2} \rho \frac{I+\sigma_z}{2} + \frac{I-\sigma_z}{2} \rho \frac{I-\sigma_z}{2} . \tag{2.6}$$

Choi [9] has classified all sets of decomposition operators that give rise to the same operation. The result is that two sets of decomposition operators, $\{A_k\}$ and $\{B_j\}$, give rise to the same quantum operation if and only if they are related linearly by a square unitary matrix u:

$$B_j = \sum_k u_{jk} A_k . \tag{2.7}$$

It is generally necessary to add some zero decomposition operators to the set with the smaller number of elements so that both sets have the same number of decomposition operators. We call a decomposition *minimal* if no decomposition into a smaller number of operators exists; a decomposition is minimal if and only if the operators in the decomposition are linearly independent.

We say that an operation $\mathcal{E}$ is *pure* if it can be written in terms of an operator-sum decomposition that contains only one operator; that is, there exists an operator A such that

$$\mathcal{E}(\rho) = A\rho A^\dagger \; . \tag{2.8}$$

The unitary transformation (2.1) is an example of a pure quantum operation.

We say that an operation $\mathcal{D}$ is *deterministic* or *trace-preserving* if $\mathrm{tr}(\mathcal{D}(\rho)) = 1$ whenever the input is a normalized density operator ρ. For a deterministic operation, the decomposition operators D_j satisfy a completeness relation

$$\sum_j D_j^\dagger D_j = I \; , \tag{2.9}$$

which implies that $\mathrm{tr}(\mathcal{D}(\rho)) = 1$. Notice that a pure deterministic operation must be a unitary transformation. Any deterministic quantum operation $\mathcal{D}$ can be obtained by adjoining an ancilla system to the system of interest, allowing the system plus ancilla to interact unitarily, and then discarding the ancilla. Such a dynamics leads to a state change of the form

$$\rho \to \mathrm{tr}_A(V(\rho \otimes \sigma^A)V^\dagger) \equiv \mathcal{D}(\rho) \; , \tag{2.10}$$

where tr_A denotes tracing out the ancilla, σ^A is the initial state of the ancilla, and V is the unitary operator for the joint dynamics of the system and ancilla.

For a general quantum operation,

$$\mathrm{tr}(\mathcal{E}(\rho)) = \mathrm{tr}\left(\rho \sum_j A_j^\dagger A_j\right) \tag{2.11}$$

is generally less than one, so $\mathcal{E}(\rho)$ must be renormalized, as in Eq. (2.2), to produce an output density operator. A general quantum operation cannot be represented solely in terms of the joint unitary dynamics of the system and an ancilla, for that always leads to a deterministic operation, as in Eq. (2.10). A general quantum operation can be obtained, however, if the joint dynamics is followed by a measurement on the ancilla; the quantum operation corresponds to particular measurement outcome described by an ancilla projection operator P^A. The resulting state change, once the ancilla is discarded, is given by [7,8]

$$\rho \to \frac{\mathrm{tr}_A((I \otimes P^A)V(\rho \otimes \sigma^A)V^\dagger(I \otimes P^A))}{\mathrm{tr}((I \otimes P^A)V(\rho \otimes \sigma^A)V^\dagger(I \otimes P^A))} \equiv \frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))} \; . \tag{2.12}$$

Notice that $\mathrm{tr}(\mathcal{E}(\rho))$ is the probability of the measurement result described by P^A. A deterministic operation arises in the special case $P^A = I^A$. We note that for any quantum operation, there is a representation of the form (2.12) in which the initial ancilla state σ^A is a pure state.

Suppose the measurement on the ancilla is described by a complete set of orthogonal projection operators P_i^A, where the index i labels the measurement outcomes. Outcome i corresponds to a quantum operation

$$\mathcal{E}_i(\rho) = \mathrm{tr}_A((I \otimes P_i^A)V(\rho \otimes \sigma^A)V^\dagger(I \otimes P_i^A)) = \mathrm{tr}_A((I \otimes P_i^A)V(\rho \otimes \sigma^A)V^\dagger) \; , \tag{2.13}$$

which gives the unnormalized post-measurement state of the system, conditioned on outcome i. The probability for result i is

$$p_i = \mathrm{tr}(\mathcal{E}_i(\rho)) \ . \tag{2.14}$$

If one discards the measurement outcome, the output density operator is obtained by averaging over the outcomes, and the state change is given by a deterministic quantum operation:

$$\rho \to \sum_i p_i \frac{\mathcal{E}_i(\rho)}{\mathrm{tr}(\mathcal{E}_i(\rho))} = \sum_i \mathcal{E}_i(\rho) = \mathcal{D}(\rho) \ . \tag{2.15}$$

Thus a deterministic quantum operation can always be regarded as describing a measurement on the ancilla, whose result is discarded.

In the case of a deterministic operation, the ancilla can be regarded as the system's environment; interaction with the environment gives rise to the nonunitary evolution that is described by the operation. For a general operation, the ancilla can also be regarded as an environment that can be observed and thus that has features of a measuring apparatus. Keeping these connotations in mind, we use the terms ancilla and environment interchangeably in the remainder of the paper.

B. Operations and generalized measurements

The connection of quantum operations to quantum measurements is easy to explain. Standard textbook treatments describe quantum measurements in terms of a complete set of orthogonal projection operators for the system being measured. This formalism, however, does not describe many of the measurements that can be performed on a quantum system. The most general type of measurement that can be performed on a quantum system is known as a *generalized measurement* [7,10,11].

Generalized measurements can be understood within the framework of quantum operations, because any generalized measurement can be performed by allowing the system to interact with an ancilla and then doing a standard measurement described by orthogonal projection operators on the ancilla. Thus, as we can see from Eq. (2.13), the most general type of quantum measurement is described by a set of quantum operations $\mathcal{E}_i$, where the index i labels the possible measurement outcomes. The sum of the operations for the various outcomes is required to be a deterministic quantum operation, as in Eq. (2.15).

Since we can give an operator-sum decomposition for each operation,

$$\mathcal{E}_i(\rho) = \sum_j A_{ij}\rho A_{ij}^\dagger \ , \tag{2.16}$$

we can also say that the generalized measurement is completely described by the system operators A_{ij}, which are labeled by two indices, i and j, and which satisfy the completeness relation

$$\sum_{i,j} A_{ij}^\dagger A_{ij} = I \ . \tag{2.17}$$

If result i occurs, the unnormalized state of the system immediately after the measurement is given by

$$\mathcal{E}_i(\rho) = \sum_j A_{ij}\rho A_{ij}^\dagger \ . \tag{2.18}$$

The probability for result i to occur is

$$p_i = \mathrm{tr}(\mathcal{E}_i(\rho)) = \mathrm{tr}\left(\rho \sum_j A_{ij}^\dagger A_{ij}\right) \ . \tag{2.19}$$

This form makes the connection to the formalism of positive-operator-valued measures. The operators

$$E_i \equiv \sum_j A_{ij}^\dagger A_{ij} \tag{2.20}$$

are elements of a decomposition of the unit operator into positive operators, as in the completeness relation (2.17). Such a decomposition of unity is called a *positive-operator-valued measure* (POVM).

We say a measurement is *pure* if for each measurement result i, the corresponding quantum operation $\mathcal{E}_i$ is pure; that is, there exist operators A_i such that

$$\mathcal{E}_i(\rho) = A_i\rho A_i^\dagger \ . \tag{2.21}$$

The probability that result i occurs is given by

$$p_i = \mathrm{tr}(\rho A_i^\dagger A_i) \ , \tag{2.22}$$

It can be shown that pure measurements correspond to extracting the maximum amount of information about the system from the state of the apparatus to which the system is coupled.

C. Reversal of a quantum operation

When we talk about reversing a quantum operation $\mathcal{E}$, we generally do not mean that $\mathcal{E}$ can be reversed for all input states, but rather only that $\mathcal{E}$ can be reversed for all input density operators ρ whose support lies in a subspace M of the total state space L. In the case of a trace-preserving operation $\mathcal{E}$, the subspace M is sometimes called a *quantum error-correcting code* or simply a *code*. It makes sense to talk about reversing $\mathcal{E}$ on a subspace M only if $\mathcal{E}(\rho) \neq 0$ for all ρ whose support lies in M, and we assume this condition henceforth. We say that a quantum operation $\mathcal{E}$ is *reversible* on a subspace M if there exists a *deterministic* quantum operation $\mathcal{R}$, acting on the total state space L, such that for all ρ whose support lies in M,

$$\rho = \mathcal{R}\left(\frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))}\right) = \frac{\mathcal{R}\circ\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))} \ . \tag{2.23}$$

Here $\mathcal{R}\circ\mathcal{E}$ denotes the composition of $\mathcal{R}$ with $\mathcal{E}$, that is, $\mathcal{R}\circ\mathcal{E}(\rho)\equiv\mathcal{R}(\mathcal{E}(\rho))$. We require the reversal operation $\mathcal{R}$ to be deterministic because we want the reversal definitely to occur, not just to occur with some probability, conditional on some measurement result or ancilla state.

In [2] the problem of reversing deterministic quantum operations was considered. This case is of particular interest in situations where one is unable to obtain information about the environment. In contrast, [3] and [4] considered reversal of operations representing measurements, in which case information about the environment is available.

We say that a measurement is *reversible* on a subspace M of the total state space L if for each measurement result i, the corresponding quantum operation is reversible. Outcomes that have zero probability on M are irrelevant, because they never occur, and thus they can be discarded. We could define measurements that are only sometimes reversible by requiring that only some of the measurement results have reversible quantum operations. Although we do not deal explicitly with such sometimes reversible measurements in this paper, the results obtained in Secs. IV A and V A, since they are derived for individual quantum operations, apply to sometimes reversible measurements.

III. INFORMATION-THEORETIC MEASURES FOR QUANTUM OPERATIONS

Schumacher [1] introduced entanglement fidelity and entropy exchange as useful information-theoretic measures for characterizing deterministic quantum operations. This section extends to general quantum operations the definitions of entanglement fidelity and entropy exchange and generalizes the properties of those quantities obtained in [1] and in [2]. We begin by outlining the particular method for characterizing quantum operations that was used in [1] and [2] and that we use throughout the remainder of this paper.

A. Method for characterizing quantum operations

Suppose we have a quantum system, denoted henceforth by Q, and a quantum operation $\mathcal{E}$ that acts on states of Q. We denote the dimension of the Hilbert space of Q by D. It is convenient to introduce two mathematical artifices, a reference system R, whose Hilbert space has the same dimension, D, as the Hilbert space of Q, and an environment E, which has a Hilbert space of arbitrary dimension.

