Theory of Quantum Error Correction for General Noise

Emanuel Knill,[1,*] Raymond Laflamme,[1,†] and Lorenza Viola[2,‡]
[1]*Los Alamos National Laboratory, MS B265, Los Alamos, New Mexico 87545*
[2]*d'Arbeloff Laboratory for Information Systems and Technology, Department of Mechanical Engineering,*
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 10 September 1999)

A measure of quality of an error-correcting code is the maximum number of errors that it is able to correct. We show that a suitable notion of "number of errors" e makes sense for any quantum or classical system in the presence of arbitrary interactions. Thus, e-error-correcting codes protect information without requiring the usual assumptions of independence. We prove the existence of large codes for both quantum and classical information. By viewing error-correcting codes as subsystems, we relate codes to irreducible representations of operator algebras and show that noiseless subsystems are infinite-distance error-correcting codes.

PACS numbers: 03.67.Lx, 89.70.+c

The chief reason for the robustness of quantum computation [1–4] is the ability to use quantum error-correcting codes [5,6] to maintain information stored in qubits (two-state particles) subject to environmental noise. Quantum error-correcting codes are defined as subspaces of the qubits' state space with the property that an initial state in this subspace can be recovered if sufficiently few of the qubits experience errors. Provided the noise affecting different qubits is independent and not too intense, any quantum state stored in the subspace can then be regained with high fidelity. This view suffers from several disadvantages. Notably, it is neither obvious whether collective errors can also be corrected well nor is it clear in what sense the information is preserved *before* it is recovered by correcting the errors. In addition, the present theory does not directly lend itself to the application of similar ideas to physical systems that are not canonically decomposable into qubits or are subject to different interaction Hamiltonians.

In this Letter, we overcome the above inconveniences by introducing a description of arbitrary system-environment couplings in terms of an *interaction algebra.* The *degree* of an operator in this algebra determines the temporal order with which the operator can affect the system, regardless of the internal evolution of the environment. For qubits with independent one-qubit interactions, the degree is given by the "number of errors" as defined in combinatorial error analysis. We find that the generalization of minimum distance relates to error correction in the usual way and show that large codes exist depending solely on the dimension of the linear space of errors of a given degree.

Using algebras to classify errors naturally leads to algebraic methods for describing error-correcting codes. The basic idea is to revisit the notion of error-correcting codes as "abstract particles" [7] that are associated with irreducible representations (irreps) of operator algebras closed under Hermitian conjugation (†-closed). Accordingly, error-correcting codes can be viewed as *subsystems* (i.e., tensor factors of subspaces), which makes it clear where the protected information resides. This generalizes a trivial example: Suppose that errors affect all but the first qubit. Then information in the first qubit is clearly safe, and the qubit can be regarded as *noiseless.* We show that *noiseless subsystems* are equivalent to infinite-distance error-correcting codes, and provide the most general method of noise-free information storage, thereby substantially extending the concept of noiseless subspaces [8,9].

Systems and noise.—Let S be a quantum system with state space $\mathcal{S}$. (Fonts are used to distinguish between systems and their state spaces.) $\mathcal{S}$ is an N-dimensional Hilbert space. S interacts with the environment B via an interaction Hamiltonian J that can be written in the form

$$J = \sum_i J_i \otimes B_i \,, \tag{1}$$

where the B_i's are linearly independent environment operators. We assume that the internal evolution of B is removed from J by requiring that $\mathrm{tr}(J_i) = 0$ for all i. The internal evolution of S is retained. If desirable, the latter can be absorbed into a rotating frame, at the expense of making the operators J_i time dependent. This can be accommodated within the present formalism through appropriate redefinitions of the relevant quantities. Our analysis depends on the noise-inducing interaction (1) only through the overall *noise strength* $\lambda = |J|$, where $|J|$ is the maximum eigenvalue of $\sqrt{J^\dagger J}$. The above quantity can be infinite in situations involving infinite-dimensional environments, e.g., the modes of an electromagnetic field. In such cases a redefinition of λ is necessary, based on additional information about the initial state and the internal evolution of the environment. A prototype example is Markovian noise, which will be discussed later.

