CS191 – Fall 2014
Lecture 16: Error suppression and prevention techniques

Mohan Sarovar
(Dated: October 23, 2014)

I. RECAP: OPEN QUANTUM SYSTEMS

Errors and noise in quantum computing arise from the coupling of our computing system (*e.g.*, a collection of qubits) to an "environment" that we cannot fully control or observe. This coupling leads to open quantum system dynamics and in the last couple of lectures you've seen three different ways in which we model such dynamics:

1. Specify the Hamiltonian for the entire setup including system, environment and the way in which they are coupled:

$$H_{SE} \;=\; H_S + H_E + H_I = h_0^S \otimes I^E + I^S \otimes h_0^E + \sum_{k=1}^{K} h_k^S \otimes h_k^E$$

Sometimes this representation is approximated by a classical fluctuating model for the environment:

$$H_S(t) \;=\; H_S + H_{\text{noise}} = h_0^S + \sum_{k=1}^{K} \lambda_k(t) h_k^S$$

This approach requires explicit modeling of the environmental degrees of freedom, and hence we call it the *extended Hamiltonian formulation* of open system dynamics.

2. An *operator sum representation* or *Kraus map* description of the system dynamics:

$$\rho = \mathcal{E}(\rho_0) = \sum_{k=1}^{K} A_k \rho A_k^\dagger,$$

with $\sum_k A_k^\dagger A_k = I$.

3. A *Lindblad master equation* describing the time evolution of the system, including the effects of environmental coupling:

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[h_0 + h_{LS}, \rho(t)] + \sum_{k=1}^{K} \gamma_k \left(L_k \rho(t) L_k^\dagger - \frac{1}{2} L_k^\dagger L_k \rho(t) - \frac{1}{2}\rho(t) L_k^\dagger L_k \right),$$

While the first two models were fully general, this master equation formulation is valid only for systems weakly coupled to Markovian environments (recall the two critical approximations that went into deriving this evolution equation).

II. ERROR CONTROL

Now we will examine ways in which to mitigate the effects of open system dynamics and the errors that it can generate. We can categorize mitigation strategies into two types: (i) passive, error prevention (EP) strategies, and (ii) active error correction (EC) strategies. EP strategies typically require more knowledge about the system and environment (*e.g.*, an accurate model of the system-environment interaction Hamiltonian), while EC strategies are more generally applicable but require greater resources to implement. Examples of EP strategies are *decoherence free subspaces* (DFS), and *dynamical decoupling* (DD), while the prime example of an EC strategy is *fault-tolerant quantum error correction.* The modern view is that some combination of both will be required to develop a quantum computer, with the EP strategies operating at the physical layer (*e.g.*, hardware or firmware level) while the EC strategies operate at the logical layer (*e.g.*, software level).

The above categorization is a little coarse in the sense that there exist techniques that incorporate ideas that fall into both categories, but we will not cover such advanced techniques here. In this lecture we will go over the two main EP strategies: decoherence free subspaces and dynamical decoupling, while the following lectures will look at error correction.

III. DECOHERENCE FREE SUBSPACES

A decoherence free subspace (DFS) is a subspace of system Hilbert space that satisfies certain symmetries with respect to the system-environment interaction. These symmetries mean that a DFS is a "quiet" region of Hilbert space that is unaffected by noise.

Let's begin by looking at an example. Assume the relevant error process is the phase-flip channel

$$\rho = \mathcal{E}_{\rm pf}(\rho_0) = (1-p)\rho_0 + p\sigma_z\rho_0\sigma_z, \tag{1}$$

for $0 \leq p \leq 1$. We have seen that a general state of one qubit will be dephased by this channel. That is,

$$\alpha\left|0\right\rangle + \beta\left|1\right\rangle \to |\alpha|^2\left|0\right\rangle\left\langle 0\right| + |\beta|^2\left|1\right\rangle\left\langle 1\right|, \tag{2}$$

under the action of this channel. Now, assume the physical noise that causes this dephasing is homogenous in space, meaning that if we have two qubits the error process is collective:

$$\rho = \mathcal{E}_{\rm coll}(\rho_0) = (1-p)\rho_0 + p\big(\sigma_z\otimes\sigma_z\big)\rho_0\big(\sigma_z\otimes\sigma_z\big),$$

where ρ and ρ_0 are two qubit density matrices now.