The joint state of the system Q and the reference system R is chosen so as to *purify* the initial state of Q; that is, RQ is initially in a pure state $\rho^{RQ}=|\Psi^{RQ}\rangle\langle\Psi^{RQ}|$ satisfying

$$\mathrm{tr}_R\Big(|\Psi^{RQ}\rangle\langle\Psi^{RQ}|\Big)=\rho^Q\ , \tag{3.1}$$

where ρ^Q is the initial state of system Q. To reduce the clutter in the notation, we drop the Q superscript when it is clear that we are dealing with the primary quantum system Q. The initial state of the environment E is assumed to be a pure state $\rho^E=|e\rangle\langle e|$, which is uncorrelated with the system RQ. Thus the initial state of the overall system is also pure:

$$|\Psi^{RQE}\rangle=|\Psi^{RQ}\rangle\otimes|e\rangle. \tag{3.2}$$

The joint system QE is subjected to a two-part dynamics consisting of a unitary operation, U^{QE}, followed by a projection onto the environment alone, described by a projector P^E. The reference system R has no internal dynamics and does not interact with Q or E. This two-part dynamics leaves the overall state pure. As we mentioned above, it is always possible to find $|e\rangle$, U^{QE}, and P^E such that

$$\mathcal{E}(\rho^Q) = \mathrm{tr}_E(P^E U^{QE}(\rho^Q \otimes \rho^E) U^{QE\dagger} P^E) = \mathrm{tr}_{RE}(P^E U^{QE}(\rho^{RQ} \otimes \rho^E) U^{QE\dagger} P^E) \,. \tag{3.3}$$

We denote the *normalized* states of the different systems R, Q, and E *after* this evolution by primes. Of special interest is the joint state of RQ after the dynamics, which is given by

$$\rho^{RQ'} = \frac{\mathrm{tr}_E(P^E U^{QE}(\rho^{RQ} \otimes \rho^E) U^{QE\dagger} P^E)}{\mathrm{tr}(P^E U^{QE}(\rho^{RQ} \otimes \rho^E) U^{QE\dagger} P^E)} = \frac{(\mathcal{I}^R \otimes \mathcal{E})(\rho^{RQ})}{\mathrm{tr}((\mathcal{I}^R \otimes \mathcal{E})(\rho^{RQ}))} = \frac{(\mathcal{I}^R \otimes \mathcal{E})(\rho^{RQ})}{\mathrm{tr}(\mathcal{E}(\rho^Q))} \,, \tag{3.4}$$

where $\mathcal{I}^R$ is the identity operation for the reference system. Using the operator-sum decomposition (2.3) of $\mathcal{E}$, we can write $\rho^{RQ'}$ in terms of an ensemble of unnormalized pure states, that is, as a sum of terms each of which is proportional to a one-dimensional projection operator:

$$\rho^{RQ'} = \sum_j \frac{(I^R \otimes A_j)|\Psi^{RQ}\rangle\langle\Psi^{RQ}|(I^R \otimes A_j^\dagger)}{\mathrm{tr}(\mathcal{E}(\rho^Q))} \,. \tag{3.5}$$

The state of the reference system after the dynamics is given by

$$\rho^{R'} = \frac{\mathrm{tr}_{QE}(P^E U^{QE}(\rho^{RQ} \otimes \rho^E) U^{QE\dagger} P^E)}{\mathrm{tr}(P^E U^{QE}(\rho^{RQ} \otimes \rho^E) U^{QE\dagger} P^E)} = \mathrm{tr}_Q(\rho^{RQ'}) \,. \tag{3.6}$$

We emphasize that $\rho^{R'}$ is generally not the same as ρ^R, because of the presence of the environment projector P^E in Eq. (3.6). This is in contrast to the case of a trace-preserving operation $\mathcal{E}$, for which the environment projector is absent, that is, $P^E = I^E$ in Eq. (3.6) and, hence, for which $\rho^{R'} = \rho^R$.

B. Entanglement fidelity and entropy exchange

Following Schumacher [1,2], we define the *entanglement fidelity* to be the fidelity with which the joint state of RQ is preserved by the dynamics:

$$F_e(\rho, \mathcal{E}) \equiv \langle\Psi^{RQ}|\rho^{RQ'}|\Psi^{RQ}\rangle = \frac{\langle\Psi^{RQ}|(\mathcal{I}^R \otimes \mathcal{E})(\rho^{RQ})|\Psi^{RQ}\rangle}{\mathrm{tr}(\mathcal{E}(\rho^Q))} \,. \tag{3.7}$$

Using the form (3.5) of $\rho^{RQ'}$ and noting that

$$\langle\Psi^{RQ}|(I^R \otimes A_j)|\Psi^{RQ}\rangle = \mathrm{tr}\Big(|\Psi^{RQ}\rangle\langle\Psi^{RQ}|(I^R \otimes A_j)\Big) = \mathrm{tr}(\rho^Q A_j) \,, \tag{3.8}$$

we can put the entanglement fidelity in a form that, as implied by our notation $F_e(\rho, \mathcal{E})$, manifestly depends only on the initial state of Q and the operation that is applied to Q,

$$F_e(\rho,\mathcal{E}) = \sum_j \frac{|\mathrm{tr}(\rho A_j)|^2}{\mathrm{tr}(\mathcal{E}(\rho))} \; . \tag{3.9}$$

The *entropy exchange* is defined to be

$$S_e(\rho,\mathcal{E}) \equiv S(\rho^{RQ'}) = S(\rho^{E'}) \; , \tag{3.10}$$

where $S(\rho) = -\mathrm{tr}(\rho \log \rho)$ denotes the von Neumann entropy of the density operator ρ and where the latter equality follows from the fact that the overall state after the two-part dynamics is pure. This generalizes the definition of entropy exchange in [1] to general quantum operations $\mathcal{E}$.

The entropy exchange obeys several inequalities that follow from the subadditivity of von Neumann entropy [12–14] and the purity of the overall state of RQE after the dynamics:

$$S_e = S(\rho^{RQ'}) \le S(\rho^{R'}) + S(\rho^{Q'}) \; , \tag{3.11}$$
$$S(\rho^{Q'}) = S(\rho^{RE'}) \le S(\rho^{R'}) + S(\rho^{E'}) = S(\rho^{R'}) + S_e \; , \tag{3.12}$$
$$S(\rho^{R'}) = S(\rho^{QE'}) \le S(\rho^{Q'}) + S(\rho^{E'}) = S(\rho^{Q'}) + S_e \; . \tag{3.13}$$

Each inequality here is an expression of subadditivity, with equality holding if and only if the joint density operator on the left factors into a product of the two density operators on the right (for example, equality holds in Eq. (3.11) if and only if $\rho^{RQ'} = \rho^{R'} \otimes \rho^{Q'}$). The last two of the above inequalities can be combined into a single inequality,

$$S_e(\rho,\mathcal{E}) = S(\rho^{RQ'}) \ge |S(\rho^{Q'}) - S(\rho^{R'})| \; , \tag{3.14}$$

sometimes known as the *Araki-Lieb inequality* [12,14].

If $\mathcal{E}$ is trace preserving, then as noted below Eq. (3.6), the state of the reference system does not change under the dynamics, that is, $\rho^{R'} = \rho^R$; moreover, since the initial state of RQ is pure, we always have that $S(\rho^R) = S(\rho^Q)$. Thus, for a trace-preserving operation, we can use $S(\rho^{R'}) = S(\rho^Q)$ to eliminate the reference system from the above inequalities, leaving the inequalities obtained by Schumacher [1]. For a general quantum operation, it is not true that $S(\rho^{R'}) = S(\rho^Q)$, and the inequalities must be left in the form given above.

The von Neumann entropy of $\rho^{RQ'}$ is the same as the entropy of the matrix of inner products formed from the unnormalized pure states that contribute to the ensemble for $\rho^{RQ'}$ in Eq. (3.5). Explicitly, in terms of a positive, unit-trace matrix W, whose components are

$$W_{jk} = \frac{\langle \Psi^{RQ}|(I^R \otimes A_k^\dagger)(I^R \otimes A_j)|\Psi^{RQ}\rangle}{\mathrm{tr}(\mathcal{E}(\rho^Q))} \; , \tag{3.15}$$

the entropy exchange is given by

$$S_e(\rho,\mathcal{E}) = S(W) = -\mathrm{tr}(W \log W) \; . \tag{3.16}$$

The components of W can be simplified to the form

$$W_{jk} = \frac{\mathrm{tr}\Big(|\Psi^{RQ}\rangle\langle\Psi^{RQ}|(I^R \otimes A_k^\dagger)(I^R \otimes A_j)\Big)}{\mathrm{tr}(\mathcal{E}(\rho^Q))} = \frac{\mathrm{tr}(A_j \rho A_k^\dagger)}{\mathrm{tr}(\mathcal{E}(\rho))} \; , \tag{3.17}$$

Notice that the diagonal elements of W,

$$q_j = W_{jj} = \frac{\mathrm{tr}(A_j \rho A_j^\dagger)}{\mathrm{tr}(\mathcal{E}(\rho))} \;, \tag{3.18}$$

are the probabilities with which the pure states in Eq. (3.5) contribute to the ensemble.

Relative to a particular density operator ρ, there is a "canonical decomposition" of the quantum operation $\mathcal{E}$. Suppose the matrix W arises from ρ and a particular operator-sum decomposition of $\mathcal{E}$ in terms of decomposition operators A_j, as in Eq. (3.17), and suppose we diagonalize W with a unitary matrix u,

$$\sum_{l,m} u_{jl} W_{lm} u_{km}^* = \lambda_j \delta_{jk} \;, \tag{3.19}$$

where the nonnegative real numbers λ_j are the eigenvalues of W. Now define new operators

$$\tilde{A}_j \equiv \sum_k u_{jk} A_k \;. \tag{3.20}$$

These operators being a unitary remixing of the original decomposition operators, they give another operator-sum decomposition for $\mathcal{E}$, with an associated matrix

$$\tilde{W}_{jk} = \frac{\mathrm{tr}(\tilde{A}_j \rho \tilde{A}_k^\dagger)}{\mathrm{tr}(\mathcal{E}(\rho))} = \sum_{l,m} u_{jl} W_{lm} u_{km}^* = \lambda_j \delta_{jk} \;. \tag{3.21}$$

We say that a decomposition $\tilde{A}_j$ satisfying Eq. (3.21) is a *canonical decomposition* of $\mathcal{E}$ with respect to ρ. The canonical decomposition is unique up to degeneracies in the eigenvalues λ_j and up to (trivial) phase changes in the canonical decomposition operators $\tilde{A}_j$. The entropy exchange can be written as

$$S_e = S(W) = S(\tilde{W}) = -\sum_j \lambda_j \log \lambda_j \equiv H(\vec{\lambda}\,) \;, \tag{3.22}$$

where $H(\vec{\lambda}\,)$ is the Shannon information constructed from the probability distribution λ. It can be shown that

$$S_e \le H(\vec{q}\,) = -\sum_j q_j \log q_j \;, \tag{3.23}$$

equality holding only for a canonical decomposition.

