ARTICLES

Bosonic quantum codes for amplitude damping

Isaac L. Chuang,[1,2] Debbie W. Leung,[1] and Yoshihisa Yamamoto[1]
[1]*ERATO Quantum Fluctuation Project, Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305*
[2]*Institute for Theoretical Physics, University of California, Santa Barbara, California 93106*
(Received 29 October 1996)

Traditional quantum error correction involves the redundant encoding of k quantum bits using n quantum bits to allow the detection and correction of any t bit error. The smallest general $t=1$ code requires $n=5$ for $k=1$. However, the dominant error process in a physical system is often well known, thus inviting the following question: Given a specific error model, can more efficient codes be devised? We demonstrate alternative codes that correct just amplitude damping errors that allow, for example, a $t=1$, $k=1$ code using effectively $n=4.6$. Our scheme is based on using bosonic states of photons in a finite number of optical modes. We present necessary and sufficient conditions for the codes and describe construction algorithms, physical implementation, and performance bounds.
[S1050-2947(97)02208-7]

PACS number(s): 03.65.−w, 89.70.+c, 89.80.+h, 02.70.−c

I. INTRODUCTION

Information is classically measured in units of bits, which are deterministic two-state systems that are said to exist either as logical zero 0_L or logical one 1_L. However, such a representation is only an approximation of reality, which is described at small size scales by the laws of quantum physics. Quantum-mechanical two-state systems [quantum bits (''qubits'')] differ from the classical bit in that they may exist in a superposition of the two states, for example, as $|\psi\rangle=a|0_L\rangle+b|1_L\rangle$, where a and b are arbitrary complex coefficients that satisfy $|a|^2+|b|^2=1$. Two continuous real parameters are needed to describe the state of a qubit and, in this sense, more information is somehow carried in it than by a classical bit. Furthermore, qubits may not be cloned [1] and even more importantly they may exist in *entangled states* where, for example, two qubits only carry one qubit of *quantum information*.

Unfortunately, quantum information is (partially) lost whenever a quantum system is observed (whether deliberately or inadvertently). This *decoherence* process plays a role analogous to noise in a classical communication channel. A major advance in quantum information theory has been the discovery that quantum information can be redundantly encoded in such a manner that it may be efficiently transmitted with arbitrarily high fidelity through a decohering quantum channel. The *quantum error correction codes* [2,3] that make this possible are analogous to classical codes for binary memoryless channels. Corresponding codes for classical linear codes and Reed-Muller codes have been found [4,5].

These quantum coding schemes are based on a model for the decoherence of qubits, in which three kinds of errors can occur: bit flips ($|0\rangle\leftrightarrow|1\rangle$), phase flips ($|1\rangle\leftrightarrow-|1\rangle$), and both simultaneously. This model is general; it describes all possible decoherence mechanisms for a qubit. However, in a given physical system, the dominant decoherence process is of a specific nature that may admit a simpler description. For example, in phase damping, no bit flips occur. The following question therefore arises: Given a particular decoherence process, what is the optimal quantum error correction scheme?

We do not yet know how to handle this general problem. However, in this paper we report on progress towards a solution, by demonstrating a different class of quantum error correction codes that correct only one particular decoherence process known as amplitude damping. Our approach is similar in philosophy to that of [6], but in contrast to that and other previous schemes, instead of qubits, which live in a two-dimensional Hilbert space, we utilize bosonic systems that occupy the Hilbert space $|0\rangle\cdots|N\rangle$. We are unaware of any classical analog to our codes. We present possible physical implementations of our scheme and conclude with a comparison with existing binary codes.

II. AMPLITUDE DAMPING MODEL

Noise is a fundamental process that accompanies the dynamics of any open system. Traditionally, the dynamics of an open quantum system are described by a ''master equation.'' We begin with an outline derivation of this formalism for a particular noise process that is physically important: amplitude damping. We then turn to an alternate description of the same noise process, using the formalism of quantum operations [7], which will later prove to be useful in understanding and deriving our alternative quantum codes.

A master equation is generally derived in the following manner. The system a, described by $\rho(t)$, couples to an environment b, described by ρ_b, through an interaction Hamiltonian H_I. Evolution generates an entangled state of the total system, causing quantum information originally in the system alone to dissipate into the environment. Tracing over the environmental degrees of freedom gives the reduced matrix

for the system alone, and in the Born and Markov approximations (the interaction only weakly perturbs the state of the environment, which is furthermore memoryless) the Schrödinger equation for the system state has the form [8,9]

$$\dot{\rho}(t) = -\frac{1}{\hbar^2}\int_0^{\tau_c} dt' \mathrm{Tr}_b[H_I(t),[H_I(t-t'),\rho(t)\otimes\rho_b]] , \tag{2.1}$$

where the operators are in given in the interaction picture and τ_c is the correlation time of the environment.

One of the simplest possible interactions between a system and an environment is a bilinear coupling, a product of the elementary system and environment coordinate operators. The effects of this kind of noise are conveniently studied by modeling the system as a simple harmonic oscillator (we shall dispense with self-Hamiltonians in the following analysis to avoid unnecessary complications). For example, this model provides a good description of the scattering of photons from a single-mode optical fiber to the outside. The interaction Hamiltonian in the Schrödinger picture is

$$H_I = \chi(a^\dagger b + b^\dagger a). \tag{2.2}$$

where a,b are the annihilation operators of the system and the environment, respectively (using a single-mode model for the environment, which is sufficient to capture the dynamics of interest), and χ is a coupling constant. When τ_c is much smaller than the time scales in which $\rho(t)$ or $H_I(t)$ change significantly, Eq. (2.1) can be approximated to give

$$\dot{\rho}(t) = -\eta\frac{\tau_c}{\hbar^2}\mathrm{Tr}_b[H_I,[H_I,\rho(t)\otimes\rho_b]] , \tag{2.3}$$

where η is a prefactor resulting from the integration. Substituting Eq. (2.2) gives an equation of motion for the system density matrix

$$\dot{\rho} = -\frac{\lambda}{2}(a^\dagger a\rho + \rho a^\dagger a - 2a\rho a^\dagger) , \tag{2.4}$$

where $\lambda = 2\eta\chi^2\tau_c$. We have set $\langle b^\dagger b\rangle = 0$ in Eq. (2.4) to reflect an environment at temperature kT much smaller than the system's energy scale $\hbar\omega$. This master equation describes the gradual loss of energy from the system to a zero-temperature environment and is known as *amplitude damping* [8,9].

Often, because of experimental reality (we put photons in one end of the fiber and observe the output at the other end), what one is interested in is the state change between two definite times rather than in the continuous evolution behavior. In this context, we may use an alternative and equivalent formalism for quantum noise. Mathematically, the evolution of a density matrix between times t and $t+\Delta t$ due to a particular process may be described as a linear transformation from one density matrix ρ to another ρ'. This may be expressed in the operator sum representation [10] as

$$\rho' = \sum_k A_k \rho A_k^\dagger , \tag{2.5}$$

where A_k are linear operators, sometimes referred to as ''Kraus effects [7],'' which are related to the Lindblad operators appearing in Eq. (2.4). In this equation, k denotes different possible final states of the environment; modeling the environment as a single simple harmonic oscillator, using the number state basis $|\ \rangle_b$ (or, equivalently, choosing k to be the number of photons lost from the system), and taking the environment to initially be in the ground state gives the A_k operators

$$A_k = {}_b\langle k| e^{i\chi\Delta t(a^\dagger b + b^\dagger a)}|0\rangle_b , \tag{2.6}$$

where $\gamma = 1-\cos^2(\chi\Delta t)$ is the probability of losing a single photon from the system during time Δt. Operator algebra techniques [8] can be used to evaluate explicitly the inner product, giving

$$A_k = \sum_n \sqrt{\binom{n}{k}}\ \sqrt{(1-\gamma)^{n-k}\gamma^k}\ |n-k\rangle\langle n| . \tag{2.7}$$

Note that the same result can be obtained by direct integration of the master equation (2.4) and manipulation of the final answer into the form of Eq. (2.5). One must then identify the damping rate as $\gamma = 1-e^{-\lambda\Delta t}$ (this is reasonable; we expect the damping to be exponential with time). The functional dependence of γ on Δt is slightly different in the two derivations because we have used a phenomenological single-mode model of the environment in our outline derivation.