The second concept we introduce is the *interaction algebra,* which is the algebra $\mathcal{J}$ generated by $\mathcal{J}_1 = \mathrm{span}\{I, J_1, J_2, \ldots\}$, I denoting the identity. Thus, elements of $\mathcal{J}$ are linear combinations of products of operators in $\mathcal{J}_1$. The linearly closed set $\mathcal{J}_1$ consists of the operators of *degree* (at most) one. Next define $\mathcal{J}_d = \mathcal{J}_1^d$, the linear span of products of d or less operators in $\mathcal{J}_1$. These are the operators of *degree* (at most) d [10]. $\mathcal{J}_1$ is well defined in the sense that it is independent of the choice of operators B_i, provided that they are linearly independent. Because the interaction Hamiltonian is Hermitian, $\mathcal{J}_1$ is $\dagger$-closed, implying that $\mathcal{J}_d$ and $\mathcal{J}$ are $\dagger$-closed.

This general formalism applies to qubits with the standard error models. If S consists of n qubits, then $N = 2^n$, and a *linear* interaction satisfies that each J_i involves only Pauli operators $\sigma_u^{(k)}$ acting on one qubit. Here k is a qubit label, $k \in \{1, \dots, n\}$, and u is one of x, y, or z. *Collective* linear interactions involve global operators $J_u = \sum_k \sigma_u^{(k)}$, corresponding to a situation where a single environment couples symmetrically to all qubits [8]. In most discussions of quantum error correction, the interaction is assumed to be *independent*, meaning that each qubit interacts with its own environment [7]. Independent interactions are linear. For linear interactions with qubits, $\mathcal{J}_d$ consists of linear combinations of products of at most d Pauli operators. Note that the Lie algebra generated by $\mathcal{J}_1$ need not include the higher-order errors: The Lie algebra generated by the linear interactions contains only linear interactions, while the effect of an environment coupled linearly can include any other higher-order operator and is not restricted to the unitary group generated by the Lie algebra.

Minimum distance and error correction.—Noise for S can now be analyzed purely in terms of $\mathcal{J}_1$ and λ. By straightforward generalization of the definitions for qubits and independent interactions, we can define a *minimum distance d* quantum code for S and $\mathcal{J}_1$ as a code that *detects* [11] all errors in $\mathcal{J}_{d-1}$. Recall that a (quantum) code of S is a subspace $\mathcal{C} \subset S$, which can be defined through the associated projector $\Pi_{\mathcal{C}}$. Error E is detected by $\mathcal{C}$ if the following protocol works: (1) Prepare a state $|\psi\rangle$ in $\mathcal{C}$. (2) Allow error E to occur, so that the new state is $E|\psi\rangle$. (3) Make a measurement to detect whether the state is in $\mathcal{C}$ or in the orthogonal complement; the outcome is either $\Pi_{\mathcal{C}} E|\psi\rangle$ or $(I - \Pi_{\mathcal{C}})E|\psi\rangle$. (4) Accept the state in the former case and reject it otherwise. The protocol is correct if accepted states are proportional to the initial state, i.e., formally, $\Pi_{\mathcal{C}} E \Pi_{\mathcal{C}} = \alpha_E \Pi_{\mathcal{C}}$.

In many cases we need to preserve only *classical* information. We define a code to have *minimum c-distance d* if a basis of $\mathcal{C}$ exists, such that the above protocol is correct when restricted to basis elements. We will use the term *c code* to denote a code intended only for transmission of classical information in some basis. The notion of error detection can be extended to c codes if a transmission basis is provided. Thus, we say that the c code $\mathcal{C}$ with orthonormal basis $|c_1\rangle, |c_2\rangle, \dots$ *detects* E if $\langle c_i|E|c_j\rangle = \alpha_{i,E}\delta_{i,j}$.

An *e-error-correcting code* permits correction of all errors in $\mathcal{J}_e$, which means that an initial state in the code can be recovered by some fixed quantum operation after an error in $\mathcal{J}_e$ has occurred. Minimum distance is related to error correction in the usual way.