> **Exercise:** This homogenous (or collective) dephasing error process is different from just applying $\mathcal{E}_{\rm fp}$ to each of the two qubits. Show this by computing the Kraus decomposition of $\mathcal{E}_{\rm pf}\otimes\mathcal{E}_{\rm pf}$.

Now note that the error operator in this process satisfies the following relations:

$$\begin{aligned} \sigma_z\otimes\sigma_z\left|00\right\rangle &= \left|00\right\rangle \\ \sigma_z\otimes\sigma_z\left|11\right\rangle &= \left|11\right\rangle \end{aligned}$$

Therefore, the error operator does the same thing (which happens to be nothing in this case) to both orthogonal states. This motivates us to encode the orthogonal basis states of a single qubit using these two special states of two qubits. That is,

$$\begin{aligned} \left|0_L\right\rangle &= \left|00\right\rangle \\ \left|1_L\right\rangle &= \left|11\right\rangle, \end{aligned}$$

where the subscript L indicates a logical state. Then an arbitrary single qubit state is encoded as $\left|\psi\right\rangle = \alpha\left|0_L\right\rangle + \beta\left|1_L\right\rangle = \alpha\left|00\right\rangle + \beta\left|11\right\rangle$. Letting the initial density matrix be $\rho_0 = \left|\psi\right\rangle\left\langle\psi\right|$, we can verify that

$$\mathcal{E}_{\rm coll}(\rho_0) = \rho_0, \tag{3}$$

which means that any state in this two-dimensional subspace of the (four dimensional) two qubit Hilbert space is unaffected by this noise process. For this reason, we refer to this as a *decoherence free subspace*, or DFS.

Some DFS properties that can already be seen from this example are:

1. We exploited symmetry in the noise process or open system dynamics in order to define the DFS (in this example the symmetry is that the error operate homogeneously). All DFS constructions require such symmetry in the noise process. This may seem restrictive but in fact there are physical situations of interest where such symmetries are present. An example is magnetic field noise in quantum dot spin qubits. The spatial variation of fluctuating magnetic fields is typically larger than the separation between quantum dots, and therefore the magnetic field noise seen by both qubits is almost the same.

2. To exploit symmetries in the system's open system dynamics, we have to have significant knowledge about the form of the noise (*e.g.*, the Kraus decomposition for the process) or the system-environment interaction.

3. DFSs utilize encoded states, which comes with two associated costs: (i) logical qubits are encoded in multiple physical qubits, and (ii) logical gates can be complicated since the logical states are distributed across multiple qubits. To see an example of the latter cost, notice that for the example above, the logical σ_x gate is:

$$X_L\left|0_L\right\rangle = \left|1_L\right\rangle,$$

with $X_L = \sigma_x\otimes\sigma_x$.

Now let us go beyond the above example and look at DFSs more formally. First, a formal definition.

> ***Definition:*** [From `quant-ph/0301032`] A system with Hilbert space $\mathcal{H}$ is said to have a decoherence free subspace $\bar{\mathcal{H}}$, if the evolution inside $\bar{\mathcal{H}} \subset \mathcal{H}$ is purely unitary.

That is, a DFS is spanned by a set of states $\left|\bar{k}\right\rangle \in \bar{\mathcal{H}}$ that evolve unitarily despite the open system evolution of the system as a whole. Note that we are not requiring that the states do not evolve at all, that would be a stricter requirement. Physically, the reason unitary evolution is allowed is because such evolution is reversible, and therefore even if the states within the DFS were evolved unitarily by the system-environment coupling, we would be able to reverse this evolution. In the above example, $\left|\bar{k}_0\right\rangle = |00\rangle$ and $\left|\bar{k}_1\right\rangle = |11\rangle$, and the unitary evolution within the subspace spanned by these states that is induced by the collective dephasing process is just the identity (*i.e.*, no evolution).

Given a model of open system dynamics, an important task is to identify whether any DFS exists for the dynamics. If so, then one can think of encoding in it to gain robustness against the noise. There are criteria for identifying DFSs when the open system dynamics is specified in any of the ways outlined in section I. We will go over the criterion when this specification is given as an OSR, but refer to `quant-ph/0301032` (see References) for a full list of criteria.