If the initial state of the system is pure, it may be written as $\rho = |\psi\rangle\langle\psi|$. The final state ρ' may be elegantly described as an explicit mixture of pure states given by

$$[\,\psi'\rangle = \bigoplus_{k=0}^{N} A_k|\psi\rangle , \tag{2.8}$$

where N is the maximum occupation number of a single bosonic mode. Here the ''$\oplus$'' symbol represents a tensor sum of states and $[\,\psi'\rangle$ is a convenient shorthand used to denote a mixed state, as distinguished from a pure state $|\psi\rangle$. In other words,

$$\rho' = [\,\psi'\rangle\langle\psi'\,] = \sum_{k=0}^{N} A_k|\psi\rangle\langle\psi|A_k^\dagger . \tag{2.9}$$

The mixed state $[\,\psi'\rangle$ is a tensor sum of $N+1$ (unnormalized) pure states that describe the $N+1$ possible final states of the system; one may interpret these as noninterfering ''alternative histories'' [11]. The normalization of each pure state gives its probability of occurrence. As previously mentioned, k describes the number of photons lost to the environment. It is important that even when no photons are lost to the environment, then the state of the system is changed.

So far, we have described the effect of amplitude damping on a single-mode system. Consider now a system with m modes and let us use A_{kj} to denote the action of the effect A_k on the jth mode of a state $j\in[1,m]$. After amplitude damping, the initial pure state

$$|\psi_{in}\rangle = |n_1\cdots n_m\rangle \tag{2.10}$$

becomes the mixed state

$$[\psi_{out}\rangle=\left[\bigoplus_{k=0}^{N}A_{k0}|n_0\rangle\right]\cdots\left[\bigoplus_{k=0}^{N}A_{km}|n_m\rangle\right], \quad (2.11)$$

where there are now $(N+1)^m$ possible final states. It is convenient to use the shorthand notation

$$A_{\tilde{k}}=A_{k_0 0}\cdots A_{k_m m}, \quad (2.12)$$

where k_j is the jth digit of the number $\tilde{k}$ written in base $N+1$, so that we may rewrite Eq. (2.11) as

$$[\psi_{out}\rangle=\bigoplus_{\tilde{k}=0}^{(N+1)^m-1}A_{\tilde{k}}|\psi_{in}\rangle . \quad (2.13)$$

Note that identical states in a tensor sum can be combined using the rule

$$a[\psi\rangle\oplus b[\psi\rangle=\sqrt{|a|^2+|b|^2}[\psi\rangle , \quad (2.14)$$

since an overall phase does not matter (assuming no entanglement with other systems).

As an example, amplitude damping of the state

$$|\psi_{in}\rangle=a|01\rangle+b|10\rangle \quad (2.15)$$

gives, using

$$A_0=|0\rangle\langle 0|+\sqrt{1-\gamma}|1\rangle\langle 1|, \quad (2.16)$$

$$A_1=\sqrt{\gamma}|0\rangle\langle 1| , \quad (2.17)$$

the output state

$$[\psi_{out}\rangle=A_{00}|\psi_{in}\rangle\oplus A_{01}|\psi_{in}\rangle\oplus A_{10}|\psi_{in}\rangle\oplus A_{11}|\psi_{in}\rangle \quad (2.18)$$

$$=\sqrt{1-\gamma}|\psi_{in}\rangle\oplus\sqrt{\gamma}|00\rangle . \quad (2.19)$$

This result can be understood intuitively: The original state only contains a single photon and thus, whenever it is lost, the final state must be the vacuum. This example indicates that the state of Eq. (2.15) is useful for detection of a single-photon loss. However, no useful information about a and b can be extracted from the vacuum state and so it is not useful for error *correction*.

III. EXAMPLES

Let us motivate the remainder of this paper by considering the following example: We *encode* the logical zero and one states of a single qubit as

$$|0_L\rangle=\left[\frac{|40\rangle+|04\rangle}{\sqrt{2}}\right] \quad |1_L\rangle=|22\rangle , \quad (3.1)$$

such that the initial state is the arbitrary qubit

$$|\psi_{in}\rangle=a|0_L\rangle+b|1_L\rangle . \quad (3.2)$$

The possible outcomes after amplitude damping may be written as

$$[\psi_{out}\rangle=\bigoplus_{\tilde{k}}|\phi_{\tilde{k}}\rangle=\bigoplus_{\tilde{k}}A_{\tilde{k}}|\psi_{in}\rangle , \quad (3.3)$$

where we shall express $\tilde{k}$ as a base 5 numeral and $|\phi_{\tilde{k}}\rangle$ is an unnormalized pure state (the norm of which gives its probability for occurring in the mixture). For small loss probability γ, the most likely final state will be

$$|\phi_{00}\rangle=(1-\gamma)^2|\psi_{in}\rangle , \quad (3.4)$$

corresponding to no quanta being lost to the bath. The next most likely states result from the loss of a single photon:

$$|\phi_{01}\rangle=\sqrt{2\gamma}(1-\gamma)^{3/2}[a|03\rangle+b|21\rangle], \quad (3.5)$$

$$|\phi_{10}\rangle=\sqrt{2\gamma}(1-\gamma)^{3/2}[a|30\rangle+b|12\rangle] . \quad (3.6)$$

States resulting from the loss of more than one quantum occur with probabilities of order γ^2. Therefore, we limit our correction scheme to errors losing at most one quantum. Each such error E_i takes $|0_L\rangle$ and $|1_L\rangle$ to states $|0_L\rangle_i$ and $|1_L\rangle_i$, respectively. The key is that $|0_L\rangle$, $|1_L\rangle$, $|0_L\rangle_i$, and $|1_L\rangle_i$ $\forall i$ are mutually orthogonal and so are $|\phi_{00}\rangle$, $|\phi_{01}\rangle$, and $|\phi_{10}\rangle$. In principle, a (''quantum-nondemolition'') measurement scheme can detect all error syndromes. Furthermore, for each i, the norms of $|0_L\rangle_i$ and $|1_L\rangle_i$ are equal. After detecting an error syndrome, one can apply an appropriate unitary transformation converting $|0_L\rangle_i$ and $|1_L\rangle_i$ to $|0_L\rangle$ and $|1_L\rangle$, respectively. This makes possible the correction

$$a|0_L\rangle_i+b|1_L\rangle_i\rightarrow\alpha[a|0_L\rangle+b|1_L\rangle] , \quad (3.7)$$

where α is independent of a,b. Note that this is done without any information about a,b and without diminishing the amplitude of the erroneous state. For this particular code, the output state has fidelity [12,13] [see also Eq. (7.4)] $\mathcal{F}=1-6\gamma^2$ with respect to the input.

As a comparison, consider the code

$$|0_L\rangle=|11\rangle, \quad |1_L\rangle=|22\rangle, \quad (3.8)$$

with the most probable state

$$|\phi_{00}\rangle=a(1-\gamma)|11\rangle+b(1-\gamma)^2|22\rangle . \quad (3.9)$$

No unitary transformation will bring it back to

$$a|11\rangle+b|22\rangle \quad (3.10)$$

unless a,b are predetermined (a nonunitary transformation can revert the change, but it will reduce the fidelity of the correction process by an amount first order in γ.

In the remainder of the paper, we shall describe the criteria for a scheme in which k qubits may be encoded so that loss up to t quanta may be corrected. For small t, a scheme will be exhibited.

IV. CODE CRITERIA

Quantum error correction is just the reversing of some effect due to decoherence. General criteria for this to be possible have been given in the literature [13–16]. In this particular case, we may express the required conditions in the following manner. Let $\{|c_0\rangle\cdots|c_l\rangle\cdots|c_{l_o}\rangle\}$ be l_o+1 code-

words that encode orthogonal logical states within the m-mode Hilbert space with maximum total photon number N and define $\mathcal{K}(t)$ as the set of all m-digit base $N+1$ numbers whose digits sum to t (corresponding to t errors). The logical states must satisfy

$$\langle c_{l_1}|A_{\tilde{k}}^{\dagger}A_{\tilde{k}'}|c_{l_2}\rangle=0 \quad \text{for } l_1\neq l_2 \quad \text{or} \quad \tilde{k}\neq\tilde{k}', \tag{4.1}$$

$$\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle=g_{\tilde{k}} \quad \forall l \tag{4.2}$$

for all $\tilde{k},\tilde{k}'\in\cup_{s\leq t}\mathcal{K}(s)$. Here $g_{\tilde{k}}$ is some constant that depends only on $\tilde{k}$. The first condition requires that all erroneous states be orthogonal and the second requires that the encoded Hilbert space not be deformed. Here we present an explicit statement of these two conditions as algebraic conditions on the code construction.