Notice that for a pure quantum operation, the decomposition that contains only a single decomposition operator, as in Eq. (2.8), is the canonical decomposition with respect to any density operator ρ. The canonical W matrix is the one-dimensional unit matrix, and hence the entropy exchange S_e is zero.

Consider the canonical decomposition of an operation $\mathcal{E}$ with respect to the unit density operator I/D, where D is the dimension of the system Hilbert space. Such a canonical decomposition, whose decomposition operators satisfy

$$\mathrm{tr}(\tilde{A}_j \tilde{A}_k^\dagger) = \lambda_j \mathrm{tr}(\mathcal{E}(I)) \delta_{jk} \;, \tag{3.24}$$

is a minimal decomposition of $\mathcal{E}$. All minimal decompositions can be obtained from this one by unitary remixings that leave the number of operators in the decomposition unchanged.

The entanglement fidelity and the entropy exchange obey the *quantum Fano inequality*,

$$S_e(\rho,\mathcal{E}) \le h(F_e(\rho,\mathcal{E})) + (1 - F_e(\rho,\mathcal{E}))\log(D^2 - 1) \ , \tag{3.25}$$

where $h(p) \equiv -p\log p - (1-p)\log(1-p)$ and D is the dimension of the system Hilbert space. The quantum Fano inequality was first derived by Schumacher [1] for trace-preserving operations, but Schumacher's proof goes through unchanged for general quantum operations.

C. Data-processing inequality

In the following we often consider a two-step operation that consists of two successive operations. The situation of interest here is that of reversing an operation, as discussed in Sec. II C: the first of the two operations is the operation to be reversed, and the second is a reversal operation, which is necessarily deterministic. Schumacher and Nielsen [2] derived an important inequality, called the data-processing inequality, for the case in which both operations in a two-step operation are trace preserving. Here we show that the data-processing inequality remains valid for any two-step operation in which the first operation $\mathcal{E}$ is arbitrary, but the second operation $\mathcal{D}$ is deterministic.

In this situation we use E to denote the environment used in the first step and A to denote the ancilla or environment used in the second step. We use double primes to distinguish the normalized states of the various systems after the second step. We can draw several conclusions about the von Neumann entropies of various states in this scenario. In particular, since the overall state after both steps is pure, we have that $S(\rho^{REA''}) = S(\rho^{Q''})$ and that $S(\rho^{EA''}) = S(\rho^{RQ''}) = S_e(\rho,\mathcal{D}\circ\mathcal{E})$ is the entropy exchange for the two-step operation. Moreover, since the state of RQE is pure after the first step and since R and E do not participate in the second step, we have that $S(\rho^{RE''}) = S(\rho^{RE'}) = S(\rho^{Q'})$ and that $S(\rho^{E''}) = S(\rho^{E'}) = S(\rho^{RQ'}) = S_e(\rho,\mathcal{E})$ is the entropy exchange in the first step.

The strong subadditivity property of von Neumann entropy [13–15] constrains the entropies after the two-step dynamics:

$$S(\rho^{REA''}) + S(\rho^{E''}) \le S(\rho^{RE''}) + S(\rho^{EA''}) \ . \tag{3.26}$$

Substituting the entropy relations just derived and re-arranging yields the *data-processing inequality*,

$$S(\rho^{Q'}) - S_e(\rho,\mathcal{E}) \ge S(\rho^{Q''}) - S_e(\rho,\mathcal{D}\circ\mathcal{E}) \ . \tag{3.27}$$

The left-hand side of the data-processing inequality is constrained by Eq. (3.12), leading to the double inequality

$$S(\rho^{R'}) \ge S(\rho^{Q'}) - S_e(\rho,\mathcal{E}) \ge S(\rho^{Q''}) - S_e(\rho,\mathcal{D}\circ\mathcal{E}) \ . \tag{3.28}$$

If $\mathcal{E}$ is trace preserving, then $S(\rho^{R'}) = S(\rho^{Q})$ and this double inequality reduces to the form found by Schumacher and Nielsen [2].

IV. INFORMATION-THEORETIC CHARACTERIZATION OF REVERSIBLE QUANTUM OPERATIONS

A. General information-theoretic characterization

In this section we demonstrate that a general quantum operation $\mathcal{E}$ is reversible on a subspace M of the total state space L if and only if the following two conditions are satisfied:

$$\textit{Condition 1:} \quad \mathrm{tr}(\mathcal{E}(\rho)) = \mu^2 \quad \text{for all } \rho \text{ whose support is confined to } M, \tag{4.1}$$

where μ is a real constant satisfying $0 < \mu \leq 1$;

Condition 2:

$$S(\rho) = S\left(\frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))}\right) - S_e(\rho, \mathcal{E}) \quad \text{for any } \textit{one } \rho \text{ whose support is the } \textit{entirety} \text{ of } M, \tag{4.2}$$

(and then for all ρ whose support is confined to M).

Condition 1 is equivalent to

$$P_M E P_M = \mu^2 P_M \ , \tag{4.3}$$

where

$$E \equiv \sum_j A_j^\dagger A_j \tag{4.4}$$

is the POVM element corresponding to $\mathcal{E}$ and P_M is the projector onto M. Condition 1 has the appealing intuitive interpretation that if we view $\mathcal{E}$ as a dynamics for the system, conditional on some measurement result or post-interaction environment state, knowledge of that result or state gives no information about the initial system state ρ. Condition 2, though less intuitive, states essentially that for initial states whose support lies in M, no quantum information is lost in the dynamics described by $\mathcal{E}$.

We begin by proving necessity. Suppose that $\mathcal{E}$ is reversible on M. Then it was shown in [3] that Condition 1 follows. The reason is that reversibility implies that $\mathcal{R} \circ \mathcal{E}(\rho) = \mathrm{tr}(\mathcal{E}(\rho))\rho$ for all ρ whose support is confined to M; $\mathcal{R} \circ \mathcal{E}$ being linear, the only way this equation can be satisfied is if $\mathrm{tr}(\mathcal{E}(\rho))$ has a constant value $\mu^2 > 0$. Let $\mathcal{E}_M$ be the restriction of $\mathcal{E}$ to M, that is,

$$\mathcal{E}_M(\rho) \equiv \sum_j A_j P_M \rho P_M A_j^\dagger \ . \tag{4.5}$$

Notice that $\mathcal{E}_M(\rho) = \mathcal{E}(\rho)$ if ρ has support lying wholly in M. Let $\overline{M}$ be the subspace that is the orthocomplement of M and $P_{\overline{M}}$ be the projector onto $\overline{M}$. Now introduce a new quantum operation $\mathcal{F}$, whose action on *any* ρ is given by

$$\mathcal{F}(\rho) \equiv \frac{\mathcal{E}_M(\rho)}{\mu^2} + P_{\overline{M}} \rho P_{\overline{M}} \ . \tag{4.6}$$

The reason for introducing $\mathcal{F}$ is that it is a deterministic operation with the property that

$$\mathcal{F}(\rho) = \frac{\mathcal{E}(\rho)}{\mu^2} = \frac{\mathcal{E}(\rho)}{\mathrm{tr}(\mathcal{E}(\rho))} \tag{4.7}$$

for states ρ whose support lies wholly in M. Thus $\mathcal{E}$ is reversible on M if and only if $\mathcal{F}$ is reversible on M. Since $\mathcal{F}$ is deterministic, however, the necessary and sufficient condition for its reversibility is the condition already obtained by Schumacher and Nielsen [2]: $\mathcal{F}$ is reversible on M if and only if

$$S(\rho) = S(\mathcal{F}(\rho)) - S_e(\rho, \mathcal{F}) \tag{4.8}$$

for any one ρ whose support is the entirety of M (and then for all ρ whose support lies in M). We complete the proof of necessity by noting that for states ρ whose support is confined to M, the W matrices of $\mathcal{E}$ and $\mathcal{F}$ are the same, which implies that

$$S_e(\rho, \mathcal{E}) = S_e(\rho, \mathcal{F}) \; . \tag{4.9}$$

Substituting Eqs. (4.9) and (4.7) into Eq. (4.8) yields the second condition (4.2).

The sufficiency of Conditions 1 and 2 is proved in an obviously similar way, but one point should be stressed. For $\mathcal{F}$ to be reversible on M, it is sufficient that Eq. (4.8) hold for any *one* ρ whose support is the *entirety* of M. Thus for $\mathcal{E}$ to be reversible, it is sufficient that the second condition (4.2) hold for any one such ρ.

Before going on, one further point deserves mention. If the initial state of Q is the unit density operator in the subspace M—that is, $\rho = P_M/d$, where $d \le D$ is the dimension of M—then we can dispense with Condition 1. What we are claiming is the following equivalence: $\mathcal{E}$ is reversible on M if and only if

$$\log d = S(P_M/d) = S\left(\frac{\mathcal{E}(P_M/d)}{\mathrm{tr}(\mathcal{E}(P_M/d))}\right) - S_e(P_M/d, \mathcal{E}) \; . \tag{4.10}$$

The necessity of Eq. (4.10) has already been shown. We now demonstrate sufficiency by showing that Eq. (4.10) implies the first condition (4.1).