Theorem 1: *A minimum (c-) distance* $2e + 1$ *code is an e-error-correcting (c) code.*

Proof: Recall the necessary and sufficient conditions for a code $\mathcal{C}$ to permit correction of the errors in $\mathcal{J}_e$ [7,12]: $\mathcal{C}$ detects the operators in $\mathcal{J}_e^\dagger \mathcal{J}_e$. This condition is also valid for c codes with a transmission basis. Since $\mathcal{J}_e^\dagger \mathcal{J}_e \subset \mathcal{J}_{2e}$, the result follows. ■

Error bounds.—To make the analysis based on minimum distance and e-error correction useful, it is necessary to show that e-error-correcting codes protect information well. We give a quantitative relationship for the worst-case error as a function of time t. Error is measured in terms of error amplitude, which is the amplitude of the part of the state orthogonal to the intended state.

Theorem 2: *The error amplitude of information protected in an e-error-correcting (c) code is at most* $(\lambda t)^{e+1}/(e + 1)!$.

We defer the proof until after error-correcting codes have been characterized as subsystems. Note that independence from the internal Hamiltonian of the environment implies that even if the latter is subject to arbitrary, adversarial manipulation, the error-correcting code still effectively protects information on a time scale of $O(1/\lambda)$.

Existence of large codes.—A goal of constructing good error-correcting codes is to maximize the dimension of minimum (c-)distance d codes. The greedy algorithm for constructing good minimum-distance classical codes works well in the general case. Let $\{E_1 = I, E_2, \dots, E_D\}$ be a basis of $\mathcal{J}_{d-1}$, with dimension D, and let $\lceil x \rceil$ denote the least integer $\geq x$.

Theorem 3: *There exist codes of S with minimum c-distance d of dimension at least* $\lceil \frac{N}{D} \rceil$.

Proof: Minimum c distance is equivalent to the existence of an orthonormal basis $|c_1\rangle, \dots, |c_k\rangle$ of the c code such that, for each operator E_l to be detected,

$$\langle c_i|E_l|c_j\rangle = \alpha_{i,l}\delta_{i,j} \,. \qquad (2)$$

The proof greedily constructs such a basis. Let $|c_1\rangle$ be any state of S. Suppose that $|c_1\rangle, \dots, |c_k\rangle$ have been found, fulfilling (2). Choose $|c_{k+1}\rangle$ orthogonal to the vectors $E_i|c_j\rangle$, $i = 1, \dots, D$, $j = 1, \dots, k$. Such a state exists provided that $kD < N$. The new set of $|c_i\rangle$ satisfies (2). Upon continuing until the set cannot be extended, a c code of dimension at least N/D is found. ■

Our best general construction of good codes for *quantum* information is based on finding a subcode of a c code.

Theorem 4: *There exist minimum-distance d codes of S of dimension at least* $\lceil \frac{N}{D} \rceil \frac{1}{D+1}$.

Proof: Let $\mathcal{C}$ be a c code of S of dimension at least $\lceil \frac{N}{D} \rceil$ with basis $|c_i\rangle$ satisfying (2). Let Y be the set of indices of the basis vectors. To construct a large quantum code, we seek a partition of Y into subsets Y_i and non-negative coefficients $\beta_{i,j}$, satisfying $\sum_{j \in Y_i} \beta_{i,j} = 1$. Let $|q_i\rangle = \sum_{j \in Y_i} \sqrt{\beta_{i,j}}\, |c_j\rangle$. Then the orthonormal vectors $|q_i\rangle$ span the desired code provided

$$\langle q_i|E_l|q_j\rangle = \gamma_l\, \delta_{i,j}, \qquad \forall\ i, j, l \,. \qquad (3)$$

Compute $\gamma_{l,i} = \sum_{j \in Y_i} \beta_{i,j}\alpha_{j,l}$, the $\alpha_{j,l}$'s being given in (2). We need the $\gamma_{l,i}$ to be independent of i. This problem can be cast in terms of a convex sets problem. We need to find as many disjoint subsets of the set of vectors $\vec{\alpha}_j = \{\alpha_{j,l}\}_l$ with the property that their convex closures have a common intersection. Since $\mathcal{J}_{d-1}$ is †-closed, the $\vec{\alpha}_j$ live in a subspace of real dimension D. By invoking a generalization of Radon's theorem [13], a necessary condition for the existence of at least r such sets is $r(D + 1) - D \leq \lceil N/D \rceil$. The result follows. ■