We consider first a code word $|c_l\rangle$, which may be expressed as an equally weighted sum of N_l energy eigenstates $|n_1\ldots n_m\rangle$ [we shall borrow from [17] the name *quasiclassical states* (QCSs) because these states resemble the classical code words; QCSs do not refer to coherent states in this paper]. When all the QCSs are equally weighted, we call the code ''balanced.'' Otherwise, the code is referred to as ''unbalanced.'' Each code word can be represented by a matrix with m columns and N_l rows, each row being one of the QCSs in the code word. For instance, if

$$|c_l\rangle=\frac{1}{\sqrt{N_l}}[|n_{11}\cdots n_{1m}\rangle+\cdots+|n_{N_l1}\cdots n_{N_lm}\rangle] \ , \tag{4.3}$$

then the corresponding matrix $\mathcal{M}_l$ is

$$\begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1m} \\ n_{21} & n_{22} & \cdots & n_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ n_{N_l1} & n_{N_l2} & \cdots & n_{N_lm} \end{bmatrix} . \tag{4.4}$$

For $t=0$ errors, we have $\mathcal{K}(0)=\{0\}$ and $A_0|n_{i1}\cdots n_{im}\rangle=(1-\gamma)^{\mathcal{R}_i/2}|n_{i1}\cdots n_{im}\rangle$, where the row sum $\mathcal{R}_i=\Sigma_{j=1}^{m}n_{ij}$. Criteria given by Eq. (4.2) require that the amplitudes of $A_0|c_l\rangle$ be the same for all $|c_l\rangle$, that is,

$$\frac{1}{N_l}\sum_{i=1}^{N_l}(1-\gamma)^{\mathcal{R}_i/2} \tag{4.5}$$

be the same for all $|c_l\rangle$. A *sufficient* condition for this is the equality of all the $\mathcal{R}_i$ for all i and for all code words $|c_l\rangle$. Alternatively, one may say that the sum of any row from any $\mathcal{M}_l$ equals N, the same total photon number in all the QCSs. Denote the set of QCSs with m modes and total photon number N as $\mathcal{Q}(N,m)$. It follows that if we construct all the code words from states in $\mathcal{Q}(N,m)$, then the nondeformation constraint Eq. (4.2) is satisfied for $t=0$. Physically, this requirement stems from the fact that a state with higher number of quanta decays faster. To preserve *a posteriori* probabilities of each code word, we must encode them in a subspace in which the decay probabilities are equal for all of them.

Similarly, for $t=1$ errors, we have $\mathcal{K}(1)=\{0\cdots1,0\cdots10,\ldots,1\cdots0\}$ and, for example,

$$A_{0\ldots1}|c_l\rangle=A_{0\ldots1}\frac{1}{\sqrt{N_l}}\sum_{i=1}^{N_l}|n_{i1}\cdots n_{im}\rangle \tag{4.6}$$

$$=\sum_{i=1}^{N_l}\sqrt{\frac{n_{im}\gamma(1-\gamma)^{N-1}}{N_l}}|n_{i1}\cdots n_{im}-1\rangle \tag{4.7}$$

and

$$\langle c_l|A_{0\ldots1}^{\dagger}A_{0\ldots1}|c_l\rangle=\frac{\gamma(1-\gamma)^{N-1}}{N_l}\sum_{i=1}^{N_l}n_{im}. \tag{4.8}$$

The nondeformation criteria require the above sum to be the same for all code words. Equivalently, the column sum of the mth column of each code word divided by N_l has to be independent of the code word. Similar expressions for other $A_{\tilde{k}}$ give rise to similar criteria for each column separately. We therefore have the following.

Lemma 1. Let each code word be expressed as an m column, N_l row matrix with elements n_{ij}. If we choose code words such that $\Sigma_i n_{ij}/N_l=y_j$ for all $|c_l\rangle$, then $\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle=g_{\tilde{k}}\ \forall\tilde{k}\in\mathcal{K}(1)$.

Proof. This is proved above. ■

These criteria correspond to certain symmetry requirements among the various code words. For $t=2$, $\mathcal{K}(2)=\{0\cdots02,0\cdots20,\ldots,20\cdots0,0\cdots11,0\cdots110,\ldots,11\cdots0\}$. Working out $A_{\tilde{k}}|c_l\rangle$ for each $\tilde{k}$ and applying the criteria for $t=0,1,2$, one arrives at the following.

Lemma 2. We adopt the same notations as in Lemma 1. Let us choose code words that satisfy the nondeformation criteria for $t=1$ and such that $\Sigma_i n_{ij_1}n_{ij_2}/N_l=y_{j_1,j_2}$ for all $|c_l\rangle$, where $j_1,j_2\in[1,N_l]$ and j_1,j_2 may or may not be distinct. Then $\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle=g_{\tilde{k}}\ \forall\tilde{k}\in\mathcal{K}(2)$.

Proof. See Appendix A. ■

The generalization to arbitrary t is as follows.

Theorem 1. Let each code word be expressed as an m column, N_l row matrix with elements n_{ij}. If we choose code words such that $\Sigma_i n_{ij_1}n_{ij_2}\cdots n_{ij_t}/N_l=y_{j_1,j_2,\ldots,j_s}$ independently of $|c_l\rangle\ \forall l$, $\forall(j_1,j_2,\ldots,j_s)\in[1,N_l]^s$, and $\forall s\in[1,t]$, then $\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle=g_{\tilde{k}}\ \forall\tilde{k}\in\cup_{s=1}^{t}\mathcal{K}(s)$, $\forall l$.

Proof. See Appendix A. ■

The above theorem can be generalized to unbalanced codes in which code words are unequally weighted superpositions of QCSs. If the amplitudes of the QCSs in $|c_l\rangle$ are $(\sqrt{\mu_1},\sqrt{\mu_2},\ldots,\sqrt{\mu_{N_l}})$, we replace the sum $\Sigma_i n_{ij_1}n_{ij_2}\cdots n_{ij_t}/N_l$ by $\Sigma_i\mu_i n_{ij_1}n_{ij_2}\cdots n_{ij_t}$, i.e., we replace the equal weights $1/N_l$ by the μ_i's (the derivation of Theorem 1 in the unbalanced case is a straightforward generalization of the balanced case and we will skip the proof).

As t increases, the nondeformation criteria become very restrictive. We have found unbalanced codes by a numerical search correcting up to $t\leq4$ (Sec. VIII), which have no analogs in the balanced codes. On the other hand, for $t\leq2$, we

found simple construction algorithms for balanced codes with no apparent counterparts for the unbalanced codes.

It should also be noted that the $t=0$ nondeformation criterion, that row sums (total number of excitations in each QCS) be equal for all rows and for all matrices (code words), is not a necessary condition. An example is Shor's nine-bit code [2]

$$|0_L\rangle=(|000\rangle+|111\rangle)^{\otimes 3}, \tag{4.9}$$

$$|1_L\rangle=(|000\rangle-|111\rangle)^{\otimes 3}\ . \tag{4.10}$$

The QCSs have different numbers of 1's, but Eq. (4.5) is equal for $|0_L\rangle$ and $|1_L\rangle$. However, code criteria for $t>1$ will be extremely complicated when row sums are different, and treatment of such codes are outside the scope of the present discussion.

The other criteria, the orthogonality constraints given by Eq. (4.1), can be satisfied as follows. Let $|u\rangle=|u_1\cdots u_m\rangle$ and $|v\rangle=|v_1\cdots v_m\rangle$ be two states in $\mathcal{Q}(N,m)$. We define the *distance* between u and v as

$$\mathcal{D}(u,v)=\frac{1}{2}\sum_i\ |u_i-v_i|\ . \tag{4.11}$$

Clearly, $0\leqslant\mathcal{D}\leqslant N$. Moreover, $\mathcal{D}(u,v)=\mathcal{D}(v,u)$, $\mathcal{D}(u,u)=0$, and

$$\mathcal{D}(u,v)+\mathcal{D}(v,w)=\frac{1}{2}\sum_i\ |u_i-v_i|+|v_i-w_i| \tag{4.12}$$

$$\geqslant\frac{1}{2}\sum_i\ |u_i-w_i| \tag{4.13}$$

$$=\mathcal{D}(u,w)\ . \tag{4.14}$$

Thus $\mathcal{D}$ is a metric on the discrete space $\mathcal{Q}(N,m)$. (For binary states, this is half of the Hamming distance.) Define the distance between two code words $|c_1\rangle$ and $|c_2\rangle$ to be the minimum of $\mathcal{D}(u_1,u_2)$ with $|u_1\rangle,|u_2\rangle$ being QCSs in $|c_1\rangle$ and $|c_2\rangle$, respectively. For code words with non-negative amplitudes of the constituent QCSs, two code words are orthogonal if and only if their distance is nonzero. We therefore have the following.