For this purpose, notice that when $\rho^Q = P_M/d$, the initial pure state of RQ is an entangled state of the form

$$|\Psi^{RQ}\rangle = \frac{1}{\sqrt{d}} \sum_{m=1}^{d} |\chi_m^R\rangle \otimes |\phi_m^Q\rangle \; . \tag{4.11}$$

Here the kets $|\phi_m^Q\rangle$ are an orthonormal basis for the d-dimensional subspace M, and the kets $|\chi_m^R\rangle$ are a set of d orthonormal vectors for R. Substituting this entangled state into Eq. (3.5) yields a new expression for the joint state of RQ after the dynamics:

$$\rho^{RQ'} = \sum_{m,n} |\chi_m^R\rangle\langle\chi_n^R| \otimes \frac{1}{d} \sum_j \frac{A_j |\phi_m^Q\rangle\langle\phi_n^Q| A_j^\dagger}{\mathrm{tr}(\mathcal{E}(P_M/d))} \; . \tag{4.12}$$

The state of the reference system after the dynamics now assumes the form

$$\rho^{R'} = \mathrm{tr}_Q(\rho^{RQ'}) = \sum_{m,n} \frac{1}{d} \frac{\langle\phi_n^Q|E|\phi_m^Q\rangle}{\mathrm{tr}(\mathcal{E}(P_M/d))} |\chi_m^R\rangle\langle\chi_n^R| \ , \tag{4.13}$$

where E is the POVM element associated with $\mathcal{E}$ [cf. Eq. (4.4)]. Since the support of $\rho^{R'}$ lies within the d-dimensional subspace spanned by the vectors $|\chi_m^R\rangle$, we know that $S(\rho^{R'}) \le \log d$. Condition (4.10), when combined with the left inequality in Eq. (3.28), implies that $S(\rho^{R'}) = \log d$. This means that the matrix elements of $\rho^{R'}$ in Eq. (4.13) must be those of the unit density operator P_M/d on M. Hence we conclude that

$$P_M E P_M = \mathrm{tr}(\mathcal{E}(P_M/d)) P_M \ , \tag{4.14}$$

which as already noted in Eq. (4.3), is equivalent to Condition 1, with $\mu^2 = \mathrm{tr}(\mathcal{E}(P_M/d))$.

B. Applications of information-theoretic conditions for reversibility

A number of useful results follow from the information-theoretic characterization of reversible quantum operations found in the preceding subsection. Among these is a general characterization of teleportation schemes, which has been discussed in [3] and is the subject of continuing work. This subsection describes several simpler, but still useful applications of the information-theoretic characterization.

We first show that an operation is reversible by a unitary operation if and only if it acts like a multiple of a unitary operation when restricted to the reversal subspace.

Theorem. A quantum operation $\mathcal{E}$, with decomposition operators A_j, is reversible by a unitary operator U on a subspace M if and only if there exist complex constants c_j such that

$$A_j P_M = c_j U^\dagger P_M \ . \tag{4.15}$$

The constants satisfy

$$\sum_j |c_j|^2 = \mu^2 \ , \tag{4.16}$$

where μ^2 is the constant value of $\mathrm{tr}(\mathcal{E}(\rho))$ on M.

Proof. The sufficiency of the condition (4.15) is obvious. The proof of necessity is to notice that for all ρ, not just those whose support is confined to M, we have

$$U \frac{\mathcal{E}_M(\rho)}{\mu^2} U^\dagger = P_M \rho P_M \ , \tag{4.17}$$

where $\mathcal{E}_M$ is the restriction of ρ to M. Rewritten as

$$\mathcal{E}_M(\rho) = \sum_j A_j P_M \rho P_M A_j^\dagger = \mu^2 U^\dagger P_M \rho P_M U \ , \tag{4.18}$$

this shows that $\mathcal{E}_M$ is a pure operation whose canonical decomposition contains the single operator $\mu U^\dagger P_M$. By the result (2.7) that relates operator-sum decompositions, the conclusion follows. This completes the proof.

A second theorem shows that an operation that is reversible for all initial states acts like a multiple of a unitary operation.

Theorem. A quantum operation $\mathcal{E}$ that is reversible on the entire state space of the system is a positive multiple of a unitary operation; that is,

$$\mathcal{E}(\rho) = \mu^2 U \rho U^\dagger \tag{4.19}$$

for some constant μ satisfying $0 < \mu \leq 1$.

Proof. A simple proof can be obtained by examining the reversal operation constructed in [2] and verifying that it is unitary. The result follows from this and the fact that $\mu^2 \equiv \mathrm{tr}(\mathcal{E}(\rho))$ is a constant.

We present here, however, a purely information-theoretic proof that does not require the explicit construction of a reversal operation. A quantum operation that is reversible on the entire D-dimensional state space of Q satisfies

$$\log D = S(I/D) = S\left(\frac{\mathcal{E}(I/D)}{\mu^2}\right) - S_e(I/D, \mathcal{E})\ . \tag{4.20}$$

Since the entropy is maximized by I/D and the entropy exchange is nonnegative, we see immediately that $S(\mathcal{E}(I/D)/\mu^2) = \log D$ and $S_e(I/D, \mathcal{E}) = 0$. The first of these conclusions means that

$$\mathcal{E}(I/D) = \mu^2 (I/D)\ . \tag{4.21}$$

The second means that the W matrix for initial state I/D is of rank one; thus there exists a unitary matrix u such that

$$\sum_{l,m} u_{jl} W_{lm} u^*_{km} = \delta_{j1}\delta_{k1}\ . \tag{4.22}$$

Defining a canonical decomposition of $\mathcal{E}$ as in Eq. (3.20), the canonical W matrix becomes

$$\tilde{W}_{jk} = \frac{\mathrm{tr}(\tilde{A}_j \tilde{A}_k^\dagger)}{D\mu^2} = \delta_{j1}\delta_{k1}\ . \tag{4.23}$$

It follows that only the first operator in the canonical decomposition, $\tilde{A}_1$, is nonzero. Combining this result with Eq. (4.21) yields

$$\mathcal{E}(I/D) = \tilde{A}_1 \tilde{A}_1^\dagger / D = \mu^2 (I/D)\ , \tag{4.24}$$

which implies that $\tilde{A}_1 = \mu U$. This completes the proof.

This second result has a useful application to teleportation. Recall the basic set-up for teleportation [6]. Alice possesses an unknown quantum state ρ, which she wishes to teleport to Bob. Alice also sends Bob some classical information, which we represent by i. It was shown in [3] that the state of Bob's system, conditioned on the information i, is related to the state of Alice's system by a quantum operation $\mathcal{E}_i$,

$$\rho \rightarrow \frac{\mathcal{E}_i(\rho)}{\mathrm{tr}(\mathcal{E}_i(\rho))}. \tag{4.25}$$

If Bob wishes to achieve teleportation then he must be able to reverse the operation $\mathcal{E}_i$. What the above result shows is that Bob can use a unitary operation to do the reversal, since his reversal must work over the entire space of initial states ρ. No generality is introduced by allowing Bob to use nonunitary reversal operations, that is, by allowing Bob to employ an ancilla to assist in teleportation. Considering only unitary reversals, as was done in [3], is thus sufficient for the study of teleportation.

We can also compare the results obtained in this paper to earlier characterizations of reversible pure and deterministic quantum operations, obtained in [2] and [3] (where pure operations are called ideal operations), and show that these earlier results are special cases of the general characterization embodied in Conditions 1 and 2.

For a deterministic quantum operation $\mathcal{E}$, it is certainly true that $\mathrm{tr}(\mathcal{E}(\rho)) = 1$ is a constant, so Condition 1 is automatic. Thus reversibility for a deterministic operation is equivalent to Condition 2 alone, that is, to $S(\rho) = S(\mathcal{E}(\rho)) - S_e(\rho, \mathcal{E})$, which is the reversibility condition obtained in [2].

For a pure quantum operation $\mathcal{E}$, Condition 1 is equivalent to $P_M A^\dagger A P_M = \mu^2 P_M$. This implies, using the polar-decomposition property of operators, that $AP_M = U^\dagger \sqrt{P_M A^\dagger A P_M} = \mu U^\dagger P_M$ for some unitary operator U. Hence $\mathcal{E}$ can be reversed by U. Thus reversibility for a pure operation is equivalent to Condition 1 alone, that is, to $\mathrm{tr}(\mathcal{E}(\rho)) = \mu^2$, which is the reversibility condition obtained in [3]. For a pure operation, Condition 2 can be dispensed with because it follows from Condition 1: for any pure operation the entropy exchange is zero, and Condition 1 implies that $\mathcal{E}$ acts like a multiple of a unitary on M, which means that $S(\rho) = S(\mathcal{E}(\rho)/\mu^2)$.

V. ALGEBRAIC CHARACTERIZATION OF REVERSIBLE OPERATIONS

Up to this point we have taken an information-theoretic approach to the reversal of quantum operations. In this section we switch to an algebraic point of view. The algebraic results obtained in this section are particularly powerful when used in combination with the information-theoretic viewpoint, as we illustrate in the next section on the thermodynamic cost of error correction. We begin with the theorem that establishes algebraic conditions for reversibility.

A. Reversibility theorem

Theorem. A quantum operation $\mathcal{E}$, with decomposition operators A_j, is reversible on M if and only if there exists a positive matrix m such that

$$P_M A_k^\dagger A_j P_M = m_{jk} P_M \ . \tag{5.1}$$

The trace of m,

$$\sum_j m_{jj} = \mu^2 \ , \tag{5.2}$$

is the constant value of $\mathrm{tr}(\mathcal{E}(\rho))$ on M. (Under a unitary remixing of the decomposition operators, the matrix m undergoes a unitary transformation, which leaves the trace invariant.)

This result was proved by Knill and Laflamme [16] and by Bennett *et al.* [17]. We give a different proof, particularly of the sufficiency of condition (5.1). The construction used in our proof is crucial to our subsequent analysis of the thermodynamics of error correction.

Proof. We deal first with the necessity of condition (5.1) and notice that for all density operators, not just those whose support is confined to M, we have

$$\mathcal{R} \circ \mathcal{E}_M(\rho) = \sum_{l,j} R_l A_j P_M \rho P_M A_j^\dagger R_l^\dagger = \mu^2 P_M \rho P_M \;, \tag{5.3}$$

where the operators R_k make up an operator-sum decomposition for the reversal operation $\mathcal{R}$. Equation (5.3) means that $\mathcal{R} \circ \mathcal{E}_M$ is a pure operation, whose canonical decomposition consists of the single operator μP_M. By the result (2.7) that relates operator-sum decompositions, we can conclude that there exist constants c_{jl} such that

$$R_l A_j P_M = c_{jl} P_M \;. \tag{5.4}$$

The constants satisfy

$$\sum_{j,l} |c_{jl}|^2 = \mu^2 \;. \tag{5.5}$$

Using the trace-preserving property of the reversal operation $\mathcal{R}$, we can write

$$P_M A_k^\dagger A_j P_M = \sum_l P_M A_k^\dagger R_l^\dagger R_l A_j P_M = \left(\sum_l c_{jl} c_{kl}^* \right) P_M = m_{jk} P_M \;, \tag{5.6}$$

where the matrix $m = cc^\dagger$ is manifestly positive.