Subsystems.—If a system consists of a number of qubits, the obvious subsystems are the qubits. If the system consists of a number of photon modes, each mode is a subsystem. However, in order to use these modes as qubits, one could choose the two polarization states for a single photon in a mode as the computational basis. The relevant system is then the subspace where each mode is occupied by exactly one photon, and it is in this subspace that we can identify the qubit subsystems. In both examples, subsystems appear as factors (in the tensor product sense) of subspaces of a larger state space. To avoid working with explicit bases and states, it is convenient to resort to a general algebraic definition. We shall characterize a subsystem of S in terms of a subalgebra of operators acting on S together with an irrep of the subalgebra. This is motivated by the following result from the representation theory of †-closed operator algebras [14].

Theorem 5: *Let $\mathcal{A}$ be a †-closed algebra of operators on S, including the identity. Then S is isomorphic to a direct sum,*

$$S \sim \sum_i C_i \otimes Z_i\,, \tag{4}$$

in such a way that in the representation on the right-hand side, $\mathcal{A} = \sum_i \mathrm{Mat}(C_i) \otimes I^{(Z_i)}$ and the commutant of $\mathcal{A}$ is given by $Z(\mathcal{A}) = \sum_i I^{(C_i)} \otimes \mathrm{Mat}(Z_i)$.

Here, $\mathrm{Mat}(\mathcal{H})$ means the set of all linear operators from $\mathcal{H}$ to itself, while $Z(\mathcal{A})$ is the space of all operators commuting with $\mathcal{A}$. Formally, each *factor* Z_i (C_i) in Theorem 5 defines a *subsystem* of S with associated *state space* $\mathcal{Z}_i$ ($\mathcal{C}_i$). Accordingly, subsystems are naturally definable in terms of either algebras or their commutants.

Noiseless subsystems.—Consider the interaction algebra $\mathcal{J}$ associated with (1). Since $\mathcal{J}$ is †-closed, the representation of Theorem 5 applies. For each subsystem Z_i, states in $\mathcal{Z}_i$ are completely immune to the interaction, as the interaction operators act only on the cofactor $\mathcal{C}_i$. Thus, Z_i is a *noiseless subsystem,* i.e., a subsystem where information is intrinsically stabilized against the effects of the noise with no need for corrective action. Noiseless subspaces [8,9] can be recognized as special cases of the general decomposition (4) for interaction algebras supporting one-dimensional irreps, in which case $\dim\mathcal{C}_i = 1$ for some i's. However, noiseless subsystems can exist in the absence of noiseless subspaces.

Example.—Let us consider three qubits A, B, C with collective linear interactions. The interactions are the generators for spatial rotations. As pointed out in [8], no state of three qubits is invariant under spatial rotations, the minimal implementation of a noiseless subspace requiring $n = 4$ qubits. However, the state space decomposes into one spin-$\frac{3}{2}$ and two spin-$\frac{1}{2}$ irreducible subspaces. The two spin-$\frac{1}{2}$ components together are representable as the product of two two-state spaces as in Theorem 5, with $\mathcal{J}$ acting only on the first. Thus the second one is a noiseless subsystem. Another method of finding this subsystem is to observe that the commutant $Z(\mathcal{J})$ is nontrivial. In particular, it includes the scalars under spatial rotations,

$$s_1 = \sigma_x^{(A)}\sigma_x^{(B)} + \sigma_y^{(A)}\sigma_y^{(B)} + \sigma_z^{(A)}\sigma_z^{(B)}, \tag{5}$$

$$s_2 = \sigma_x^{(A)}\sigma_x^{(C)} + \sigma_y^{(A)}\sigma_y^{(C)} + \sigma_z^{(A)}\sigma_z^{(C)}, \tag{6}$$

which are generating observables for the noiseless subsystem. Equivalently, the latter is seen to support one of the irreps of the algebra generated by the scalars.