Theorem 2: Let $|c_1\rangle$ and $|c_2\rangle$ be two code words formed from states in $\mathcal{Q}_1$ and $\mathcal{Q}_2$ respectively, where $\mathcal{Q}_1,\mathcal{Q}_2\subset\mathcal{Q}(N,m)$ and $\mathcal{D}(u_1,u_2)>t\ \ \forall u_1\in\mathcal{Q}_1,u_2\in\mathcal{Q}_2$. Then $\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}'}|c_2\rangle=0\ \ \forall\tilde{k},\tilde{k}'\in\cup_{s\leqslant t}\mathcal{K}(s)$.

Proof. Let $A_{\tilde{k}'}|c_1\rangle=|d_1\rangle$, $A_{\tilde{k}}|c_2\rangle=|d_2\rangle$, and $|v_1\rangle,|v_2\rangle$ be QCSs in $|d_1\rangle,|d_2\rangle$, respectively, such that $\mathcal{D}(d_1,d_2)=\mathcal{D}(v_1,v_2)$. Let $|u_1\rangle$, $|u_2\rangle$ be the original QCSs in $|c_1\rangle$, $|c_2\rangle$ before the error. Then $\mathcal{D}(u_1,v_1)=\mathcal{D}(u_2,v_2)\leqslant t/2$ and $\mathcal{D}(v_1,v_2)+\mathcal{D}(u_1,v_1)+\mathcal{D}(u_2,v_2)\geqslant\mathcal{D}(u_1,u_2)>t$. Hence $\mathcal{D}(v_1,v_2)>0$ and $\mathcal{D}(d_1,d_2)>0$. Therefore, $|d_1\rangle$ and $|d_2\rangle$ are orthogonal states. ■

In other words, by forming code words using QCSs that are sufficiently far apart, then the orthogonality conditions are easily satisfied.

V. EXISTENCE OF CODES

How large must N, m, and N_l be to satisfy both the nondeformation constraint, (4.2) and the orthogonality constraint, (4.1)? We now show that an unbalanced code exists for arbitrarily large t if N is allowed to be arbitrarily large and give an upper bound for the required N.

Let $|c_0\rangle, |c_1\rangle, \ldots, |c_l\rangle, \ldots, |c_{l_o}\rangle$, be l_o+1 code words, each being an unequally weighted superposition of N_l QCSs in $\mathcal{Q}(n,m)$. For convenience, define

$$\mathcal{P}(n,m)=\mathcal{C}(n+m-1,m-1)\equiv\frac{(n+m-1)!}{n!(m-1)!} \tag{5.1}$$

as the number of all possible partitions of the integer n, i.e., the number of x_i such that $x_1+x_2+\cdots+x_m=n$ [18]. Then if we choose $N=nd$ such that

$$\mathcal{P}(n,m)\geqslant N_1+N_2+\cdots+N_{l_o}=N_{QCS}, \tag{5.2}$$

where N_{QCS} is the total number of QCS in the code words, then by Theorem 2 all the QCSs involved can be chosen to be distinct and multiplication of the number states by $d=t+1$ allows the orthogonality condition to be satisfied.

On the other hand, the nondeformation condition involves satisfying a certain number of constraint equations, given by the total number of possible errors times l_o. The number of errors involving losing s photons from m modes is just the number of partitions of s into m parts $\mathcal{P}(s,m)$. Take the QCSs to be arbitrary and solve the nondeformation constraint equations (of Theorem 1, generalized to include unbalanced codes) as linear equations for the weights of the QCSs. As long as the number of variables (N_{QCS}) is no fewer than the number of equations, solutions always exist. We may also augment the system of equations by l_o+1 equations to ensure the correct normalization of each code word. Hence, for N that satisfy

$$1+l_o+l_o\sum_{s=0}^{t}\mathcal{P}(s,m)\leqslant N_{QCS}\leqslant\mathcal{P}(N/(t+1),m)\ , \tag{5.3}$$

codes with m modes correcting t errors exist.

We simplify Eq. (5.3) by writing explicitly the expression for $\mathcal{P}(N/(t+1),m)$ and $\mathcal{P}(s,m)$ and performing the summation (by writing the summand as a telescopic sum). We obtain

$$m!(1+l_o)+l_o\left[\frac{(t+m)!}{t!}\right]\leqslant m\left[\frac{\left(\dfrac{N}{t+1}+m-1\right)!}{\dfrac{N}{t+1}!}\right]\ . \tag{5.4}$$

For example, when $m=2$, Eq. (5.4) becomes

$$\frac{l_o(t+2)(t+1)}{2}+l_0+1\leqslant\frac{N}{t+1}+1\ , \tag{5.5}$$

which gives a scaling law $N\approx t^3l_o/2$. The scaling of N as a function of t for arbitrary but fixed m can be obtained by approximating the factorials involving N and t in Eq. (5.4)

using the Stirling approximation. We found that $N \approx (l_o/em)^{1/(m-1)} t^{(2m-1)/(m-1)}$. We have also assumed $N/(t+1)$ to be large in obtaining the scaling law, and this is a consistent assumption. Note that this upper bound is generally much larger than necessary, as can be seen in the examples for $t=3$ or $t=4$. Much more efficient codes may be obtained because the QCSs may be chosen to give redundant constraint equations. This may be accomplished either systematically (next section) or by numerical search (Sec. VIII).

VI. CONSTRUCTION ALGORITHM FOR $T \leqslant 2$ BALANCED CODES

So far, we have established criteria for and proved the existence of bosonic codes for amplitude damping. We now present an explicit procedure that obtains a class of balanced codes to correct for $t=1$ and $t=2$ errors.

To correct the $t=1$ error, consider ordered m-tuples $(x_1, x_2, \ldots, x_m)$ such that $x_1 + x_2 + \cdots + x_m = n$. We will use the same symbol $\mathcal{Q}(n,m)$ for the space of all such m-tuples as well as the space of all QCSs $\{|x_1 x_2 \cdots x_m\rangle\}$. We define an operation $\mathcal{T}$ on $\mathcal{Q}(n,m)$ that takes $(x_1, x_2, \ldots, x_m)$ to $(x_2, \ldots, x_m, x_1)$, i.e., the symbols are cyclically permuted. Define the order of an element to be the size of its orbit under $\mathcal{T}$. It follows that the order p must divide m and let $m=pq$. An element of order p looks like $(x_1, \ldots, x_p, x_1, \ldots, x_p, \ldots, x_1, \ldots, x_p)$ with the string $(x_1, \ldots, x_p)$ repeated q times. The orbit looks like

$$(x_1, x_2, \ldots, x_p, \ldots, x_1, x_2, \ldots, x_p), \tag{6.1}$$

$$(x_2, x_3, \ldots, x_1, \ldots, x_2, x_3, \ldots, x_1), \tag{6.2}$$

$$\ldots, \tag{6.3}$$

$$(x_p, x_1, \ldots, x_{p-1}, \ldots, x_p, x_1, \ldots, x_{p-1}) \, . \tag{6.4}$$

We form states by taking equal-weight superposition of QCSs in each orbit:

$$|c\rangle = \frac{1}{\sqrt{p}} (|x_1 x_2 \cdots x_p \cdots x_1 x_2 \cdots x_p\rangle \tag{6.5}$$

$$+ |x_2 x_3 \cdots x_1 \cdots x_2 x_3 \cdots x_1\rangle \tag{6.6}$$

$$+ \cdots \tag{6.7}$$

$$+ |x_p x_1 \cdots x_{p-1} \cdots x_p x_1 \cdots x_{p-1}\rangle) \, . \tag{6.8}$$

States formed by distinct orbits are orthogonal, as the orbits partition $\mathcal{Q}(n,m)$. Furthermore, we multiply each number in the QCSs by d. The minimal separation of distinct QCSs will be at least d since distances come as multiples of d only. Hence all the code words will remain orthogonal after errors of $t<d$ occur. Code words are now in the form

$$|c\rangle = \frac{1}{\sqrt{p}} (|dx_1 dx_2 \cdots dx_p \cdots dx_1 dx_2 \cdots dx_p\rangle \tag{6.9}$$

$$+ |dx_2 dx_3 \cdots dx_1 \cdots dx_2 dx_3 \cdots dx_1\rangle \tag{6.10}$$

$$+ \cdots \tag{6.11}$$

$$+ |dx_p dx_1 \cdots dx_{p-1} \cdots dx_p dx_1 \cdots dx_{p-1}\rangle) \, . \tag{6.12}$$

For the nondeformation criteria, the row sum is $nd=N$ by construction. The column sum divided by the normalization factor squared is (by Lemma 1)

$$\frac{dx_1 + \cdots + dx_p}{p} = \frac{dn}{m} \tag{6.13}$$

in any code word, independent of the order of the constituent QCSs. Codes in examples (1)–(3) in Sec. VIII are constructed in this way.