We now demonstrate that condition (5.1) is sufficient for reversibility. Let u be a unitary matrix that diagonalizes m, that is,

$$\sum_{l,n} u_{jl} m_{ln} u_{kn}^* = d_j \delta_{jk} \;, \tag{5.7}$$

where the nonnegative real numbers d_j are the nonnegative eigenvalues of m. Relative to a new decomposition of $\mathcal{E}$, defined by

$$\tilde{A}_j \equiv \sum_j u_{jk} A_k \;, \tag{5.8}$$

condition (5.1) becomes

$$P_M \tilde{A}_k^\dagger \tilde{A}_j P_M = d_j \delta_{jk} P_M \;. \tag{5.9}$$

The diagonal ($j = k$) elements of Eq. (5.9) imply, by the polar-decomposition property, that there exist unitary operators U_j such that

$$\tilde{A}_j P_M = U_j \sqrt{P_M \tilde{A}_j^\dagger \tilde{A}_j P_M} = \sqrt{d_j} U_j P_M \;. \tag{5.10}$$

Notice that if $d_j = 0$, the corresponding decomposition operator $\tilde{A}_j$ is irrelevant to the operation of $\mathcal{E}$ within M, although such an $\tilde{A}_j$ is generally important to the action of $\mathcal{E}$ on density operators whose support is not confined to M. When there are such decomposition operators, i.e., when $d_j = 0$ for some j, the subspace M is called a *degenerate code*; we discuss the meaning and significance of degenerate codes in Sec. V B. For $d_j \neq 0$ we let M_j be the subspace that M is mapped to by U_j, and we let $P_j \equiv U_j P_M U_j^\dagger$ be the projector onto M_j. The off-diagonal elements of Eq. (5.9) imply that these subspaces are orthogonal, that is,

$$P_k P_j = \delta_{jk} P_j \ . \tag{5.11}$$

The action of $\mathcal{E}$ on any density operator whose support is confined to M takes the following form in terms of the new decomposition:

$$\mathcal{E}(\rho) = \sum_j \tilde{A}_j P_M \rho P_M \tilde{A}_j^\dagger = \sum_j d_j U_j P_M \rho P_M U_j^\dagger = \sum_j d_j P_j U_j \rho U_j^\dagger P_j \ . \tag{5.12}$$

It is easy now to construct an operation that reverses $\mathcal{E}$. Let N be the subspace that is the direct sum of the orthogonal subspaces M_j, and let $\overline{N}$ be the orthocomplement of N. The projector onto $\overline{N}$ is given by

$$P_{\overline{N}} = I - P_N = I - \sum_{\{j|d_j \neq 0\}} P_j \ . \tag{5.13}$$

Now we define the action of a putative reversal operation by

$$\mathcal{R}(\rho) \equiv \sum_{\{j|d_j \neq 0\}} U_j^\dagger P_j \rho P_j U_j + P_{\overline{N}} \rho P_{\overline{N}} \ . \tag{5.14}$$

This reversal operation is trace preserving, as required, since

$$\sum_{\{j|d_j \neq 0\}} P_j U_j U_j^\dagger P_j + P_{\overline{N}} P_{\overline{N}} = \sum_{\{j|d_j \neq 0\}} P_j + P_{\overline{N}} = I \ , \tag{5.15}$$

and simple algebra shows that for all ρ whose support is confined to M,

$$\mathcal{R} \circ \mathcal{E}(\rho) = \mu^2 \rho \ , \tag{5.16}$$

where

$$\mu^2 = \sum_j d_j = \mathrm{tr}(\mathcal{E}(\rho)) \ . \tag{5.17}$$

Thus $\mathcal{R}$ is indeed a reversal operation for $\mathcal{E}$ on the subspace M. This completes the proof.

B. Discussion of algebraic conditions for reversibility

The proof has a compelling physical interpretation. It shows that an operation that is reversible on M has an operator-sum decomposition in which the decomposition operators $\tilde{A}_j$ map M *unitarily* to *orthogonal* subspaces M_j. The operation on the entire space is generally not representable as an ensemble of unitary operations, but as far as its action on the reversible subspace M is concerned, the operation *can* be represented by unitary operators

U_j, which are applied randomly with probabilities $\lambda_j = d_j/\mu^2$ and which, moreover, take M to orthogonal subspaces M_j. Reversal can be effected by first measuring in which of the orthogonal subspaces the state lies after the operation, thus determining which unitary operator U_j occurred, and then applying the corresponding inverse unitary operator $U_j^\dagger$ to restore the initial state. We stress that one can *always* effect reversal in this way, by using a measurement—indeed, a pure, projection-valued measurement—followed by a unitary conditioned on the result of the measurement. It is equally important to emphasize, however, that the deterministic reversal operation can also be constructed without measurements, by using an ancilla as described in Sec. II. The two methods of reversal lead, of course, to the same reversal operation.

The operator-sum decomposition $\tilde{A}_j$ used in the above proof is obviously quite special. It is a canonical decomposition for $\mathcal{E}$ relative to the initial state P_M/d, where d is the dimension of M, as can be seen directly by taking the trace of Eq. (5.9). More interesting is that the operators $\tilde{A}_j$ are a canonical decomposition for $\mathcal{E}$ relative to *any* initial density operator ρ whose support lies in M; that is, the W matrix is diagonal,

$$\tilde{W}_{jk} = \frac{\mathrm{tr}(\tilde{A}_j \rho \tilde{A}_k^\dagger)}{\mu^2} = \frac{\mathrm{tr}(\tilde{A}_j P_M \rho P_M \tilde{A}_k^\dagger)}{\mu^2} = \frac{d_j}{\mu^2}\delta_{jk} = \lambda_j \delta_{jk} \; , \tag{5.18}$$

with eigenvalues λ_j. It follows that

$$S_e(\rho, \mathcal{E}) = S(\tilde{W}) = H(\vec{\lambda}) \; . \tag{5.19}$$

Moreover, for any ρ whose support lies in M, Eq. (5.12) shows that the density operator after application of $\mathcal{E}$ is given by

$$\rho' = \mathcal{E}(\rho)/\mu^2 = \sum_j \lambda_j \rho_j \; , \tag{5.20}$$

where

$$\rho_j \equiv U_j \rho U_j^\dagger \tag{5.21}$$

is the unitary image of ρ in M_j. It follows that $S(\rho_j) = S(\rho)$ and since the density operators ρ_j are orthogonal, that

$$S(\mathcal{E}(\rho)/\mu^2) = H(\vec{\lambda}) + S(\rho) = S(\rho) + S_e(\rho, \mathcal{E}) \; . \tag{5.22}$$

This is an explicit demonstration of condition (4.2).

The existence of the canonical decomposition $\tilde{A}_j$ of Eq. (5.8) clarifies the notion of a degenerate code. A common way of defining degeneracy is to say that a code is degenerate if any of the off-diagonal elements of the matrix m in Eq. (5.1) are nonzero. This definition is flawed, however, because it is not invariant under changes in the operator-sum decomposition of $\mathcal{E}$. The off-diagonal elements of m can always be made to vanish by transforming to the canonical decomposition $\tilde{A}_j$.

Loosely speaking, what degeneracy is supposed to capture is the idea that some of the "errors" produced by $\mathcal{E}$ are irrelevant within the code subspace M. This idea must be translated into a mathematical form that is independent of the operator-sum decomposition of $\mathcal{E}$. One way of doing so was introduced by Gottesman [18]: suppose that the operators A_j constitute a minimal (and thus linearly independent) decomposition of $\mathcal{E}$; if the restricted operators A_jP_M, which form a decomposition of the restricted operation $\mathcal{E}_M$, are linearly dependent, Gottesman calls the code degenerate. The reason for this definition is that if the operators A_jP_M are linearly dependent, then transformation to a minimal decomposition of $\mathcal{E}_M$ reduces the number of decomposition operators, i.e., reduces the number of "errors." The canonical decomposition provides just such a minimal decomposition of $\mathcal{E}_M$, that is, the decomposition consisting of the restricted operators $\tilde{A}_jP_M$. The reduction in the number of errors shows up in that the operators $\tilde{A}_j$ that have $d_j = 0$ are irrelevant to the operation of $\mathcal{E}$ within M. Thus we arrive at the equivalent definition of degeneracy introduced in the above proof: a code is degenerate if one or more of the eigenvalues d_j vanishes.

This discussion leads to a manifestly invariant way of defining degeneracy in terms of operator subspaces: a code is degenerate if the operator subspace spanned by the decomposition operators of $\mathcal{E}$ has higher dimension than the operator subspace spanned by the decomposition operators of $\mathcal{E}_M$. Moreover, it is now clear why degeneracy is considered a possible means of beating the "quantum Hamming bound." That bound is derived from counting the number of possible linearly independent errors, not restricted to the code subspace, and assuming that error correction requires for each error an orthogonal subspace the same size as the reversible subspace M.

We turn now to properties of the reversal operation. In Eq. (5.14) we introduced a particular reversal operation $\mathcal{R}$, which is defined in terms of an operator-sum decomposition that consists of the operators

$$\tilde{R}_j = U_j^\dagger P_j = P_M U_j^\dagger \quad \text{for } j \text{ such that } d_j \neq 0, \tag{5.23}$$

and the operator

$$\tilde{R}_{\overline{N}} = P_{\overline{N}} \; . \tag{5.24}$$

The important part of $\mathcal{R}$ is its restriction to the subspace N; the action of the restricted operation $\mathcal{R}_N$ is defined by

$$\mathcal{R}_N(\rho) \equiv \sum_j \tilde{R}_j P_N \rho P_N \tilde{R}_j^\dagger + \tilde{R}_{\overline{N}} P_N \rho P_N \tilde{R}_{\overline{N}} = \sum_j \tilde{R}_j \rho \tilde{R}_j^\dagger = \sum_{\{j \mid d_j \neq 0\}} U_j^\dagger P_j \rho P_j U_j \; . \tag{5.25}$$

The first question we address is the extent to which the reversal operation is unique. For that purpose, consider another operation $\mathcal{T}$, with decomposition operators T_l, which reverses $\mathcal{E}$ on M. The action of $\mathcal{T}$ on an arbitrary density operator can be written as

$$\mathcal{T}(\rho) = \sum_l T_l P_N \rho P_N T_l^\dagger + \sum_l T_l P_{\overline{N}} \rho P_{\overline{N}} T_l^\dagger + \sum_l T_l (P_N \rho P_{\overline{N}} + P_{\overline{N}} \rho P_N) T_l^\dagger \; . \tag{5.26}$$

The first and second terms in this expression are the restrictions of $\mathcal{T}$ to the orthogonal subspaces N and $\overline{N}$, respectively, and the third term is an additional contribution that can arise when ρ is not block-diagonal with respect to N and $\overline{N}$. The second and third terms are unaffected by the requirement that $\mathcal{T}$ be a reversal operation; the only restrictions on the second and third terms come from the requirement that $\mathcal{T}$ be trace preserving. The first term defines the action of the restriction of $\mathcal{T}$ to the subspace N, that is,

$$\mathcal{T}_N(\rho) \equiv \sum_l T_l P_N \rho P_N T_l^\dagger \ , \tag{5.27}$$

The restricted operation $\mathcal{T}_N$ is the important part of $\mathcal{T}$ for reversal. What we show now is that $\mathcal{T}_N$ is the same operation as $\mathcal{R}_N$.