Error-correcting codes as subsystems.—The traditional view of error-correcting codes involves encoding the information and correcting errors after the information carriers are transmitted through a noisy channel. The concept of a noiseless subsystem shows that, for the purposes of information maintenance, it is not necessary to correct errors, insofar as they affect components independent of the system where information is stored. In general, we wish to protect the information against all errors in $\mathcal{J}_e$ for some reasonably large e. Since a subsystem unaffected by the operators in $\mathcal{J}_e$ is automatically noiseless, but in most cases of interest noiseless subsystems do not exist, one needs to take an active role in maintaining information. Rather than using error correction to restore the overall state of the system *after* errors happened, we propose to use a quantum operation *before* the latter occur, in such a way that the net effect of the quantum operation followed by errors in $\mathcal{J}_e$ assures preservation of the information in a subsystem. A quantum operation is described by a family $\mathcal{A} = \{A_i\}_i$ of linear operators acting on $\mathcal{S}$, evolving the system density operator as $\rho \mapsto \sum_i A_i \rho A_i^\dagger$. Assuming that no error takes place during the quantum operation $\mathcal{A}$, the combined action of $\mathcal{A}$, followed by errors in $\mathcal{J}_e$, is represented by the product of an operator $E \in \mathcal{J}_e$ and one of the operators $A_i \in \mathcal{A}$. Thus, a state of a noiseless subsystem of the †-closed algebra generated by $\mathcal{J}_e\mathcal{A}$ is preserved in this process.

Theorem 6: *Every e-error-correcting code arises as a noiseless subsystem of $\mathcal{J}_e\mathcal{A}$ for some $\mathcal{A}$ with the property that $I \in \mathrm{span}(\mathcal{A}^\dagger\mathcal{A})$. Conversely, every noiseless subsystem of $\mathcal{J}_e\mathcal{A}$ with $\mathcal{A}$ satisfying the above condition corresponds to an e-error-correcting code.*

Proof: The fact that error-correcting codes yield such noiseless subsystems follows from Theorem III.5 of [7] by letting $\mathcal{A}$ consist of operators that return the state of the error system to the state $|\mathcal{E}(0)\rangle$ (using the language of and in the notation of [7]). Conversely, the condition $I \in \mathrm{span}(\mathcal{A}^\dagger\mathcal{A})$ ensures the existence of a quantum operation whose operators are in $\mathcal{A}$. Thus, the process suggested earlier protects the information against errors in $\mathcal{J}_e$. Because of the necessity of the conditions for error-correcting codes [7], there exists an associated e-error-correcting code in the usual sense. ■

As a consequence, noiseless subsystems are infinite-distance quantum error-correcting codes.

Error analysis.—We now prove Theorem 2 by viewing error-correcting codes as subsystems protected by an initial quantum operation $\mathcal{A}$.

Proof of Theorem 2: By purifying the environment [15], we can assume that the environment's initial state is $|\psi\rangle_B$. The initial state of the system has the intended state in the subsystem associated with the error-correcting code. Again, by purifying and by adding the reference system to S, we can assume that the state is given by $|\phi_0\rangle_S$. The quantum operation $\mathcal{A}$ can be assumed to arise from a unitary evolution U applied to $|\phi_0\rangle_S|0\rangle_A$, A being an ancillary system. Let $|\phi\rangle = U|\phi_0\rangle_S|0\rangle_A$ and consider the subsequent interaction with the environment over time t. By slicing t into intervals of duration t/n, the overall evolution up to time t can be written as ($\hbar = 1$)

$$\lim_{n\to\infty}\prod_{k=1}^{n}\delta U_k^{(S)}\delta U_k^{(B)}|\phi\rangle|\psi\rangle_B\,, \qquad (7)$$