To correct for $t=2$ errors, the $t=1$ criteria have to be satisfied as well. We take a subset of the $t=1$ code words that will survive the extra nondeformation criteria for $t=2$. We also replace $d=2$ by $d \geqslant 3$. For $m>2$, pairs of code words in the form

$$|0_L\rangle = \frac{1}{\sqrt{m}} (|dx_1 dx_2 \cdots dx_m\rangle \tag{6.14}$$

$$+ |dx_2 dx_3 \cdots dx_1\rangle \tag{6.15}$$

$$+ \cdots \tag{6.16}$$

$$+ |dx_m dx_1 \cdots dx_{m-1}\rangle), \tag{6.17}$$

$$|1_L\rangle = \frac{1}{\sqrt{m}} (|dx_m \cdots dx_2 dx_1\rangle \tag{6.18}$$

$$+ |dx_{m-1} \cdots dx_1 dx_m\rangle \tag{6.19}$$

$$+ \cdots \tag{6.20}$$

$$+ |dx_1 dx_m \cdots dx_2\rangle) \tag{6.21}$$

will always satisfy the nondeformation criteria for $t=2$ (proof omitted). Examples (4) and (5) in Sec. VIII are constructed in this way. This encodes only one qubit; we are still looking for $t=2$ codes that can encode more qubits.

For $t \geqslant 3$, we performed a numerical search for special QCSs in which the system of linear equations for the weights is linearly dependent. In the best case, the number of linear equations to be solved can be much reduced. Therefore, we can find code words involving fewer QCSs, fewer modes, and fewer quanta. Although encoding is certainly possible with a much smaller Hilbert space, we have not found a systematic way to generate such QCSs. Codes correcting $t \leqslant 4$ errors are exhibited in Sec. VIII.

VII. RATES AND FIDELITIES

The performance of these bosonic quantum codes can be characterized by their *rate*, the number of qubits communicated per qubit transmitted, and by their *fidelity*, the worst-

case qubit degradation after decoding and correction. We discuss these two measures here.

The rate r is given by the ratio of the number of encoded qubits to the maximum number of qubits that can be accommodated in our Hilbert space:

$$r=\frac{k}{m\log_2(nd+1)}\,, \tag{7.1}$$

where 2^k is the number of code words and $(nd+1)^m$ is the size of the Hilbert space in our code. The exact number of possible code words depends on the choice of N (the maximum number of excitations in any single mode) and m (the number of modes). For $t=1$, we have worked out a counting scheme, but omit the details here. However, the majority of the QCSs have order m. Hence, to a good approximation, the number of code words obtained is

$$2^k=\frac{\mathcal{P}(n,m)}{m}\,. \tag{7.2}$$

Thus the asymptotic rate of our codes for large n is $(m-1)/m$. For small n, code words involving fewer than m QCSs allow slightly more qubits to be encoded compared to Eq. (7.2) [see examples (1)–(3) in Sec. VIII]. This small gain can be important in applications such as key distribution in quantum cryptography. We have shown in Sec. V that for arbitrary but fixed t, $N\approx(l_o/em)^{1/(m-1)}t^{(2m-1)/(m-1)}$ is large enough to guarantee the existence of a code with l_o+1 code words. This implies a loose asymptotic lower bound to the achievable rate for our code, of $r=(m-1)/m$.

We now turn to the code fidelity $\mathcal{F}$, which we desire to know as a function of the parameters N, m, and t. The fidelity is defined as [13]

$$\mathcal{F}=\min_{\psi_{in}}\langle\psi_{in}|\rho_f|\psi_{in}\rangle\,, \tag{7.3}$$

where ρ_f is the final output state after correction. When the correction criteria are satisfied, recovery procedures exist for the correctable errors that recover the input states. Therefore, the fidelity equals the total probability that some correctable errors occur, which is given by

$$\mathcal{F}=\min_{\psi_{in}}\sum_{\tilde{k}\in\cup_{s\in[1,t]}\mathcal{K}(s)}\langle\psi_{in}|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|\psi_{in}\rangle\,. \tag{7.4}$$

Let the input state be expressed as a sum of code words $|\psi_{in}\rangle=\Sigma_l\alpha_l|c_l\rangle$. Then using the orthogonality and nondeformation conditions, we find that

$$\langle\psi_{in}|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|\psi_{in}\rangle=\left|\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle\right|\,, \tag{7.5}$$

with $|c_l\rangle$ any one of the code words on the right-hand side. Now, if we write each code word as

$$|c_l\rangle=\sqrt{\mu_1}|n_{11}n_{12}\cdots n_{1m}\rangle \tag{7.6}$$

$$+\sqrt{\mu_2}|n_{21}n_{22}\cdots n_{2m}\rangle \tag{7.7}$$

$$+\cdots \tag{7.8}$$

$$+\sqrt{\mu_{N_l}}|n_{N_l1}n_{N_l2}\cdots n_{N_lm}\rangle\,, \tag{7.9}$$

then, for $\tilde{k}=(k_1,k_2,\ldots,k_m)\in\mathcal{K}(s)$,

$$\left|\langle c_l|A_{\tilde{k}}^{\dagger}A_{\tilde{k}}|c_l\rangle\right|=(1-\gamma)^{N-s}\gamma^s\sum_{i=1}^{N_l}\mu_i\mathcal{C}(n_{ij},k_j)\,, \tag{7.10}$$

and using the relation for binomial coefficients

$$\mathcal{C}(N,s)=\sum_{\tilde{k}\in\mathcal{K}(s)}\sum_{i=1}^{N_l}\mu_i\mathcal{C}(n_{i1},k_1)\mathcal{C}(n_{i2},k_2)\cdots\mathcal{C}(n_{im},k_m)\,, \tag{7.11}$$

we find that the fidelity is

$$\mathcal{F}=\sum_{s=1}^{t}(1-\gamma)^{N-s}\gamma^s\mathcal{C}(N,s) \tag{7.12}$$

$$=1-\mathcal{C}(N,t+1)\gamma^{t+1}+O(\gamma^{t+2})\,. \tag{7.13}$$

This expression holds for balanced codes as well as unbalanced codes. The amazing feature is that given a code that satisfies the orthogonality and nondeformation constraints, $\mathcal{F}$ is independent of m; it is determined only by N and t.

One should note that although codes can be constructed to correct an arbitrary number of photon loss, the more errors one wishes to correct, the more photons are required. On the other hand, having larger photon number states means a higher probability for the system as a whole to suffer loss of quanta. These two effects compete against each other to give an upper bound on the fidelity, which we can estimate as follows. Let N be the required maximum photon number and t the total number of errors to be corrected. As previously discussed, due to the constraint equations that hold for error correction to be possible, the two parameters can be reduced to one degree of freedom, in terms of which we may estimate the optimal achievable fidelity. In terms of t, the optimum fidelity for fixed γ is obtained by setting

$$\frac{d}{dt}\ln(1-\mathcal{F})=0\,. \tag{7.14}$$

From Eq. (7.13), this gives, to first order in γ,

$$\frac{1}{\mathcal{C}}\frac{\partial\mathcal{C}}{\partial N}\frac{dN}{dt}+\frac{1}{\mathcal{C}}\frac{d\mathcal{C}}{dt}+\ln\gamma=0, \tag{7.15}$$

where $\mathcal{C}$ is a shorthand notation for $\mathcal{C}(N,t+1)$. Using the Stirling approximation for the factorials in $\mathcal{C}$, we obtain

$$\ln\left(\frac{N}{N-t-1}\right)\frac{dN}{dt}+\ln\left(\frac{N-t-1}{t+1}\right)+\ln\gamma=0\,. \tag{7.16}$$

In general, N is much larger than t, which allows further simplification of Eq. (7.16):

$$\frac{dN}{dt}\frac{t}{N}-\ln\left(\frac{t}{N}\right)+\ln\gamma=0\,. \tag{7.17}$$

The exact dependence of N on t is generally very complicated. This point can be appreciated from the explicit code examples following this section. In particular, for a fixed number of modes, the minimum number of encoding excitations needed to be able to correct for a fixed number of photons lost depends on the existence of ''good solutions'' to the equations representing the code criteria and it is analytically intractable to find the optimal fidelity for the general case. We thus have only a bound on the fidelity of a bosonic code with arbitrary QCSs. There is no theoretical bound on the number of correctable errors.