We proceed by noting that the decomposition operators T_l must satisfy Eq. (5.4) for any decomposition of $\mathcal{E}$ and, in particular, must satisfy it when the decomposition operators for $\mathcal{E}$ are chosen to be the canonical decomposition $\tilde{A}_j$; that is, we must have

$$\sqrt{d_j} T_l U_j P_M = T_l \tilde{A}_j P_M = \tilde{c}_{jl} P_M \ , \tag{5.28}$$

where the constants $\tilde{c}_{jl}$ determine the diagonalized m matrix of Eq. (5.7),

$$\sum_l \tilde{c}_{jl} \tilde{c}^*_{kl} = d_j \delta_{jk} \ . \tag{5.29}$$

We now discard the values of the index j for which $d_j = 0$; this eliminates rows of zeroes from the matrix $\tilde{c}$. For the remaining values of j, we have that

$$T_l P_j U_j = T_l U_j P_M = \frac{\tilde{c}_{jl}}{\sqrt{d_j}} P_M \equiv v_{lj} P_M \ , \tag{5.30}$$

The columns of the matrix v are orthonormal, that is,

$$\sum_l v^*_{lk} v_{lj} = \frac{1}{\sqrt{d_j d_k}} \sum_l \tilde{c}_{jl} \tilde{c}^*_{kl} = \delta_{jk} \tag{5.31}$$

(this means, in particular, that the number of rows of v is not smaller than the number of columns), and thus by adding columns, v can be extended to be a unitary matrix. By moving the unitary operator U_j in Eq. (5.30) to the other side of the expression, we obtain

$$T_l P_j = v_{lj} P_M U_j^\dagger = v_{lj} U_j^\dagger P_j = v_{lj} \tilde{R}_j \ , \tag{5.32}$$

which implies that

$$T_l P_N = \sum_j T_l P_j = \sum_j v_{lj} \tilde{R}_j \ . \tag{5.33}$$

Since the decomposition operators $T_l P_N$ are related to the decomposition operators $\tilde{R}_j$ by a unitary matrix, we can conclude, as promised, that $\mathcal{T}_N$ and $\mathcal{R}_N$ are the same operation.

The upshot is that the part of a reversal operation that actually effects the reversal—that is, the restriction of the reversal operation to the subspace N—is uniquely determined. In what follows this permits us to make general statements about all reversal operations.

The decomposition used to define $\mathcal{R}$ in Eqs. (5.23) and (5.24) is special. For any ρ whose support is confined to M, this decomposition is a canonical decomposition for $\mathcal{R}$ relative to the output state $\rho' = \mathcal{E}(\rho)/\mu^2$ of Eq. (5.20). This fact is crucial to our later analysis of the thermodynamic efficiency of error correction. To prove it, notice that for any ρ whose support is confined to M, we have $P_j \rho' P_k = \lambda_j \delta_{jk} \rho_j$ and $P_{\overline{N}} \rho' = \rho' P_{\overline{N}} = 0$. It follows that the W matrix is diagonal, that is,

$$\tilde{W}_{jk} = \mathrm{tr}(\tilde{R}_j \rho' \tilde{R}_k^\dagger) = \mathrm{tr}(U_j^\dagger P_j \rho' P_k U_k) = \lambda_j \delta_{jk} \tag{5.34}$$

and

$$\tilde{W}_{j\overline{N}} = \tilde{W}_{\overline{N}Nj} = \tilde{W}_{\overline{N}\,\overline{N}} = 0 \; . \tag{5.35}$$

The canonical W matrices for $\mathcal{E}$ and $\mathcal{R}$ being the same, the entropy exchange in the reversal operation is the same as the entropy exchange in $\mathcal{E}$:

$$S_e(\rho', \mathcal{R}) = H(\vec{\lambda}) = S_e(\rho, \mathcal{E}) \; . \tag{5.36}$$

In addition, since $\mathcal{R}(\rho') = \rho$, Eq. (5.22) can be recast as

$$S_e(\rho', \mathcal{R}) = S(\rho') - S(\mathcal{R}(\rho')) \; . \tag{5.37}$$

This result, that the entropy exchange in the reversal operation is equal to the entropy reduction, is important for our discussion of the thermodynamics of error correction in Sec. VI. Since the entropy exchange (5.36) is determined by the restriction of $\mathcal{R}$ to N, Eqs. (5.36) and (5.37) hold for operation that reverses $\mathcal{E}$ on M. We stress, however, that Eq. (5.37) does not hold generally for trace-preserving operations; rather, as Eq. (3.14) shows, all that one can say for a general trace-preserving operation is that the entropy reduction does not exceed the entropy exchange.

We can also make some powerful observations about the reversibility of entire classes of operations. Knill and LaFlamme [16] showed that if an operation $\mathcal{E}$, with decomposition operators A_k, is reversible on M, then any operation $\mathcal{F}$, whose decomposition operators B_j are linear combinations of the A_k, is also reversible on M. This can be seen immediately from Eq. (5.1): if

$$B_j = \sum_k b_{jk} A_k \tag{5.38}$$

(since b_{jk} is not assumed to be a unitary matrix, this is not just a unitary remixing, which would yield another decomposition of $\mathcal{E}$ instead of a new operation), then

$$P_M B_k^\dagger B_j P_M = \sum_{l,m} b_{km}^* b_{jl} P_M A_m^\dagger A_l P_M = \left(\sum_{l,m} b_{jl} m_{lm} b_{km}^* \right) P_M = n_{jk} P_M \; , \tag{5.39}$$

where the matrix $n = bmb^\dagger$ is manifestly positive.

This result can be stated compactly in the language of operator subspaces: the reversibility of an operation $\mathcal{E}$ on M implies the reversibility on M of any operation whose decomposition operators span a subspace of the span of the decomposition operators of $\mathcal{E}$. We stress that the decomposition operators A_k can be written as a linear combination of the decomposition operators B_j only if the operators B_j span the entire operator subspace spanned by operators A_j.

We can go further to show that any operation that reverses $\mathcal{E}$ is also a reversal operation for $\mathcal{F}$. To do so, notice first that the decomposition operators B_j can be written as a linear combination of the canonical decomposition operators $\tilde{A}_k$ of Eq. (5.10):

$$B_j = \sum_k \tilde{b}_{jk} \tilde{A}_k \; . \tag{5.40}$$

Thus the action of $\mathcal{F}$ on any density operator ρ whose support is confined to M can be written as

$$\mathcal{F}(\rho) = \sum_j B_j P_M \rho P_M B_j^\dagger = \sum_{j,k,l} \tilde{b}_{jk} \tilde{b}_{jl}^* \tilde{A}_k P_M \rho P_M \tilde{A}_l^\dagger = \sum_{j,k,l} \sqrt{d_k d_l}\, \tilde{b}_{jk} \tilde{b}_{jl}^* P_k U_k \rho U_l^\dagger P_l \; , \tag{5.41}$$

The constant value of $\mathrm{tr}(\mathcal{F}(\rho))$ is given by

$$\nu^2 = \mathrm{tr}(\mathcal{F}(\rho)) = \sum_{j,k} d_k |\tilde{b}_{jk}|^2 \; , \tag{5.42}$$

and simple algebra shows that $\mathcal{R}$ reverses $\mathcal{F}$:

$$\mathcal{R} \circ \mathcal{F}(\rho) = \nu^2 \rho \; . \tag{5.43}$$

Moreover, since only the restriction of $\mathcal{R}$ to N is involved in the reversal and any operation that reverses $\mathcal{E}$ has the same restriction to N, we can conclude that any operation that reverses $\mathcal{E}$ also reverses $\mathcal{F}$. The converse is not true, however, because the canonical decomposition of $\mathcal{F}$ might map M to a set of orthogonal subspaces whose span is a *proper* subspace of N, in which case reversal of $\mathcal{F}$ would not entail reversal over all of N.

The reversibility theorem proved in Sec. V A can be recast in another, very compact algebraic form. Suppose a quantum operation $\mathcal{E}$, with operator-sum decomposition consisting of operators A_j, can be reversed on a subspace M. Introduce the Hilbert-Schmidt inner product for operators, defined by

$$(N, O) \equiv \mathrm{tr}(N^\dagger O) \; . \tag{5.44}$$

This inner product allows us to define an adjoint of a superoperator, that is, any linear operator on operators. For example, the adjoint of $\mathcal{E}$ is given by

$$\mathcal{E}^\dagger(\rho) = \sum_j A_j^\dagger \rho A_j \; . \tag{5.45}$$

To see that this is an adjoint with respect to the Hilbert-Schmidt inner product, notice that

$$(N, \mathcal{E}(O)) = \mathrm{tr}\left(N^\dagger \sum_j A_j O A_j^\dagger \right) = \mathrm{tr}\left(\left(\sum_j A_j^\dagger N A_j \right)^\dagger O \right) = (\mathcal{E}^\dagger(N), O) \; , \tag{5.46}$$

which is the required inner-product relation for an adjoint. The adjoint of an operation is generally not an operation, since it can be trace-increasing, but it is always a completely positive linear map.