Suppose the system evolves in Hilbert space $\mathcal{H}$ (of dimension d) and the open system dynamics is specified as an OSR (*i.e.*, the environment has been traced out):

$$\mathcal{E}(\rho_0) = \sum_{j=1}^{K} A_j \rho_0 A_j^\dagger$$

Suppose a subspace $\bar{\mathcal{H}} \subset \mathcal{H}$ is spanned by the n states $\left|\bar{k}_i\right\rangle$, and its orthogonal complement is spanned by $d-n$ states $|k_i\rangle$, and further we write each Kraus operator in a basis $\left|\bar{k}_1\right\rangle, \left|\bar{k}_2\right\rangle, \dots \left|\bar{k}_n\right\rangle, |k_1\rangle, |k_2\rangle, \dots |k_{d-n}\rangle$. Then $\bar{\mathcal{H}}$ is a DFS if and only if all the Kraus operators in this basis have the form:

$$A_j = \begin{pmatrix} g_j \bar{U} & 0 \\ 0 & \tilde{A}_j \end{pmatrix}, \tag{4}$$

for $\bar{U}$, some unitary operator on $\bar{\mathcal{H}}$, and scalar g_j. This means that each Kraus operator when restricted to the subspace $\bar{\mathcal{H}}$ acts like a unitary, and that this unitary is the same for all Kraus operators. The $\tilde{A}_j$ are arbitrary operators on the Hilbert subspace orthogonal to $\bar{\mathcal{H}}$. Since $\sum_j A_j^\dagger A_j = I$, we have

$$\sum_j A_j^\dagger A_j = \sum_j \begin{pmatrix} |g_j|^2 \bar{U}^\dagger \bar{U} & 0 \\ 0 & \tilde{A}_j^\dagger \tilde{A}_j \end{pmatrix} = I \quad \Rightarrow \sum_j |g_j|^2 = 1, \quad \sum_j \tilde{A}_j^\dagger \tilde{A}_j = \tilde{I}. \tag{5}$$

To see why this is a sufficient condition for a DFS, consider a general state in this subspace:

$$|\psi\rangle = \sum_i \alpha_i \left|\bar{k}_i\right\rangle. \tag{6}$$

The action of a Kraus operator on this state is

$$A_j |\psi\rangle = \sum_i \alpha_i g_j \bar{U} \left|\bar{k}_i\right\rangle = g_j \bar{U} |\psi\rangle. \tag{7}$$

Therefore the action of the whole OSR map is

$$\mathcal{E}(|\psi\rangle\langle\psi|) = \sum_j |g_j|^2 \bar{U} |\psi\rangle\langle\psi| \bar{U}^\dagger = \bar{U} |\psi\rangle\langle\psi| \bar{U}^\dagger, \tag{8}$$

and hence any state in this subspace evolves unitarily under the OSR, which is exactly how we defined a DFS.

> *Example 1* Let us see how the example we examined above fits this criterion. The Kraus operators for the two qubit collective dephasing channel are

$$\begin{aligned} A_0 &= \sqrt{1-p} I_4 = \sqrt{1-p} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ A_1 &= \sqrt{p} \sigma_z \otimes \sigma_z = \sqrt{p} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

These Kraus operators are written in the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. But we know the DFS, $\bar{\mathcal{H}}$ is spanned by $|00\rangle$ and $|11\rangle$. So reordering the above basis to put these elements first, and rewriting the Kraus operators yields

$$\begin{aligned} A_0 &= \sqrt{1-p}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ A_1 &= \sqrt{p}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{aligned}$$

So we see that $\bar{U} = I_2$ in this case and $g_0 = \sqrt{1-p}, g_1 = \sqrt{p}$. Note that the subspace spanned by $|01\rangle$ and $|10\rangle$ is also a DFS with $\bar{U} = I_2$ again, and $g_0 = \sqrt{1-p}, g_1 = -\sqrt{p}$ in this case.

As mentioned above, one can also define conditions for the existence of DFSs in the extended Hamiltonian formulation of open system dynamics (these become conditions on the system-environment interaction Hamiltonian), and in the Lindlbad master equation formulation of open system dynamics (these become conditions on the Lindblad operators).