For concreteness, we will illustrate the above assuming that N asymptotically follows a power scaling law in t. As illustrated in Sec. VI, N is *bounded* by such polynomials in t. Therefore, the following gives a *loose* lower bound of the upper bound of the fidelity. Suppose $N \approx f l_o t^{\alpha}$, where the prefactor f and exponent α are approximately constant. Equation (7.17) can be solved for the optimum t:

$$t_{opt} \approx (e^{-\alpha}/\gamma f l_0)^{1/(\alpha-1)} . \tag{7.18}$$

Plugging back into Eq. (7.13) would give an estimate for the optimal achievable fidelity. However, these gross estimates are not expected to be meaningful in actual applications because $N(t)$ will be the determining factor and, as previously mentioned, is analytically unobtainable.

VIII. EXPLICIT CODES

Some explicit codes resulting from our work are presented here. States are given unnormalized when the normalization factor is common for all code words. Codes are specified as $[[N,m,2^k,d]]$, where N is the total number of excitations in the QCSs, m is the number of modes for each QCS, 2^k is the number of code words, and d is the minimal distance between code words. The fidelity of all the codes is given by $\mathcal{F} \approx 1 - \mathcal{C}(N, t+1)\gamma^{t+1}$.

Example (1): $[[4,2,2,2]]$, $n=2$, $t=1$, and fidelity $\mathcal{F} \approx 1 - 6\gamma^2$. For this we have

$$|0_L\rangle = \frac{1}{\sqrt{2}}[|40\rangle + |04\rangle], \tag{8.1}$$

$$|1_L\rangle = |22\rangle. \tag{8.2}$$

Example (2): $[[12,3,10,2]]$, $n=6$, $t=1$, fidelity $\mathcal{F} \approx 1 - 66\gamma^2$, and labels given in hexadecimal ($c=12$, $a=10$). For this

$$|c_1\rangle = \frac{1}{\sqrt{3}}[|00c\rangle + |c00\rangle + |0c0\rangle], \tag{8.3}$$

$$|c_2\rangle = \frac{1}{\sqrt{3}}[|02a\rangle + |a02\rangle + |2a0\rangle], \tag{8.4}$$

$$|c_3\rangle = \frac{1}{\sqrt{3}}[|048\rangle + |804\rangle + |480\rangle], \tag{8.5}$$

$$|c_4\rangle = \frac{1}{\sqrt{3}}[|066\rangle + |606\rangle + |660\rangle], \tag{8.6}$$

$$|c_5\rangle = \frac{1}{\sqrt{3}}[|084\rangle + |408\rangle + |840\rangle], \tag{8.7}$$

$$|c_6\rangle = \frac{1}{\sqrt{3}}[|0a2\rangle + |20a\rangle + |a20\rangle], \tag{8.8}$$

$$|c_7\rangle = \frac{1}{\sqrt{3}}[|228\rangle + |822\rangle + |282\rangle], \tag{8.9}$$

$$|c_8\rangle = \frac{1}{\sqrt{3}}[|246\rangle + |624\rangle + |462\rangle], \tag{8.10}$$

$$|c_9\rangle = \frac{1}{\sqrt{3}}[|264\rangle + |642\rangle + |264\rangle], \tag{8.11}$$

$$|c_a\rangle = |444\rangle. \tag{8.12}$$

Example (3): $[[6,3,4,2]]$, $n=3=0+1+2$, $t=1$, and fidelity $\mathcal{F} \approx 1 - 15\gamma^2$. Here

$$|c_1\rangle = |600\rangle + |060\rangle + |006\rangle, \tag{8.13}$$

$$|c_2\rangle = |420\rangle + |204\rangle + |042\rangle, \tag{8.14}$$

$$|c_3\rangle = |240\rangle + |402\rangle + |024\rangle, \tag{8.15}$$

$$|c_4\rangle = |222\rangle . \tag{8.16}$$

Example (4): $[[9,3,2,3]]$, $n=3$, $t=2$, and fidelity $\mathcal{F} \approx 1 - 84\gamma^3$. Note that this code differs from the previous one from having $d=3$ instead of $d=2$. We take only $|c_2\rangle$ and $|c_3\rangle$ as code words. Thus

$$|0_L\rangle = |306\rangle + |063\rangle + |630\rangle, \tag{8.17}$$

$$|1_L\rangle = |036\rangle + |360\rangle + |603\rangle. \tag{8.18}$$

Example (5): $[[6,4,2,2]]$, $n=6=0+1+2+3$, and fidelity $\mathcal{F} \approx 1 - 15\gamma^2$. The minimal distance between QCSs is $d=2$. However, the QCSs are not generated by multiplying each number by $d=2$. Thus

$$|0_L\rangle = |0321\rangle + |1032\rangle + |2103\rangle + |3210\rangle, \tag{8.19}$$

$$|1_L\rangle = |0123\rangle + |1230\rangle + |2301\rangle + |3012\rangle . \tag{8.20}$$

Example (6): $[[7,2,2,2]]$ and fidelity $\mathcal{F} \approx 1 - 21\gamma^2$. The code words are not formed by cyclic permutations of the QCSs. Note that columns 1 and 2 have different column sums. In this case

$$|0_L\rangle = |70\rangle + |16\rangle, \tag{8.21}$$

$$|1_L\rangle = |52\rangle + |34\rangle. \tag{8.22}$$

Example (7): $[[9,2,2,3]]$ and fidelity $\mathcal{F} \approx 1 - 84\gamma^3$. This is an unbalanced code that will tolerate $t=2$ errors. Note that

one code word is formed from the other by reversing the order of the modes. (This symmetry between the two modes is a sufficient condition for balanced codes with $t=2$, $m\geqslant 3$.) Here

$$|0_L\rangle=\frac{1}{2}|90\rangle+\frac{\sqrt{3}}{2}|36\rangle, \tag{8.23}$$

$$|1_L\rangle=\frac{1}{2}|09\rangle+\frac{\sqrt{3}}{2}|63\rangle. \tag{8.24}$$

Example (8): $[[9,3,2,3]]$ and fidelity $\mathcal{F}\approx 1-84\gamma^3$. This is an unbalanced code that will tolerate $t=2$ errors, showing that the symmetry is not a necessary condition for correcting $t=2$ errors:

$$|0_L\rangle=\frac{1}{\sqrt{3}}[|036\rangle+|306\rangle+|360\rangle], \tag{8.25}$$

$$|1_L\rangle=\frac{1}{3}[\sqrt{6}|333\rangle+\sqrt{2}|009\rangle+|090\rangle]. \tag{8.26}$$

Example (9): $[[16,2,2,4]]$ and fidelity $\mathcal{F}\approx 1-1820\gamma^4$. This is an unbalanced code that will tolerate $t=3$ errors. Labels are given in base 17. c and g denote 12 and 16, respectively. We have

$$|0_L\rangle=\frac{1}{\sqrt{8}}[|0g\rangle+|g0\rangle+\sqrt{6}|88\rangle], \tag{8.27}$$

$$|1_L\rangle=\frac{1}{\sqrt{2}}[|4c\rangle+|c4\rangle]. \tag{8.28}$$

Example (10): $[[20,3,2,4]]$ and fidelity $\mathcal{F}\approx 1-4845\gamma^4$. This is another unbalanced code that will tolerate $t=3$ errors. Labels are given in base 21. c, g, and k denote 12, 16, and 20, respectively. We have

$$|0_L\rangle=\frac{1}{5}[|04g\rangle+2|40g\rangle+2\sqrt{5}|0k0\rangle], \tag{8.29}$$

$$|1_L\rangle=\frac{1}{\sqrt{5}}[\sqrt{2}|44c\rangle+\sqrt{3}|488\rangle]. \tag{8.30}$$