where $\delta U_k^{(S)}$, $\delta U_k^{(B)}$ denote the unitary evolutions during the kth interval due to J and to the environment's internal Hamiltonian, respectively. It suffices to consider a first-order expansion $\delta U_k^{(S)} = I - iJ(t/n) + O[(t/n)^2]$. The elements contributing noise all involve at least $e + 1$ factors of J. By distributing some of the sums $I - iJt/n$ starting at the first time interval, the expression inside the limit can be thought of as a sum over the branches of a binary tree of products of operators associated with the edges and nodes of the tree. The root node is labeled $\delta U_1^{(B)}$, and its two edges by I and $-iJt/n$, respectively. The two children are labeled by $\delta U_2^{(B)}$, their descendant edges by I and $-iJt/n$ and so on. We choose to terminate a branch at a point where there are $e + 1$ factors of $-iJt/n$ on its path and label the leaf with the remaining product of unitary operators. The total error is estimated by summing the error amplitudes associated with the products along each of these terminated branches. A counting argument shows that there are $\binom{n}{e+1}$ such branches. Recalling that unitary operators preserve the amplitude, the error of each such branch is bounded by $(\lambda t/n)^{e+1}$. Hence, the error amplitude is at most $\binom{n}{e+1}(\lambda t/n)^{e+1} \leq (\lambda t)^{e+1}/(e + 1)!$. ■

Markovian noise.—When the noise is to a good approximation Markovian, the state of the system evolves as $\rho \mapsto \rho_t = \lim_{n\to\infty}\mathcal{L}_{t/n}^n(\rho)$, where the superoperator $\mathcal{L}_{t/n}$ takes the form $\mathcal{L}_{t/n}(\rho) = \rho + (t/n)(\sum_i L_i\rho L_i^\dagger + V\rho + \rho V^\dagger) + O[(t/n)^2]$ for appropriate operators L_i and V [16]. Our techniques apply with $\mathcal{J}_1$ given by the linear span of I, these operators, and their Hermitian transposes. By suitably modifying the proof of Theorem 2, one can show that the same bound holds for an e-error-correcting (c) code with Markovian noise provided one replaces error amplitude with error probability and redefines λ as $\lambda = 2|V| + |L_1|^2 + |L_2|^2 + \ldots$. Since error probability is what is commonly used in the description of noise processes, this further connects our formulation of error correction to the usual one.

Conclusion.—By incorporating the description of the error process within a general algebraic setting, we showed how to reformulate quantum error correction without restricting the statistical properties of the environmental noise. The existence of large codes was established for both classical and quantum information, opening the way to accurate quantum computations in the presence of arbitrary errors. In addition to substantially strengthening the power of quantum error-correction theory, our analysis points to the notion of a noiseless subsystem as a unifying framework for quantum information protection. Full exploitation of the above concept might prove fruitful in the general context of quantum information processing.

We thank Asher Peres for the three qubit example of a noiseless subsystem. E. K. and R. L. received support from the Department of Energy, under Contract No. W-7405-ENG-36, from the NSA, and from the Newton Institute. L. V. was supported in part by DARPA/ARO under the QUIC initiative.

*Electronic address: knill@lanl.gov
†Electronic address: laflamme@lanl.gov
‡Electronic address: vlorenza@mit.edu

[1] P. W. Shor, in *Proceedings of the Symposium on the Foundations of Computer Science* (IEEE Press, Los Alamitos, CA, 1996), p. 56.

[2] D. Aharonov and M. Ben-Or, in *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing* (ACM Press, New York, 1996), p. 176.

[3] A. Yu. Kitaev, Usp. Mat. Nauk. **52**, 53 (1997) [Russ. Math. Survey **52**, 1191 (1997)].

[4] E. Knill, R. Laflamme, and W. H. Zurek, Science **279**, 342 (1998).

[5] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[6] A. Steane, Proc. R. Soc. London A **452**, 2551 (1996).

[7] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[8] P. Zanardi and M. Rasetti, Phys. Rev. Lett. **79**, 3306 (1997).

[9] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).

[10] The algebra $\mathcal{J}$ with the increasing-chain filtration $\mathcal{J}_d \subseteq \mathcal{J}_{d+1}$ defines a *filtered algebra* on $\mathcal{S}$; see W. Magnus *et al.*, *Combinatorial Group Theory* (Dover, New York, 1976).

[11] C. H. Bennett *et al.*, Phys. Rev. A **54**, 3824 (1996).

[12] M. A. Nielsen *et al.*, Proc. R. Soc. London A **454**, 277 (1998).

[13] H. Tverberg, J. London Math. Soc. **41**, 123 (1966).

[14] C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras* (Interscience Publishers, Wiley & Sons, 1962).

[15] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[16] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications* (Springer-Verlag, Berlin, 1987).