IV. DYNAMICAL DECOUPLING

Another common error prevention strategy is called dynamical decoupling (DD), which works by applying appropriately designed pulses on a system in order to average out its interaction with the environment. Dynamical decoupling has a long history, starting from the *spin echo* technique developed in the early days of nuclear magnetic resonance (NMR). In fact, the physical resonance communities (NMR, ESR, etc.) have used dynamical decoupling as a technique to suppress noise for a long time, and since the late 1990s these methods have also found application in quantum information.

Again, to appreciate the basic idea behind DD we will begin with an example. Suppose we return to the scenario examined in section II.A of lecture 14, where we had an uncertain single qubit Hamiltonian

$$H = \frac{\omega}{2}\sigma_z, \tag{9}$$

where ω is fixed but random and unknown. There we saw that if the system evolves according to this Hamiltonian for a period of time, any superposition of the computational basis states will dephase due to the uncertainty in ω feeding into an uncertainty in the phase relationship between these basis states. However, consider the following evolution

$$U = \sigma_x e^{-iH\Delta t}\sigma_x e^{-iH\Delta t}. \tag{10}$$

That is, we let the system evolve for a time Δt with the above uncertain Hamiltonian, then we hit it with a very short pulse $e^{-i\frac{\pi}{2}\sigma_x} = \sigma_x$, then we let it evolve for another interval Δt, before hitting it with another short pulse of the same form as before. Notice that we are still evolving under the uncertain Hamiltonian, but we are simply interrupting the evolution occasionally. To understand what the net evolution is, let's simplify this unitary. We will use the following useful identity to do so:

$$Ae^B A^{-1} = e^{ABA^{-1}}, \tag{11}$$

for any invertible matrix A.

Exercise: Prove this identity.

Using this identity, with $A = \sigma_x$, we get

$$\begin{aligned} U &= e^{-i\frac{\omega}{2}\Delta t(\sigma_x\sigma_z\sigma_x)}e^{-i\frac{\omega}{2}\Delta t\sigma_z} \\ &= e^{+i\frac{\omega}{2}\Delta t\sigma_z}e^{-i\frac{\omega}{2}\Delta t\sigma_z} = I, \end{aligned} \tag{12}$$

where we have used the fact $\sigma_x\sigma_z\sigma_x = -\sigma_z$. So the net evolution is the identity (no evolution)! We have "averaged away" the random rotation by reversing its direction and evolving for an equal time in the forward and backward directions. This averaging is the idea behind DD, and now let us examine the structure behind DD more generally.

Let the extended Hamiltonian for the system and environment be given by

$$H = H_E + H_{SE}, \tag{13}$$

where we drop the system component of the Hamiltonian, H_S, since we think of this as absent or we are in the interaction frame with respect to this term. Then consider evolution by this Hamiltonian (that couples the system to the environment) by small times Δt, that is interleaved by unitary pulses P_i and their inverses $P_i^\dagger$:

$$U = P_n \exp(-i\Delta t(H_B + H_{SE}))P_n^\dagger ... P_1 \exp(-i\Delta t(H_B + H_{SE}))P_1^\dagger \tag{14}$$

Then using the identity in Eq. (11) again, we get

$$U = \exp(-i\Delta t(H_B + P_n H_{SE} P_n^\dagger))... \exp(-i\Delta t(H_B + P_1 H_{SE} P_1^\dagger)) \tag{15}$$

Notice that the pulses do not modify the H_E since they are applied on the system only. Now, we wish to combine these exponentials into one exponential. This is difficult to do since the exponents do not all commute with each other. The *Baker-Cambell-Hausdorff* (BCH) formula tells us how to combine matrix exponentials:

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\frac{1}{12}([A,[A,B]]+[B,[B,A]]+...)}. \tag{16}$$

The exponent of the right hand side is an infinite series, however, if Δt is small enough that $||nH\Delta t|| \ll 1$, then we can ignore the extra terms that come from the lack of commutation and simply keep the leading order $A + B$ term when combining each of the exponents in Eq. (15). Under this approximation,

$$U \approx \exp\left[-i(n\Delta t H_B + \Delta t \sum_{i=1}^{n} P_i H_{SE} P_i^\dagger\right] \tag{17}$$