Example (11): $[[50,2,2,5]]$ and fidelity $\mathcal{F}\approx 1-2\,118\,760\gamma^5$. Note the rapid growth in the numerical factor in the second term. To correct for large number of errors, we need to encode a qubit in a large Hilbert space, but emission probabilities are large for high number states. This puts a limit of performance in our codes. The actual code involves numbers five times the numbers shown below. a denotes 10. In this case

$$|0_L\rangle=\sqrt{\frac{1}{18}}|0a\rangle+\sqrt{\frac{5}{9}}|46\rangle \tag{8.31}$$

$$+\sqrt{\frac{1}{3}}|82\rangle+\sqrt{\frac{2}{45}}|91\rangle, \tag{8.32}$$

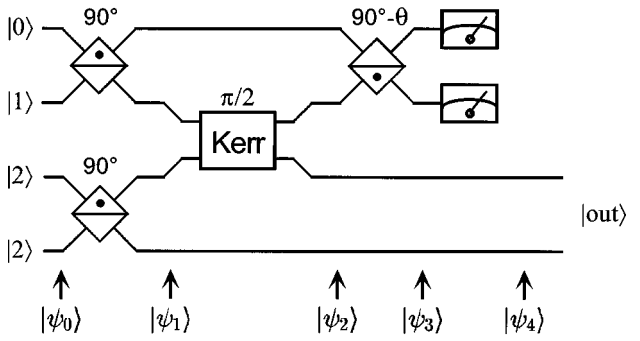

FIG. 1. Quantum circuit to encode a qubit using the code of example (1). As in [21], signals travel from left to right, wires represent optical modes, diamonds represent beam splitters, meters are ideal photon counters, and the Kerr device is an ideal nonlinear optical medium that effects cross-phase modulation of angle $\pi/2$ between single photons.

$$|1_L\rangle=\sqrt{\frac{1}{18}}|19\rangle+\sqrt{\frac{1}{6}}|28\rangle \tag{8.33}$$

$$+\sqrt{\frac{33}{90}}|55\rangle+\sqrt{\frac{1}{3}}|73\rangle+\sqrt{\frac{7}{90}}|a0\rangle\ . \tag{8.34}$$

IX. PHYSICAL IMPLEMENTATION

Encoding and decoding of the codes we have described here can be performed in principle using n-photon eigenstates, beam splitters, phase shifters, and Kerr nonlinear optical media. We demonstrate, for example, how states for our simplest code may be constructed. We then discuss how decoding and correction may be performed.

Shown in Fig. 1 is a quantum circuit that can be used to encode a qubit using the $|22\rangle$, $|04\rangle+|40\rangle$ code. Let us see how this circuit works by tracing the state at the five indicated points. The initial state is $|\psi_0\rangle=|0122\rangle$. Two 50%: 50% beam splitters act on the first and second two modes of this state to give

$$|\psi_1\rangle=\left[\frac{|10\rangle+|01\rangle}{\sqrt{2}}\right]\left[\frac{\sqrt{6}(|04\rangle+|40\rangle-2|22\rangle}{4}\right]. \tag{9.1}$$

Next, a nonlinear optical Kerr medium is used to perform a cross phase modulation between the two middle modes. This serves to ''label'' the $|22\rangle$ state, giving

$$|\psi_2\rangle=\frac{1}{4\sqrt{2}}\{(|10\rangle+|01\rangle)[\sqrt{6}(|04\rangle+|40\rangle)] +(|10\rangle-|01\rangle)[2|22\rangle]\}. \tag{9.2}$$

A final beam splitter in the first two modes now serves to turn the phase modulation into a detectable amplitude difference

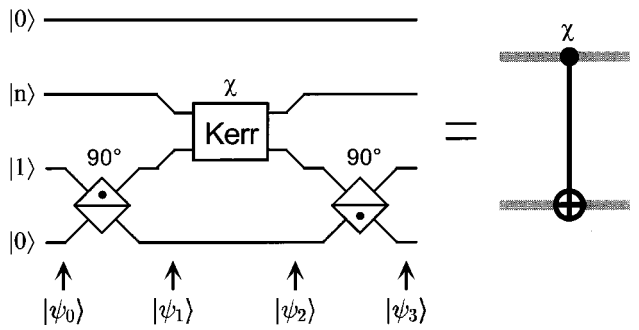

FIG. 2. (Left) Optical quantum logic gate used as a building block in the decoding procedure; example input states are shown on the left. χ is the cross-phase modulation strength of the Kerr medium, which performs the transformation $\exp(i\chi a^{\dagger}ab^{\dagger}b)$ on two modes a and b. (Right) Shorthand notation for this circuit.

$$|\psi_3\rangle=\frac{1}{2\sqrt{2}}\{|01\rangle[\sqrt{3}\sin\theta(|04\rangle+|40\rangle)+\sqrt{2}\cos\theta|22\rangle] +|10\rangle[\sqrt{3}\cos\theta(|04\rangle+|40\rangle)-\sqrt{2}\sin\theta|22\rangle]\} \quad (9.3)$$

such that if the first two modes are measured to be $|01\rangle$ (otherwise, the state is discarded) then we have the output

$$|\psi_4\rangle=\cos\theta'|22\rangle+\sin\theta'\,\frac{|04\rangle+|40\rangle}{\sqrt{2}}\,, \quad (9.4)$$

where

$$\theta'=\tan^{-1}\left[\sqrt{\frac{1}{3}}\tan\theta\right] \quad (9.5)$$

is the new effective angle. $|\psi_4\rangle$ is the desired encoded state. After transmission of $|\psi_4\rangle$ through a lossy communication link, the final state can then be measured immediately (using, for example, photon number counters) and the transmitted qubit collapsed. This would be the standard procedure for point-to-point quantum cryptography.

Alternatively, the qubit may be relayed by performing an error correction step. This involves calculation of the error syndrome, correcting any detected error, and then re-encoding the state for further transmission. In this system, we have a nonbinary state. If it can be turned into a binary state, then the entire correction procedure may be performed using standard techniques of quantum computation, with the usual binary quantum logic gates [19]. This is possible as follows.

Consider the circuit shown in Fig. 2. The structure is identical to the well-known quantum-optical Fredkin gate [20,21], but let us think of it here in a different way. The lower pair of wires may be considered to be a single ''dual-rail'' qubit [22], with logical states $|0_L\rangle=|10\rangle$ and $|1_L\rangle=|01\rangle$. With an input state $|\psi_0\rangle=|0n\rangle|0_L\rangle$ and a Kerr medium with $\chi=\pi/n$, then the output state will be $|\psi_3\rangle=|0n\rangle|1_L\rangle$, and in this manner the circuit may be thought of as a kind of *controlled-not gate* that distinguishes between a control state of $|0\rangle$ and $|n\rangle$.

We now use this bosonic controlled-not gate to construct the circuit of Fig. 3, which is based on the fact that different values of χ allow us to distinguish different values of n.

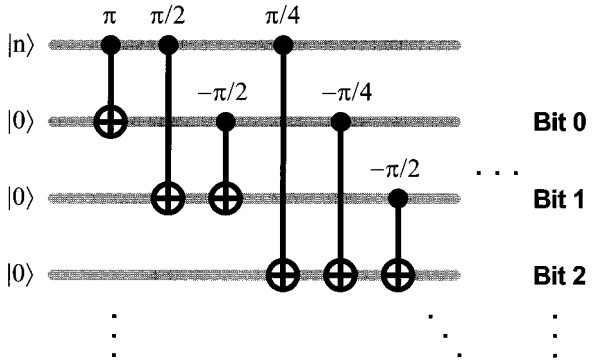

FIG. 3. Quantum circuit used to decode a bosonic state into qubits. Each thick wire represents a pair of bosonic modes. The top wire carries the bosonic state and the remaining wires are prepared as dual-rail quantum bits, which carry just one photon in each pair.

Furthermore, since the decomposition of the number n into sums of powers of 2 is unique, it is convenient to take χ to be binary fractions of π. In this manner, the first dual-rail qubit becomes the least-significant bit of n and so on. If the qubits are measured, this circuit would then function equivalently to a perfect photon number detector.

However, it is much more useful in that calculations may be performed based on the binary representation of n to determine if any error occurred and to calculate the error syndrome. The circuit is then applied in reverse to undo the entanglement with the bosonic state and the appropriate correction procedure is applied to fix the detected error. This is possible since generalized measurements may be performed on the qubit states to determine the error syndrome without destroying the superposition state of the original qubit encoded in the bosonic state.

X. CONCLUSION

Our treatment of amplitude damping errors is somewhat unusual from the standpoint of most quantum error correction theories, which deal with a bit flip and phase flip picture of errors. The relationship can be understood by expressing the A_0 and A_1 operators as coherent superpositions of such errors; from Eqs. (2.16) and (2.17),

$$A_0=\frac{1}{2}[(1+\sqrt{1-\gamma})I+(1-\sqrt{1-\gamma})\sigma_z], \quad (10.1)$$

$$A_1=\frac{\sqrt{\gamma}}{2}[\sigma_x+\sigma_y]\,. \quad (10.2)$$

With probabilities up to $O(\gamma)$, a binary code with m bits will either project $A_0^{\otimes m}|c_l\rangle$ onto a state with no errors or project $A_0^{\otimes m-1}A_1|c_l\rangle$ onto a state with only one bit flip error, resulting from a combination of $I^{\otimes m-1}$ and one of the Pauli operators σ_x,σ_y. Hence a binary code correcting for any one-bit error *will* indeed correct all amplitude damping errors up to losing one photon, although not to all orders. One reason we have studied bosonic codes is to exploit the possibilities for achieving higher efficiencies or easier physical implementation, though the study is theoretically interesting on its own.