Let $\mathcal{E}_M$ be the restriction of $\mathcal{E}$ to the subspace M, as in Eq. (4.5). Observing that

$$\mathcal{E}_M^\dagger \circ \mathcal{E}_M(\rho) = \sum_{j,k} P_M A_k^\dagger A_j P_M \, \rho \, P_M A_j^\dagger A_k P_M \; , \tag{5.47}$$

we see that condition (5.1) is equivalent to the requirement that $\mathcal{E}_M^\dagger \circ \mathcal{E}_M$ be a positive multiple of the identity operation on M. This requirement can be written as

$$\mathcal{E}_M^\dagger \circ \mathcal{E}_M(\rho) = \gamma^2 P_M \rho P_M \ , \tag{5.48}$$

where the positive constant γ^2 is the trace of m^2, that is,

$$\gamma^2 = \sum_{j,k} |m_{jk}|^2 = \sum_j d_j^2 = \mu^4 \sum_j \lambda_j^2 \ . \tag{5.49}$$

Equation (5.48) is a necessary and sufficient condition for reversibility of $\mathcal{E}$ on M. We note that $\mathcal{E}_M^\dagger$ is generally *not* the reversal operator $\mathcal{R}$, for $\mathcal{E}_M^\dagger$ has decomposition operators $P_M A_j^\dagger = \sqrt{d_j} P_M U_j^\dagger = \sqrt{d_j} U_j^\dagger P_j$, whereas the reversal operation has decomposition operators shorn of the factor $\sqrt{d_j}$. The omission is necessary so that $\mathcal{R}$ is a trace-preserving operation.

An equivalent way of writing condition (5.48) is to require that $\mathcal{E}_M$ preserve, up to the constant γ^2, the Hilbert-Schmidt inner product of operators defined on M; that is,

$$(\mathcal{E}_M(N), \mathcal{E}_M(O)) = (N, \mathcal{E}_M^\dagger \circ \mathcal{E}_M(O)) = \gamma^2 (N, P_M O P_M) = \gamma^2 (P_M N P_M, P_M O P_M) \tag{5.50}$$

for all operators N and O.

VI. ERROR CORRECTION AND THE SECOND LAW OF THERMODYNAMICS

Error correction—that is, reversal of an operation—decreases the entropy of a quantum system, so it is natural to inquire about the thermodynamic efficiency of this process. In this section we address the question of the entropy cost of error correction and show that error correction can be regarded as a sort of "refrigeration," wherein information about the system, obtained through measurement, is used to keep the system cool. Indeed, the method of operation of an error correction scheme is very similar to that of a "Maxwell demon," and the methods of analysis we use are based on those used to resolve that famous problem. As a prelude to our analysis, we review and extend the discussion of the Araki-Lieb inequality found in Sec. III B.

A. Useful inequality

The Araki-Lieb inequality [12–14] states that for two systems, 1 and 2,

$$S(1) - S(2) \le S(12) \ . \tag{6.1}$$

To see this, introduce a third system, 3, which purifies 12. Subadditivity of the von Neumann entropy and the purity of 123 imply that

$$S(1) = S(23) \le S(2) + S(3) = S(2) + S(12) \ , \tag{6.2}$$

which gives the desired result. The inequality in Eq. (6.2) is the statement of subadditivity; equality holds if and only if systems 2 and 3 are in a product state, that is,

$$\rho^{23} = \rho^2 \otimes \rho^3 \ . \tag{6.3}$$

By interchanging the roles of systems 1 and 2 in the above proof, the Araki-Lieb inequality can be written more generally as

$$|S(1) - S(2)| \leq S(12) . \tag{6.4}$$

Suppose we apply the inequality (6.4) to a deterministic quantum operation $\mathcal{D}$:

$$|S(\rho^{R'}) - S(\rho^{Q'})| \leq S(\rho^{RQ'}) . \tag{6.5}$$

For a deterministic quantum operation, we have that $S(\rho^{R'}) = S(\rho^{R}) = S(\rho^{Q})$; furthermore, it is always true that $S(\rho^{RQ'}) = S_e(\rho, \mathcal{D})$. Substituting these identities into the previous equation gives

$$|S(\rho^{Q}) - S(\rho^{Q'})| \leq S_e(\rho, \mathcal{D}) . \tag{6.6}$$

A special case of this inequality is particularly useful in our entropic analysis of error correction:

$$S_e(\rho, \mathcal{D}) \geq S(\rho^{Q}) - S(\rho^{Q'}) = -\Delta S , \tag{6.7}$$

where $\Delta S \equiv S(\rho^{Q'}) - S(\rho^{Q})$ is the change in the entropy of the system. From the equality condition (6.3), we see that equality holds in the preceding equation if and only if

$$\rho^{QE'} = \rho^{Q'} \otimes \rho^{E'} . \tag{6.8}$$

These equality conditions are crucial to the following analysis of thermodynamically efficient error correction.

B. Reversal by a "Maxwell demon"

Consider the error-correction "cycle" depicted in Fig. 1. The cycle can be decomposed into four stages:

1. The system, starting in a state ρ, is subjected to a noisy quantum evolution that takes it to a state ρ^n. We denote the change in entropy of the system during this stage by ΔS. In typical scenarios for error correction, we are interested in cases where $\Delta S \geq 0$, though this is not necessary.

2. A "demon" performs a pure measurement, described by operators $\{B_i\}$, on the state ρ^n. The probability that the demon obtains result i is

   $$p_i = \mathrm{tr}(B_i \rho^n B_i^\dagger) , \tag{6.9}$$

   and the state of the system conditioned on result i is

   $$\rho_i = B_i \rho^n B_i^\dagger / p_i . \tag{6.10}$$

   All error-correction schemes can be done in such a way that a measurement step of this type is included.

3. The demon "feeds back" the result i of the measurement as a unitary operation V_i that creates a final system state

$$\rho^c = V_i \rho_i V_i^\dagger = V_i B_i \rho^n B_i^\dagger V_i^\dagger / p_i \ , \tag{6.11}$$

which is the same regardless of which measurement result was obtained. In the case of error correction this final state is the "corrected" state.

4. The cycle is restarted. In order that this actually be a cycle and that it be a successful error correction, we must have $\rho^c = \rho$.

The second and third stages are the "error-correction" stages. The idea of error correction is to restore the original state of the system during these stages. In this section we show that the reduction in the system entropy during the error-correction stages comes at the expense of entropy production in the environment, which is at least as large as the entropy reduction.

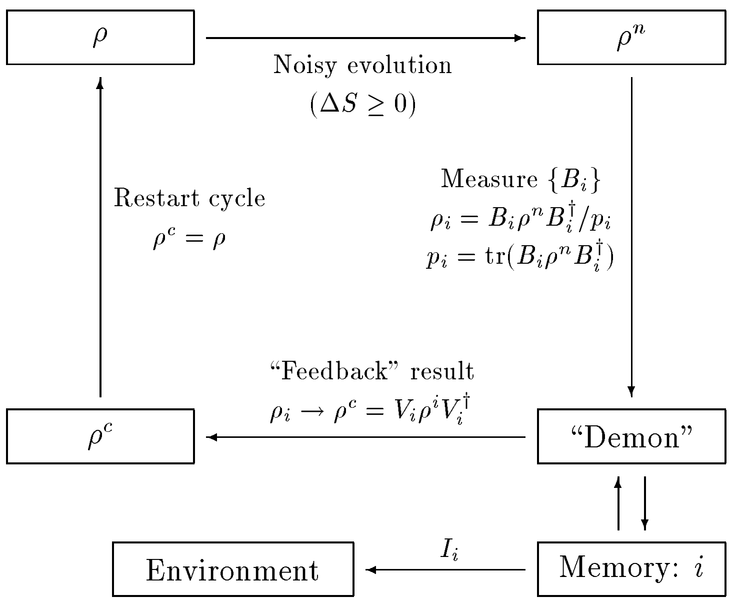

Figure 1. Error correction cycle.

To investigate the balance between the entropy reduction of the system and entropy production in the environment, we adopt the "inside view" of the demon. After stage 3 the only record of the measurement result i is the record in the demon's memory. To reset its memory for the next cycle, the demon must erase its record of the measurement result. Associated with this erasure is a thermodynamic cost, the *Landauer erasure cost* [19,20], which corresponds to an entropy increase in the environment. The erasure cost of information is equivalent to the thermodynamic cost of entropy, when entropy and information are measured in the same units, conveniently chosen to be bits. Bennett [21] used the idea of an erasure cost to resolve the paradox of Maxwell demons, and Zurek [22] and later Caves [23] showed that a correct entropic accounting from the "inside view" can be obtained by quantifying the amount of information in a measurement record by the algorithmic information content I_i of the record. Algorithmic information is the information content of the most compressed form of the record, quantified as the length of the shortest program that can be used to generate the record on a universal computer. We show here that the average thermodynamic cost of the demon's measurement record is at least as great as the entropy reduction achieved by error correction.

In a particular error-correction cycle where the demon obtains measurement result i, the total thermodynamic cost of the error-correction stages is $I_i + \Delta S_c$, where

$$\Delta S_c \equiv S(\rho^c) - S(\rho^n) \tag{6.12}$$

is the change in the system entropy in the error-correction stages. What is of interest to us is the average thermodynamic cost,

$$\sum_i p_i(I_i + \Delta S_c) = \sum_i p_i I_i + \Delta S_c \ , \tag{6.13}$$

where the average is taken over the probabilities for the measurement results. To bound this average thermodynamic cost, we now proceed through a chain of three inequalities.

The first inequality is a strict consequence of algorithmic information theory: the average algorithmic information of the measurement records is not less than the Shannon information for the probabilities p_i, that is,

$$\sum_i p_i I_i \geq H(\vec{p}\,) = -\sum_i p_i \log p_i \ . \tag{6.14}$$

Furthermore, Schack [24] has shown that any universal computer can be modified to make a new universal computer that has programs for all the raw measurement records which are at most one bit longer than optimal code words for the measurement records. On such a modified universal computer, the average algorithmic information for the measurement records is within one bit of the Shannon information H.