Thus we see that the system-environment Hamiltonian has transformed as a result of the pulses as:

$$H_{SE} \rightarrow \frac{1}{n}\sum_{i=1}^{n} P_i H_{SE} P_i^\dagger \tag{18}$$

The goal of DD is to choose the P_i such that this transformed system-environment Hamiltonian is as close to zero in as few pulses (as small n) as possible. There is considerable literature on DD pulse sequences that do this in different situations and for different H_{SE}. Note that there is some approximation in our treatment above since we have only taken the first order terms when combining the exponentials in Eq. (15) to get Eq. (17). However, we can relax this approximation by taking higher and higher order terms in the the BCH expansion, and this is typically done in more sophisticated DD sequences. In fact, one typically uses a different expansion, called the *Magnus expansion*, to incorporate the higher order terms since it has better convergence properties than the BCH expansion. See the References and further reading section for more literature on constructing DD sequences and the *effective Hamiltonian theory* that is formulated from the Magnus expansion.

Now, let us make some general remarks about the DD technique:

1. As with DFSs we require some knowledge of the system-environment interaction (or noisy Hamiltonian). In the above example we had to know the Hamiltonian was σ_z in order to know to apply σ_x pulses.

2. The pulse duration must be very fast on the time scales of the system dynamics. Intuitively the "change of direction" induced by the pulses must be quick.

3. The rate at which pulses are applied must be faster than the timescales set by the noise (or environment). For example, suppose instead of a static random ω in Eq. (9), we had a time dependent $\omega(t)$. Then if Δt were small enough such that $\omega(t)$ did not change appreciably between σ_x pulses then we would still get approximate cancelation above leading to $U \approx I$. However, if $\omega(t)$ changed appreciably in Δt it is easy to see that the cancelation would not occur and $U \neq I$. This is generally true for DD; it is only effective against noise of frequency less than $1/\Delta t$, where Δt is the average time between pulses.

I note that more advanced DD techniques can relax some of the above requirements. For example, finite duration pulses can be taken into account, but they must still be relatively fast compared to the system time scales. Also, full knowledge of the system-environment interaction (or noisy Hamiltonian) is not necessary if more complicated DD pulses (*e.g.*, universal decoupling pulses) are used.

Example 2 Suppose the system-environment Hamiltonian is

$$H_{SE} = \sigma_x^S \otimes B_x^E + \sigma_z^S \otimes B_z^E, \tag{19}$$

for some operators B_x^E and B_z^E on the environment. This environmental coupling can lead to a combination of σ_x errors and σ_z errors on the system. However, suppose we perform a length $n = 4$ DD sequence: $P_1 = \sigma_x, P_2 = \sigma_z, P_3 = \sigma_x, P_4 = \sigma_z$. What is the resulting first order effective system-environment coupling Hamiltonian when this pulse sequence is repeatedly executed?

Referring back to the first-order effective Hamiltonian derived above, Eq. (18), we get:

$$\begin{aligned} H_{SE} &\rightarrow \frac{1}{4}\sum_{i=1}^{4} P_i(\sigma_x^S \otimes B_x^E + \sigma_z^S \otimes B_z^E)P_i^\dagger \\ &= \frac{1}{4}\Big[(\sigma_x^S \otimes B_x^E - \sigma_z^S \otimes B_z^E) + (-\sigma_x^S \otimes B_x^E + \sigma_z^S \otimes B_z^E) + (\sigma_x^S \otimes B_x^E - \sigma_z^S \otimes B_z^E) + (-\sigma_x^S \otimes B_x^E + \sigma_z^S \otimes B_z^E)\Big] \\ &= 0 \end{aligned}$$

The σ_x pulses reverse the sign of the $\sigma_z^S \otimes B_z^E$ term in the system-environment interaction and the σ_z pulse reverse the sign of the $\sigma_x^S \otimes B_x^E$ term. Therefore after a $n = 4$ pulse cycle we have completely nulled out the system environment interaction.

V. REFERENCES AND FURTHER READING

1. A nice review of decoherence free subspaces is presented in *Decoherence-free subspaces and subsystems* by D. A. Lidar and K. B. Whaley.
`http://arxiv.org/abs/quant-ph/0301032`.

2. Another slightly more technical review of DFSs is *Review of Decoherence Free Subspaces, Noiseless Subsystems, and Dynamical Decoupling* by D. A. Lidar.
`http://arxiv.org/abs/1208.5791`. This paper also has a technical review of dynamical decoupling.

3. A nice physics-based introduction to DD is in these Stanford lecture notes:
`http://web.stanford.edu/~rsasaki/AP227/chap4.pdf`.