It is important to realize that amplitude damping errors are not independent bit errors since the decay factor of each QCS depends on the total number of excitations in it. This fact is also pointed out by Plenio *et al.* [6]. Any code correcting a general one-bit error can do so in the presence of possible amplitude damping and is capable of correcting any one-bit error in the projected space in which no loss occurs as well as correcting a single photon loss in the absence of any other errors. Two or more amplitude damping errors projected into the nonidentity space are viewed as separate errors. The general relation between binary codes and amplitude damping is still under investigation.

Nevertheless, some interesting comparisons may be made. Rates from our bosonic codes contrast with those achievable by the usual binary codes. For the code of example (1), $nd=4$, $m=2$, and $k=1$, so the rate is found to be $r=0.22$. This is slightly better than $r=0.20$ for the five-bit $(t,k)=(1,1)$ binary perfect code [23], and much better than the eight-bit (1,1) code of Plenio *et al.* [6], which corrects errors in the presence of A_k's in $\mathcal{K}(0)$. Similarly, for the code of example (2), $d=2$, $n=6$, $m=3$, and $2^k=10$ code words may be found, giving a rate $r=0.2994$. In comparison, a naive evaluation of the quantum Hamming bound [17] for binary codes gives a possible rate of 0.41. Nondeformation constraints are more restrictive on bosonic error correction codes than on binary codes, but the bosonic states admit coding schemes that are impossible with binary codes. There is no conclusive statement on comparing the general efficiencies of the two different type of codes, but the examples we have discovered indicate the existence of a rich variety of bosonic codes that may be useful in the future.

The code fidelities may also be compared. Our [[4,2,2,2]] code achieves $\mathcal{F}\approx 1-6\gamma^2$. In comparison, from an explicit evaluation of the effect of amplitude damping on all qubits, we have found that the five-bit (1,1) binary code achieves fidelity $\approx 1-1.75\gamma^2$, while the eight-bit (1,1) code achieves only $\approx 1-6\gamma^2$. This agreement with the bosonic code is not accidental; it stems from the use of the same total excitation number. However, it is worthwhile to point out that despite the effort to balance the code words, the five-bit code still has better performance on average due to the small number of excitation involved in the system.

In conclusion, we have given general criteria for an error correction code that encodes qubits in bosonic states. This is a generalization of the binary error correction codes. Motivated by the dominant decoherence process (amplitude damping) of a system such as photons transmitted through optical fibers, we classify our errors according to the number of excitations lost instead of the more common classification of the number of bits or modes corrupted. We have shown, in one case, that specialization to correction amplitude damping does improve the ratio of the number of encoded qubits to the number of required qubits. However, bosonic codes under amplitude damping suffer constraints involving deformation of the Hilbert space not shared by the binary codes, rendering the efficiencies lower in the bosonic case when many qubits are encoded.

It is too early to conclude on the relative performance of the binary codes and the bosonic codes. Further study will aim at improving the efficiencies, perhaps by using relative phases to maintain orthogonality of the code words instead of by using distinct QCSs, so that the QCSs can occur in more than one code word. Another possibility is to encode qubits using QCSs of a different total number of excitations. We hope that further study of bosonic codes will lead to their practical utilization in addition to the current theoretical interest.

ACKNOWLEDGMENTS

We would like to thank John Preskill for particularly useful discussions. D.W.L. was supported in part by the Army Research Office under Grant No. DAAH04-96-1-0299. I.L.C acknowledges financial support from the Fannie and John Hertz Foundation.

APPENDIX: CRITERIA FOR NONDEFORMATION OF HILBERT SPACE

Consider $\mathcal{K}(2)=\{0\cdots02,0\cdots20,\ldots,20\cdots0\}\cup\{0\cdots011,0\cdots110,\cdots,110\cdots0\}$ for $t=2$ errors. For instance,

$$A_{0\cdots02}|c_l\rangle \tag{A1}$$

$$=A_{0\cdots02}\frac{1}{\sqrt{N_l}}\sum_{i=1}^{m}|n_{i1}\cdots n_{im}\rangle \tag{A2}$$

$$=\sum_{i=1}^{m}\sqrt{\frac{\mathcal{C}(n_{im},2)\,\gamma^2(1-\gamma)^{N-2}}{N_l}}|n_{i1}\cdots n_{im}-2\rangle\ , \tag{A3}$$

where $\mathcal{C}(n,m)$ is the usual binomial coefficient. The norm square of this state is

$$\langle c_l|A^{\dagger}_{0\cdots02}A_{0\cdots02}|c_l\rangle=\frac{\gamma^2(1-\gamma)^{N-2}}{2N_l}\sum_{i=1}^{m}n_{im}(n_{im}-1)\ . \tag{A4}$$

The term linear in γ is independent of l (code-word independent) by the criteria for $t=1$; hence it follows that

$$\frac{1}{N_l}\sum_{i=1}^{m}n_{im}^2 \tag{A5}$$

has to be independent of l if the nondeformation criterion is to be satisfied. Other $A_{\tilde{k}}$ with $\tilde{k}=0\cdots02,\ldots,20\cdots0$ impose the above requirement on other columns.

Similarly, $\tilde{k}=0\cdots11$ changes the code word to

$$\sum_{i=1}^{m}\sqrt{\frac{n_{im-1}n_{im}\gamma^2(1-\gamma)^{N-2}}{N_l}}|n_{i1}\cdots n_{im-1}-1n_{im}-1\rangle\ , \tag{A6}$$

which has norm square

$$\frac{\gamma^2(1-\gamma)^{N-2}}{N_l}\sum_{i=1}^{m} n_{im-1}n_{im}. \tag{A7}$$

Equation (4.2) requires the following to be independent of l:

$$\frac{1}{N_l}\sum_{i=1}^{m} n_{im-1}n_{im}. \tag{A8}$$

A similar result is obtained for other $\tilde{k}$ with 1's at any two modes j_1 and j_2. When we allow $j_1=j_2$, we include the previous result for two photon loss at one mode. This proves Lemma 2.

For arbitrary t, we get equations involving various binomial coefficients. Using requirements involving products of fewer than t n_{ij}'s, we can replace the products of the binomial coefficients to products involving exactly t n_{ij}. By mathematical induction, the result for arbitrary t is then obtained. This completes the proof of Theorem 1.

[1] W. K. Wooters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[2] P. Shor, Phys. Rev. A **52**, 2493 (1995).
[3] A. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[4] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[5] A. Steane, Oxford University Report, quant-ph/9608026, 1996 (unpublished).
[6] M. B. Plenio, V. Vedral, and P. L. Knight, Phys. Rev. A **55**, 67 (1997).
[7] K. Kraus, *States, Effects, and Operations* (Springer-Verlag, Berlin, 1983).
[8] W. H. Louisell, *Quantum Statistical Properties of Radiation* (Wiley, New York, 1973).
[9] C. W. Gardiner, *Quantum Noise* (Springer-Verlag, New York, 1991).
[10] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[11] M. Gell-Mann and J. Hartle, Phys. Rev. D **47**, 3346 (1993).
[12] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[13] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).
[14] M. A. Nielsen, B. W. Schumacher, C. M. Caves, and H. Barnum (private communication).
[15] A. Ekert and C. Macchiavello, Phys. Rev. Lett. **77**, 2585 (1996).
[16] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, IBM Report, quant-ph/9604024, 1996 (unpublished).
[17] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).
[18] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers, 4th Ed.* (Oxford University Press, London, 1960).
[19] A. Barenco *et al.*, Phys. Rev. A **52**, 3457 (1995).
[20] Y. Yamamoto, M. Kitagawa, and K. Igeta, in *Proceedings of the Third Asia-Pacific Physics Conference*, edited by Y. W. Chan, A. F. Leung, C. N. Yang, and K. Young (World Scientific, Singapore, 1988).
[21] G. J. Milburn, Phys. Rev. Lett. **62**, 2124 (1989).
[22] I. L. Chuang and Y. Yamamoto, Phys. Rev. A **52**, 3489 (1995).
[23] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).