To obtain the second and third inequalities, notice that the corrected state ρ^c can be written as

$$\rho^c = \sum_i p_i V_i \rho_i V_i^\dagger = \sum_i V_i B_i \rho^n B_i^\dagger V_i^\dagger \equiv \mathcal{R}(\rho^n) \ , \tag{6.15}$$

where $\mathcal{R}$ is the deterministic reversal operation for the error-correction stages. The operators $V_i B_i$ make up an operator-sum decomposition for the reversal operation. The probabilities p_i are the diagonal elements of the W matrix for this decomposition,

$$p_i = \mathrm{tr}(B_i \rho^n B_i^\dagger) = \mathrm{tr}(V_i B_i \rho^n B_i^\dagger V_i^\dagger) \ . \tag{6.16}$$

Thus we have our second inequality from Eq. (3.23),

$$H(\vec{p}\,) \geq S_e(\rho^n, \mathcal{R}) \ . \tag{6.17}$$

Equality holds here if and only if the operators $V_i B_i$ are a canonical decomposition of $\mathcal{R}$ with respect to ρ^n. We stress that different measurements and conditional unitaries at stages 2 and 3 lead to the same reversal operation, but they yield quite different amounts of Shannon information.

The third inequality is obtained by applying the inequality (6.7) to $\mathcal{R}$ and ρ^n:

$$S_e(\rho^n, \mathcal{R}) + \Delta S_c \geq 0 \; . \tag{6.18}$$

As Eq. (5.37) shows, equality holds here if the operators $V_i B_i$ are a canonical decomposition of $\mathcal{R}$.

Stringing together the three inequalities, we see that the total entropy produced during the error-correction process is greater than or equal to zero:

$$\sum_i p_i I_i + \Delta S_c \geq H(\vec{p}) + \Delta S_c \geq S_e(\rho^n, \mathcal{R}) + \Delta S_c \geq 0 \; . \tag{6.19}$$

Stated another way, this result means that the total entropy change around the cycle is at least as great as the initial change in entropy ΔS, which is caused by the first stage of the dynamics. The error-correction stage can be regarded as a kind of refrigerator, similar to a Maxwell demon, achieving a reduction in system entropy at the expense of an increase in the entropy of the environment due to the erasure of the demon's measurement record.

How then does this error-correction demon differ from an ordinary Maxwell demon? An obvious difference is that the error-correction demon doesn't extract the work that is available in the first step of the cycle as the system entropy increases under the noisy quantum evolution. A subtler, yet more important difference lies in the ways the two demons return the system to a standard state, so that the whole process can be a cycle. For the error-correction demon, it is the error-correction steps that reset the system to a standard state, which is then acted on by the noisy quantum evolution. For an ordinary Maxwell demon, the noisy quantum evolution restores the system to a standard state, typically thermodynamic equilibrium, starting from different input states representing the different measurement outcomes.

Can this error correction be done in a thermodynamically efficient manner? Is there a strategy for error correction that achieves equality in the Second Law inequality (6.19)? The answer is yes, and we give such a strategy here. The proof of the Second Law inequality (6.19) uses three inequalities, $\sum_i p_i I_i \geq H$, $H \geq S_e$, and $S_e \geq -\Delta S$. To achieve thermodynamically efficient error correction, it is necessary and sufficient that the equality conditions in these three inequalities be achieved.

We have already noted that Schack has shown that the first inequality, $\sum_i p_i I_i \geq H(\vec{p})$, can be saturated to within one bit by using a universal computer that is designed to take advantage of optimal coding of the raw measurement records i. On such a universal computer the average amount of space needed to store the programs for the measurement records—that is, the encoded measurement records—is within one bit of the Shannon information H. Moreover, it is possible to reduce this one bit asymptotically to zero by the use of block coding and reversible computation. The demon stores the results of its measurements using an optimal code for a source with probabilities p_i. Thus the demon stores an encoded list of measurement results. Immediately before performing a measurement, the demon decodes the list of measurement results using reversible computation. It performs the measurement, appends the result to its list, and then reëncodes the enlarged list using optimal block coding done by reversible computation. In the asymptotic limit of large blocks, the average length of the compressed list of measurement results becomes arbitrarily close to $H(\vec{p})$ per measurement result.

The second inequality, $H(\vec{p}) \geq S_e(\rho^n, \mathcal{R})$, can be saturated by letting the measurement operators B_i and conditional unitaries V_i be those defined be the those defined by the canonical decomposition of the reversal operation $\mathcal{R}$. It should be noted that the optimal method of encoding the measurement records depends on the probabilities p_i, which in turn are ultimately determined by the initial state ρ. Thus the type of encoding needed to efficiently store the measurement record generally depends on the initial state ρ. For the canonical scheme for error correction, however, the probabilities p_i do not depend on the initial state ρ.

The third inequality, $S_e(\mathcal{R}, \rho^n) \geq -(S(\rho^c) - S(\rho^n))$, is satisfied by any error-correction procedure that corrects errors perfectly. Indeed, in Sec. V B we showed that the entropy exchange associated with any reversing operation is equal to the entropy reduction achieved by the reversing operation (see Eq. (5.37)). An alternative demonstration that perfect error correction achieves the equality $S_e = -\Delta S_c$ begins by noting that at the end of the error-correction process RQ must be in a pure state–the initial state—and therefore the overall state must be a product $\rho^{RQ''} \otimes \rho^{EA''}$ (recall that E is the environment for the noise stage, while A is the ancilla for the reversal stage). Thus the condition $\rho^{QA''} = \rho^{Q''} \otimes \rho^{A''}$ certainly holds. This is the equality condition (6.8) for the Araki-Lieb inequality, applied to the reversal operation. Hence we have $S_e = -\Delta S_c$ for the reversal operation, and we conclude that any successful error-correction procedure automatically achieves equality in Eq. (6.18). It would be interesting to see whether equality can be achieved in Eq. (6.18) by error-correction schemes that do not correct errors perfectly.

C. Discussion

Zurek [25], Milburn [26], and Lloyd [27] have analyzed examples of quantum Maxwell demons, though not in the context of error correction. Lloyd notes that "creation of new information" in a quantum measurement is an additional source of inefficiency in his scheme, which involves measuring σ_z for a spin in a static B-field applied along the z axis, in order to extract energy from it. If the spin is measured in the "wrong" basis—for example, if it is initially in a pure state not an eigenstate of σ_z—the measurement fails to extract all the available free energy of the spin, because of the disturbance to the system state induced by the measurement. In the case of error correction, something similar happens, but it is not disturbance to the system that is the source of inefficiency. Instead, if the ancilla involved in the reversal decoheres in the "wrong" basis—that is, the measurement performed by the demon is not the one defined by the canonical decomposition of the reversal operation—then the Landauer erasure cost is greater than the efficient minimum S_e. This can be thought of a "creation of new information," due to "disturbance" of the ancilla, but the change in the system state is independent of the basis in which the ancilla decoheres.

Error correction can be accomplished in ways other than that depicted in Fig. 1. The "inside view" of the preceding subsection, in which the demon makes a measurement described by some decomposition of the reversal operation, arises when the demon is decohered by an environment, the particular measurement being defined by the basis in which the environment decoheres. If the demon is isolated from everything except the system and is initially in a pure state, then its entropy gain is $S_e = -\Delta S$ for the error-correction process. One can restart the error-correction cycle by discarding the demon and bringing up a new demon, the result being an increase in the environment's entropy by the demon's entropy S_e. This way of performing error correction, which does not involve any measurement records, is equivalent to the "outside view" of the demon's operation.

The "inside view" of the demon's operation, we stress again, arises if the demon's memory is "decohered" by interaction with an environment, the measurement record thus becoming "classical information." In this case the demon has the entropy $H(\vec{p})$ of the measurement record, not just the entropy S_e. Once this decoherence is taken into account, the different decompositions of the reversal operation, corresponding to different measurements, constitute operationally different ways of reversing things, rather than just different interpretations of the same overall interaction. Keeping in mind the variety of decompositions of the reversal operation might lead one to consider a greater variety of experimental realizations, some of which may be easier to perform than others. As we emphasize above, a reversal in which the decohered measurement results correspond to a canonical decomposition of the reversal operation is the reversal method that is most efficient thermodynamically.

VII. CONCLUSION

In this paper we analyze reversible quantum operations, giving both a general information-theoretic characterization and a general algebraic characterization. Our results help in understanding quantum error correction, teleportation, and the reversal of measurements. By applying our two characterizations to a thermodynamic analysis of error correction, we show that the reduction in system entropy due to error correction is compensated by a corresponding increase in entropy of the rest of the world. Moreover, we show that error-correction schemes that correct errors perfectly can be done, in principle, in a thermodynamically efficient manner.

ACKNOWLEDGMENTS

MAN thanks W. H. Zurek for thought-provoking discussions concerning thermodynamics and error correction. This work was supported in part by the Office of Naval Research (Grant No. N00014-93-1-0116). MAN, CMC, and HB thank the Institute for Theoretical Physics for its hospitality during the ITP Workshop on Quantum Computers and Quantum Coherence and for the support of the National Science Foundation (Grant No. PHY94-07194). MAN acknowledges financial support from the Australian-American Educational Foundation (Fulbright Commission).

REFERENCES

[1] B. W. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[2] B. W. Schumacher and M. A. Nielsen, Phys. Rev. A **54**, 2629 (1996).
[3] M. A. Nielsen and C. M. Caves, Phys. Rev. A **55**, 2547 (1997).
[4] H. Mabuchi and P. Zoller, Phys. Rev. Lett. **76**, 3108 (1996).
[5] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, LANL e-print quant-ph/9608006 (1996).
[6] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).
[7] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).
[8] K. Hellwig and K. Kraus, Commun. Math. Phys. **16**, 142 (1970).
[9] M.-D. Choi, Linear Algebra and Its Applications **10**, 285 (1975).
[10] C. W. Gardiner, *Quantum Noise* (Springer-Verlag, Berlin, 1991).
[11] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer Academic, Dordrecht, 1993).
[12] H. Araki and E. H. Lieb, J. Math. Phys. **14**, 160 (1970).
[13] E. Lieb, Bull. Am. Math. Soc. **81**, 1 (1975).
[14] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[15] E. Lieb and M. B. Ruskai, J. Math. Phys. **14**, 1938 (1973).
[16] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).
[17] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[18] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).
[19] R. Landauer, IBM J. Res. Develop. **5**, 183 (1961).
[20] R. Landauer, Nature **355**, 779 (1988).
[21] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).
[22] W. H. Zurek, Phys. Rev. A **40**, 4731 (1989).
[23] C. M. Caves, in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Redwood City, California, 1990), p. 91.
[24] R. Schack, Int. J. Theor. Phys. **36**, 209 (1997).
[25] W. Zurek, in *Frontiers of Nonequilibrium Statistical Physics*, edited by G. T. Moore and M. O. Scully (Plenum, New York, 1984), p. 151.
[26] G. Milburn, preprint (1996).
[27] S. Lloyd, LANL e-print quant-ph/9612034 (1996).